SIGNAL RECOVERY

Signals, sensors and noise

Ivan Rech
Giulia Acconcia

SIGNAL RECOVERY

Ivan Rech, Giulia Acconcia

To Sergio Cova, creator of the Signal Recovery course, without whose work this book would not have existed.

INDEX

Introduction

This book was conceived as a didactic support for the course of Signal Recovery held by the authors Ivan Rech and Giulia Acconcia at Politecnico di Milano. The idea is to provide the students with organized material where they can find what was explained in class; examples, digressions, links with real applications described in class will not be reported in this text.

Signal Recovery

The title of the course to which this book refers, and the title of this text, is Signal Recovery. This course is nothing more than the evolution of the "Instrumentation and electronic measurements" course held at the Politecnico di Milano by Prof. Sergio Cova for many years. This book is based on the work of Prof. Sergio Cova to create a course and teaching material that would allow the student to fully understand the concept of signal recovery thanks to the collection of material that is difficult to recover in an organized form. Indeed, the final aim of this course is to understand how to extract the signal information when it is partially or completely buried into noise. To this aim, we will first study how to deal with signals, we will recall or introduce some key mathematical tools that will allow us to perform a deep analysis of signals first, and then we will do the same for noise.

The course is not intended to just lead the students to know and properly describe the electronic techniques and instrumentation developed for recovering signals from noise, but it rather aims at giving a good insight into problems and the approaches developed to solve them. Students need to develop critical thinking starting from a critical evaluation of the presented solutions and avoiding the attitude where sensors and electronics are designed and employed just according to some established rule or standard. It is instead necessary to clarify the reasons of any choice and decision made in the design of the acquisition chain, on the basis of the physics of phenomena involved, of the basic principles of signal and noise processing and of the actual performance of the available devices. It is necessary to clearly distinguish between the intrinsic limitations set by physical laws and the current limitations set by the state of the art, which can be overcome by technological progress. In essence, gaining insight means to progress at the pace of the technology evolution and be able to contribute to it.

Description of signals

The goal of this course is to develop a method for recovering signals, especially under critical circumstances. To this aim, the exploitation of any information that is associated only with the signal, making it different from noise and disturbances, is of the utmost importance. In this chapter we will present some key tools to describe signals and their properties.

2.1 Signals

In the most general case, a deterministic signal is a function that assigns to a variable a certain value for each time t. Intuitively, it is something that varies over time and it carries a certain amount of information. In principle, there are no restrictions on the nature of the function that describes the signal, but, in practice, a signal typically comes from a measure made in the real world, for example an electrical variable produced by a sensor. In figure 2.1 the simplified scheme of a typical chain used to acquire and measure a signal is depicted.
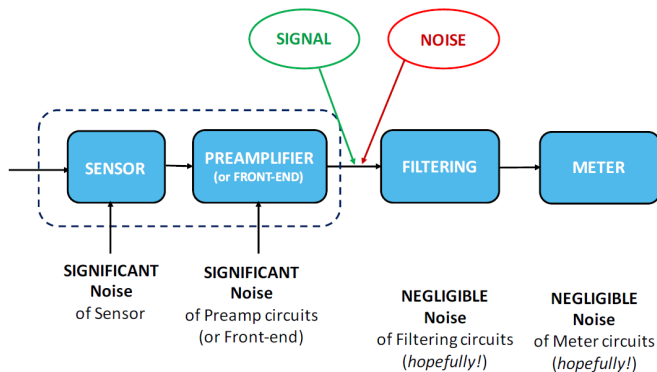
Figure 2.1 Simplified scheme of a typical acquisition chain

First of all, we need a sensor to translate a physical quantity of interest (e.g. temperature, light, etc.) into an electrical quantity. Then, we typically have some blocks for signal conditioning and recording. In particular, we usually exploit

amplifiers to acquire and amplify the signal, filtering stages to select the signal of interest, and converters to translate the signal into digital words.

As we will see in this course, one of the main issues in signal acquisition is the presence of the so-called noise. Noise is like a statistical signal that can be both related to the acquisition chain and inherently associated with the signal of interest. In both cases its existence can be justified by observing that, if we consider an arbitrary physical process, there will always be some fluctuations related to it. For instance, we will see that random thermal motion of electrons into a resistor causes a fluctuation of the voltage across it, or the statistics of photon arrivals can cause a fluctuation of the intensity in a ray of light. Both these phenomena are statistical and do not carry useful information: therefore, their output is noise.

The presence of noise corrupts in some way the information contained into the signal, so we need to use a particular class of blocks, called filters, that are designed to make the most of the differences between the signal and the noise, with the ultimate goal of reducing noise and enhancing the signal of interest. Nonetheless, before starting the study of filters and their behavior, we need to equip ourselves with some key tools that will come in handy.

The first fundamental tool is the mathematical description of both signals and noise in two domains: the time domain and the frequency domain. Knowing both these mathematical descriptions will allow us from time to time to study the problems in the domain in which the analysis and the computation are easier. For this reason, in the next section we will make a brief recall of the Fourier Transform basic principles and some of its properties that we will use during our studies. The following paragraph is only a reference, so refer to specific texts of mathematical analysis signal theory for a complete and detailed study.

2.2 Fourier transform

One of the most important mathematical tools to analyze signals is the Fourier Transform. This tool gives us the opportunity to analyze problems both in the time domain and in the frequency domain, choosing from time to time the simplest domain to solve our specific case.

The main idea behind the Fourier-transform analysis is that any arbitrary signal can be seen as superposition of sinusoids.

If we consider an arbitrary signal $x(t)$, it may be written as:

$$x(t) = \int_{-\infty}^{+\infty} X(f)e^{j2\pi ft}df$$

in which $X(f)$ is called Fourier Transform of $x(t)$. This formula is called Anti-Fourier transform. In general, the quantity $X(f)$ is a complex number and the intuition behind it is that it describes the frequency content at frequency f of the

initial signal $x(t)$. In particular, the modulus of $X(f)$ denotes the amplitude of the complex sinusoid at each frequency in the superposition, and its phase describes the phase of that complex sinusoid.

$$X(f) = \int_{-\infty}^{+\infty} x(t)e^{-j2\pi ft}dt$$

Considering at first the case of $f = 0$, we can rewrite:

$$X(0) = \int_{-\infty}^{+\infty} x(t)dt$$

This equation is just saying that the frequency content at zero frequency, i.e. the DC value of the signal $x(t)$, may be computed as something like an average over time, and this is clearly intuitive.

In the following, the main properties of the Fourier transforms that are extremely useful in this course are briefly reported.

P1. Value in zero and integrals:
Starting from the definition of Fourier transform we directly obtain:

$$X(0) = \int_{-\infty}^{+\infty} x(t)dt$$

and from the Anti-Fourier transform:

$$x(0) = \int_{-\infty}^{+\infty} X(f)df$$

This means that the value in zero of the Fourier transform equals the integral of the function over all the time domain. Analogously, the value in zero of the time-domain function is equal to the integral of the Fourier transform over all the frequency domain.

P2. Duality:
We can make an important remark at this point. If we look at the formulas that we have obtained so far for the Fourier and Anti-Fourier transforms, we can notice that they are almost equal, apart from a change in the sign of the exponent. Actually, this means that not only the function x(t) can be seen as the superposition of sinusoids, but this is also true for its Fourier transform. Indeed, the Fourier transform formula:

$$X(f) = \int_{-\infty}^{+\infty} x(t)e^{-j2\pi ft}dt$$

says that X(f) can be written as a superposition of complex exponentials, and the weights of this superposition are encoded in the function x(t) itself. The time domain and the frequency domain result as two faces of the same coin, describing the same situation in two different ways.

P3. Fourier transform of a periodized function:
A periodization in time domain corresponds to sampling in the frequency domain. Equally, sampling in time domain corresponds to a periodization in the frequency domain.

P4. Convolution and product:
The convolution of two signals in the time domain corresponds to the multiplication of their Fourier transforms in the frequency domain and vice versa.
In the same way, the multiplication of two signals in the time domain corresponds to the convolution of their Fourier transforms in the frequency domain and vice versa.

A detailed description of the Fourier Transform and all its properties is not part of this course. Here just some basic concepts that will be useful in future chapters have been reported. Just to show an example of practical use of the Fourier Transform and its properties, let's consider the case of truncated signals:
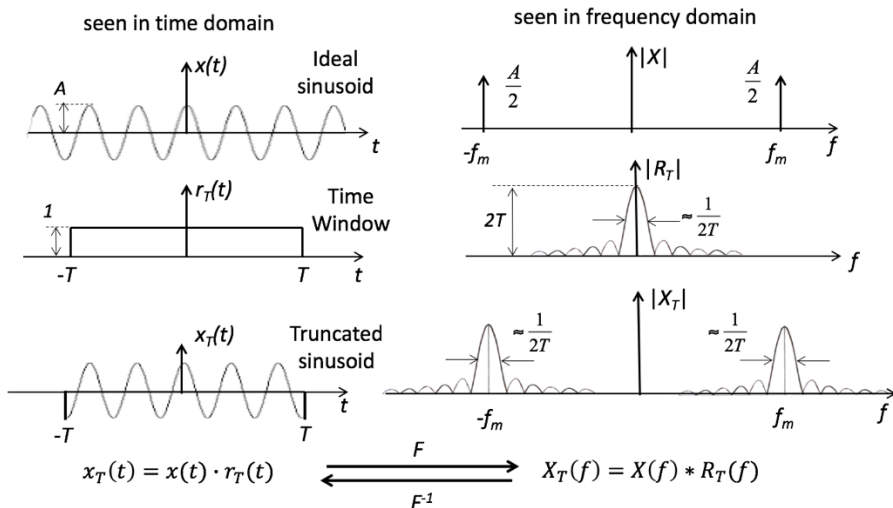
Figure 2.2 Calculation of the Fourier transform of a truncated sinusoidal

If we consider a sinusoidal signal, it can be demonstrated that its Fourier transform is a pair of Dirac deltas, as shown at the top right of Figure 2.2.

Nonetheless, in the real world we will never work with a real sinusoidal signal since we cannot observe it from minus infinite to plus infinite, but we will work on a part of it. How does this affect the information of our signal?

We can model the finite observation window of our experiments by simply multiplying the sinusoidal waveform by a rect(t) (rectangular function) featuring a time width equal to our observation window. What is the effect in the frequency domain? We are multiplying two signals (rect(t) and sinusoidal waveform) in the time domain, so the result in the frequency domain is given by the convolution of the Fourier transform of the sinusoidal waveform with the one of the rect(t), that is a cardinal *sinc* function.

The result is visible at the bottom of Fig 2.2 and it gives us the possibility to make some simple observations:

- in reality, the signal is always available over a finite time interval: therefore, in practice we always deal with truncated signals;
- signal-cropping in time corresponds to convolution of the signal in the f domain with the transform of the rectangle (*sinc* function);
- the convolution spreads the signal in the f domain, i.e. it makes it wider and smoother;
- the narrower is the window *2T*, the wider is the *sinc* and the more significant is the signal spreading in frequency domain;
- to properly apply the sampling theorem we need to observe that the sampling frequency f_S to be employed for a truncated sinusoid of frequency f_m is NOT $f_S \approx 2f_m$, but it must be remarkably higher ($f_S >> 2f_m$).

2.3 Linear systems and convolution operation

We will now discuss the effect of a generic system on a signal. A system can be seen as a "black box" having an input and an output. Among all systems, we will first consider a particular class of them: the linear systems.
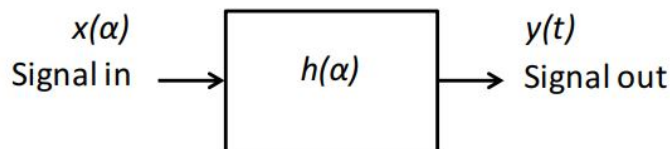

Figure 2.3.

Consider a system and let $h(\alpha)$ be the function that gives the output for each possible input. This means that, if the input signal is $x(t)$, the output is $y(t) = h[x(t)](t)$.

A system is called linear if, when scaling the input signal by a certain constant, also the output is scaled by the same constant and when the sum of two signals $x_1 + x_2$ is

applied at the input, the system produces an output equal to the sum of the individual responses to the input signals. Mathematically, a system is linear if:

$$h[\alpha x_1(t) + \beta x_2(t)](t) = \alpha h[x_1(t)](t) + \beta h[x_2(t)](t)$$

How can we exploit the property of linearity?

Let's consider this idea: if we can write a generic signal as a superposition of elementary signals and if we can calculate the response of the system to these elementary signals, then we can easily find the response to any arbitrary input. This can be done, for instance, with Dirac Delta functions. How can we write an arbitrary signal $x(t)$ as a superposition of Delta functions?

The Dirac Delta centered at time τ needs to have an area equal to the value of the signal $x(t)$ at that point, $x(\tau)$. Obviously, there are infinitely many Dirac Deltas so the superposition has to be done by means of an integral operation. We then obtain:

$$x(t) = \int_{-\infty}^{+\infty} x(\tau)\delta(t - \tau)d\tau$$

We can notice that the integral operation makes sense also for another reason. In fact, our signal is finite at each time, so it is not completely correct to say that it can be written as superposition of Deltas. It is more correct to say that it is the superposition of areas of Deltas. It is worth recalling that a Dirac Delta $\delta(t - \tau)$ is a function that is null at each time apart from $t = \tau$, where it is infinite, while its area is finite and equal to 1. From this point of view the previous formula makes sense. With the integral we basically extract the area of each Delta in the superposition.

Another interpretation could be given looking at the function $x(\tau)\delta(t - \tau)d\tau$ inside the integral. In general, by multiplying a signal $g(t)$ by a small element $d\tau$ we basically compute the local area of the function. If the function is finite, we obtain a null value, since as $d\tau$ goes to zero the area always decreases reaching the value 0. However, a Delta function is infinite for $t = \tau$ and with a finite local area, equal to 1, while its value is zero for $t \neq \tau$, so also its local area will be zero at any $t \neq \tau$ (this is actually how a Dirac Delta is defined). From this point of view the interpretation of the previous formula is clear: if we have functions that are 1 at a particular point and zero everywhere else, it is obvious that we can construct all the possible signals by taking the superposition of these elementary functions.

Knowing this fact and using linearity we can write the response of a linear system to an arbitrary input as follows:

$$y(t) = h[x(t)](t) = h\left[\int_{-\infty}^{+\infty} x(\tau)\delta(t - \tau)d\tau\right](t) = \int_{-\infty}^{+\infty} x(\tau)h[\delta(t - \tau)](t)d\tau$$

where we have used the properties of linearity to compute the last equivalence.

Looking at the final expression, we see that we can compute $y(t)$ for any $x(t)$ if we know $h[\delta(t-\tau)]$, that is the so-called *impulse response*. However, we notice that, in general, the impulse response depends on τ. This reflects the fact that the system can change over time. This means that if we apply an impulse at time $t = 0$ we could obtain an output waveform completely different from what we would get applying the impulse at a different time τ. Systems like that are called *time-variant systems*. Intuitively, they describe situations in which the topology of the system is changing over time. For instance, this is the case of a circuit in which a switch is present.

However, at the beginning we will focus our attention on a particular class of systems called Linear Time Invariant (LTI) systems.

In LTI systems, the response to an input signal does not depend on when the signal is applied, and the output waveform is "synchronized" with the input signal, i.e. if we delay the input also the output is delayed by the same quantity. Intuitively these are systems whose topology does not change with time: in this case it is obvious that the response to a certain input does not depend on where the input is applied, since the system is not changing.

Formally we will define a time-invariant system as a system such that:

$$h[x(t-\tau)](t) = h[x(t)](t-\tau)$$

This property is called *time invariancy*.

Let's see how we can exploit this property. Using the equation that we have previously found for the output signal, if we call $h(t) = h[\delta(t)](t)$ we obtain:

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)h[\delta(t-\tau)](t)d\tau = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)d\tau$$

This result is so important that the operation represented by the last integral has a specific name: it is called *convolution*.

In general, given two functions f and g, their convolution is another function ψ defined as:

$$\psi(t) = (f * g)(t) = \int_{-\infty}^{+\infty} f(\alpha)g(t-\alpha)d\alpha$$

If we want to compute the convolution at time $t = 0$, we need to solve the following integral:

$$\psi(0) = \int_{-\infty}^{+\infty} f(\alpha)g(-\alpha)d\alpha$$

This means that we need to take the function g, flip it, multiply it by the function f and compute the area of the resulting function, as depicted in figure 2.4. Similarly, the convolution at an arbitrary time can be computed as follows:

$$\psi(t) = \int_{-\infty}^{+\infty} f(\alpha)g(t - \alpha)d\alpha = \int_{-\infty}^{+\infty} f(\alpha)g(-(\alpha - t))d\alpha$$

which means that we need to take g, flip it, but before multiplying it by the function f, *we need to delay g in time by t*. Then we can compute the area of the resulting function.
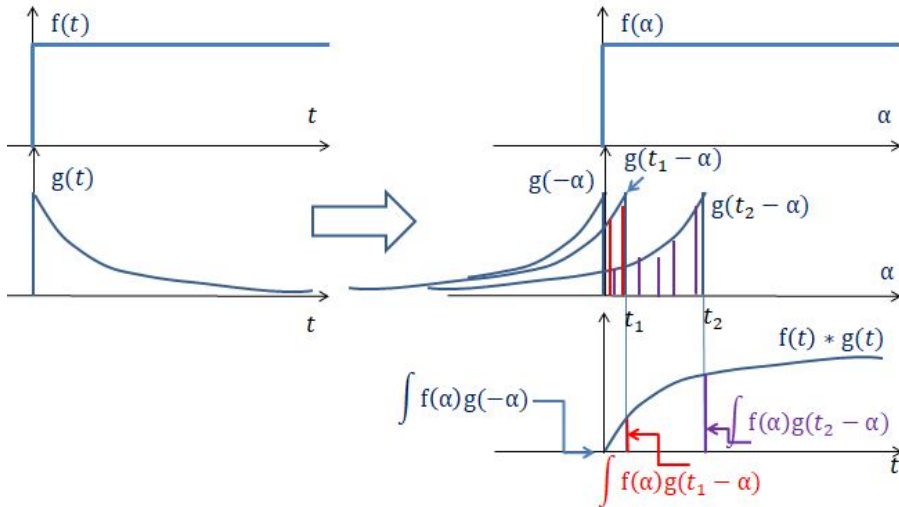
Figure 2.4 Graphical calculation of the convolution of two signals

In LTI systems, the output at time t can be obtained by computing the convolution between the input signal and the so-called **impulse response** $h(t)$. In fact, this means taking something like a linear combination of the values of the input, each one altered by the system, as will be clearer in the following chapters. This operation can be done only with linear systems.

2.4 Energy of a signal

In this paragraph, we will define one of the main properties of a signal: *the energy*. In physics the concept of energy is widely used: for example, for a particle moving with speed v we define a kinetic energy which is proportional to v^2; in the same way, a capacitor charged at a voltage V stores an electrostatic energy that is proportional to V^2; one more time, an electromagnetic wave propagating in a certain mean carries an energy proportional to the square magnitude of the electric field or magnetic field. If we want to generalize this concept to signals, we have to embed the basic properties of physical energy. In particular, the energy is always positive, it is zero only when the quantity is zero and it is proportional to the magnitude squared. Given a signal

$x(t)$ the first idea can be to consider $|x(t)|^2$. This can be a good candidate, but it is a function of time. If we want just a number, taking into account in some way the behavior of our signal on all the real line, the idea can be to take the integral of the previous quantity. We then define the *energy of a signal $x(t)$* as:

$$E[x(t)] = \int_{-\infty}^{+\infty} x(t)^2 dt$$

Signals for which this quantity exists (in a finite way) are called *energy signals.* Examples are signals with a finite duration in time, or functions going to zero at infinite quite rapidly.

2.5 Similarity of two functions, Cross-correlation operator

We will see in the next two paragraphs how the energy of a signal is strictly connected to the *autocorrelation* function. To study this new tool let's make a step back and let's start from the concept of similarity between signals. Suppose we have two functions f and g , as shown in Fig. 2.5, and we want in some way to understand if they are similar or not.
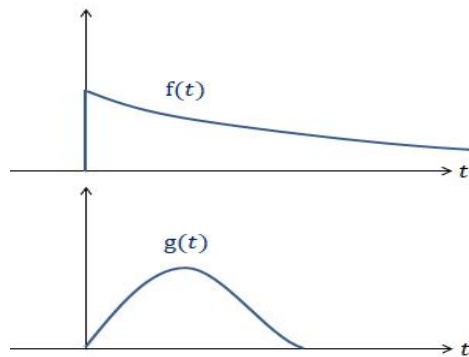

Figure 2.5.

Intuitively, what we want is a number characterizing how much these two functions resemble each other. We need an operation that in some way compares the values of the two functions at each time t and tells us, using some sort of operation, how much they are similar. However, we can understand that such a comparison cannot work. In fact, if we have, for example, two functions that are identical but shifted in time, by applying our "comparison", we will certainly obtain no similarity at all since we supposed that our operation directly compares the functions at each time t and in this case they are always different.

This means that using only a single comparison is not sufficient. To solve this problem the idea can be to make an infinite number of comparisons. We can compare the first function with the second one shifted by all the possible time intervals. This

can be done to avoid the problem of having two identical functions shifted by a certain amount of time. This means that, at the end, we will not obtain a single number, but a set of numbers related to particular shifts of the second function. These numbers can be summarized in a function that we will call *cross-correlation* function. Given two signals f and g, the cross-correlation function is defined as

$$K_{f,g}(\tau) = \int_{-\infty}^{+\infty} f(\alpha)g(\tau + \alpha)d\alpha$$

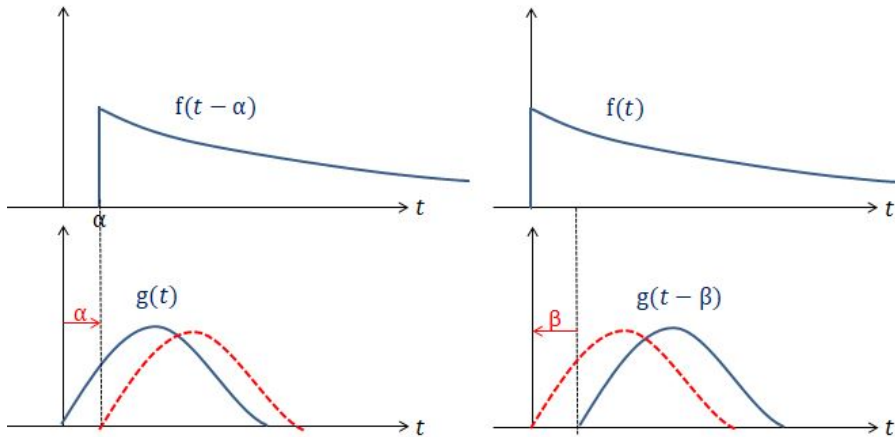Where τ is the relative time shift that we apply to g before we compare it to f.

Figure 2.6.

For example, if we want to compute the cross-correlation for $\tau = 0$ we need to compute:

$$k_{f,g}(0) = \int_{-\infty}^{+\infty} f(\alpha)g(\alpha)d\alpha$$

In this case, we take the waveform of f and the waveform of g without any shift, we multiply them together and compute the area. If we take $f = g$ we obtain the so-called *autocorrelation* function.

$$K_{x,x}(\tau) = \int_{-\infty}^{+\infty} x(\alpha)x(\tau + \alpha)d\alpha$$

Intuitively the autocorrelation function gives us information on the amount of time for which our signal is coherent with itself. Said in other terms, if we have an autocorrelation in time that is very concentrated at the origin this means that the signal is completely changing as we consider times larger than 0. Instead if we have

an autocorrelation very spread in time we expect our starting signal not to change so much for times larger than 0. Under this interpretation, looking at the autocorrelation waveform we can understand in some way where the information is localized in our signal of interest.

2.6 Properties and results related to the Cross-correlation

The cross- and autocorrelation functions are not defined for all the signals. In particular, it is possible that the integrals defining them do not converge. For this reason, these operations are defined just for a class of signals. For *energy signals*, i.e. signals with finite energy, the integrals always converge. Therefore, in the following we will always consider energy signals. For this reason the correlation operations we defined here are called energy correlations. Later we will analyze how to generalize these concepts to signals with infinite energy.

Now, we can observe that if we evaluate the autocorrelation function for $\tau = 0$, that is:

$$K_{x,x}(\tau) = \int_{-\infty}^{+\infty} x^2(t)dt.$$

we compute the same operation we have used in the previous paragraph to define the energy of a signal.

We then conclude that in general:

$$K_{x,x}(0) = E[x(t)]$$

To help the reader to become familiar with the concepts of this paragraph, we will now report an example. Let's consider an exponential signal

$$x(t) = 1(t)Ae^{\frac{-t}{T_p}}$$

in which $1(t)$ is a step function, as in Fig. 2.7. First, we will compute the value $K_{x,x}(0)$, which is equal to the energy of our signal.

$$E[x(t)] = K_{x,x}(0) = \int_0^{+\infty} x(t)^2 dt = A^2 \frac{T_p}{2}$$

To compute the value $K_{x,x}(\tau)$ we need to multiply the exponential with itself shifted to the left by τ and compute the area, as shown in Fig. 2.7.

Without doing the calculation we can immediately obtain the result. Indeed, the calculation we have to do is the same of before, but in this case the second waveform is scaled by the value $e^{-\frac{\tau}{T_p}}$. And this is true both for $\tau > 0$ and $\tau < 0$. So, also the correlation is scaled by the same quantity and we immediately obtain:

$$K_{x,x}(\tau) = K_{x,x}(0)e^{-\frac{|\tau|}{T_p}} = A^2\frac{T_p}{2}e^{-\frac{|\tau|}{T_p}}$$

This simple example shows how, starting from the graphical definition of correlation, we can immediately obtain the result without doing a lot of calculations. Let's apply, as example, the same approach to a different signal. What happens if we take x as a sum of two exponentials shifted in time as we can see in figure 2.8? For simplicity we will consider a height of 1.
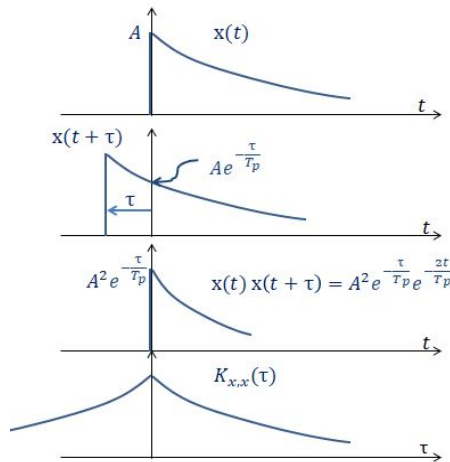

Figure 2.7.

Let's write
$$x(t) = 1(t)e^{-\frac{t}{T_p}} + 1(t - T_r)e^{-\frac{(t-T_r)}{T_p}}.$$

Calling x_1 and x_2 the two exponentials in the sum we may write, noticing that the autocorrelation is a linear operation with respect to the function we are using:

$$K_{x_1+x_2,x_1+x_2} = K_{x_1,x_1} + K_{x_2,x_2} + K_{x_1,x_2} + K_{x_2,x_1}$$

Clearly $K_{x_1,x_1} = K_{x_2,x_2}$ since the autocorrelation function of two equal waveforms is the same, independent on where the waveform is centered. This is obvious if we think to the graphical interpretation of autocorrelation. Indeed, the shifting operation is always made with respect to the first waveform, so it does not depend on where it is centered. However, the terms K_{x_1,x_2} and K_{x_2,x_1} are clearly different since now the two waveforms we are "correlating" are shifted with respect to each other. Thinking

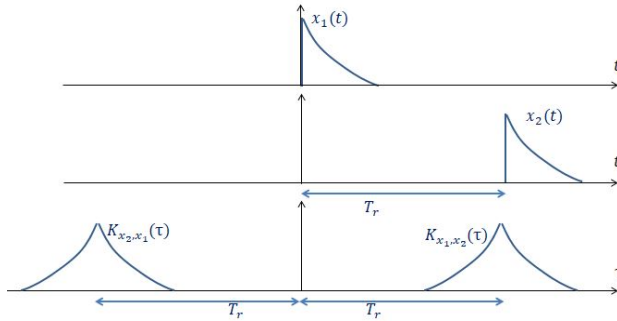to the graphical definition of the correlation let's compute K_{x_1,x_2} as we can see from figure 2.8.

Clearly for $\tau \approx 0$ we expect the correlation to be zero, since the two waveforms are far to each other. For $\tau = T_r$ we instead expect the maximum value since the two waveforms are exactly overlapping. And it is easy to see that the overall correlation is just the autocorrelation of an exponential we computed before shifted by T_r towards the right.

What about K_{x_2,x_1}?

In this case we expect the maximum value to be at $\tau = -T_r$ since we have to shift the second function by T_r towards the right. So we obtain the plot for the overall autocorrelation shown in the next figure. Formally we obtain:

$$K_{x,x}(\tau) = 2K_{x_1,x_1}(\tau) + K_{x_1,x_2}(\tau) + K_{x_2,x_1}(\tau) = T_p e^{-\frac{|\tau|}{T_p}} + \frac{T_P}{2}e^{-\frac{|\tau-T_r|}{T_p}} + \frac{T_p}{2}e^{-\frac{|t+T_r|}{T_p}}$$
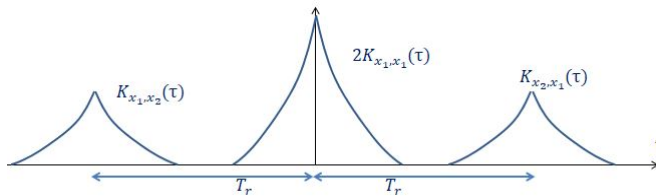

Figure 2.9.

2.7 Relation between convolution and correlation

We can observe that the procedure needed to compute the cross-correlation is similar to the operations needed to compute convolution: the only difference is that cross-correlation does not require the second function to be flipped before it is shifted and the area is computed. Also, we have seen that in convolution the shifting operation is

directed towards the right whereas in the case of the cross-correlation it is towards the left. However, we expect there is some relation between them. When we compute the cross-correlation between f and g, as we can see from the figure we need to shift g towards the left, for $\tau > 0$.
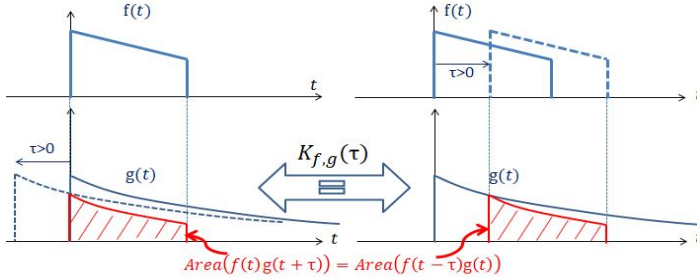

Figure 2.10

We can observe that, from the point of view of g, it is like f is shifting towards the right by the same quantity. This is good because if we want to find the convolution somewhere, we need to have a shift towards the right. For the convolution, since it is commutative, it is the same to flip and shift the first or the second function.

We know that the operation of convolution flips one of the two functions, whereas the correlation does not. The idea can then be to apply the convolution operation after a "pre-flipping" of the first function, as we can see in figure 2.11. In this way, when we flip it, the function will return to its original form as it is underlined in figure 2.11.
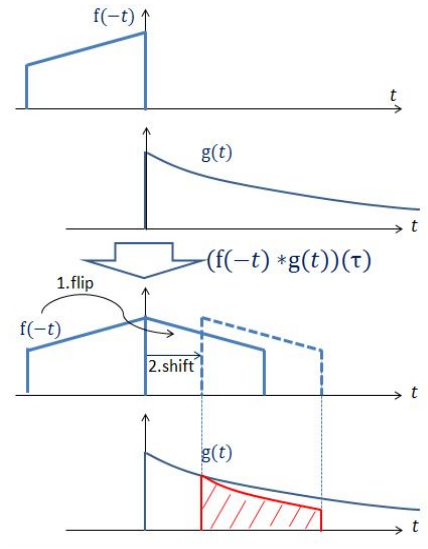

Figure 2.11.

To summarize, it seems that the cross-correlation between f and g is equivalent to the convolution between g and f, flipped with respect to the time axis. (since convolution is commutative actually the order is not important) So our guess is:

$$K_{f,g}(\tau) = (f(-t) * g(t))(\tau)$$

Let's prove it.

$$(f(-t) * g(t))(\tau) = \int_{-\infty}^{+\infty} f(-\alpha)g(\tau - \alpha)d\alpha$$

Making a change of variable $\gamma = -\alpha$ we get:

$$(f(-t) * g(t))(\tau) = \int_{-\infty}^{+\infty} f(\gamma)g(\tau + \gamma)d\gamma = K_{f,g}(\tau)$$

That is the result we wanted to prove.

2.8 Frequency interpretation of the energy

We have previously defined the energy of a signal as:

$$E[x(t)] = \int_{-\infty}^{+\infty} x(t)^2 dt$$

As always, we can try to give a view of the energy in frequency domain. Applying the Parseval theorem we can obtain:

$$E[x(t)] = \int_{-\infty}^{+\infty} x(t)^2 dt = \int_{-\infty}^{+\infty} |X(f)|^2 df$$

This result gives us an insight. The term $|X(f)|^2 df$ in the previous formula may be thought as the contribution of energy from the component at frequency f of our signal. This is why $|X(f)|^2$ is defined as the *energy spectral density* $S_x(f)$ of our signal.

2.9 Fourier transform of the Energy Autocorrelation, relation with energy spectrum

Let's see how the autocorrelation tool and the energy spectrum are strictly related. We can start from the relation between autocorrelation and energy

$$K_{x,x}(0) = E[x(t)]$$

Using the energy spectral density of $x(t)$, $S(f)$ we can also write:

$$K_{x,x}(0) = E[x(t)] = \int_{-\infty}^{+\infty} S(f)df$$

Recalling that one of the Fourier transform properties is that its integral is equal to the value in zero of the function in time, we can guess that the Fourier transform of the autocorrelation is the energy spectral density.

In order to do that, let's use the expression of the autocorrelation in terms of convolution, as we have done before:

$$K_{x,x}(\tau) = x(t) * x(-t)(\tau)$$

Taking the Fourier transform of both sides we get, using the Fourier transform of the convolution:

$$\mathcal{F}[K_{x,x}](f) = \mathcal{F}[x(t) * x(-t)](f) = \mathcal{F}[x(t)](f)\mathcal{F}[x(-t)](f)$$

It is very easy to prove that $\mathcal{F}[x(-t)](f) = \mathcal{F}[x(t)](-f)$, i.e. flipping a function with respect to the time axis flips also the Fourier transform with respect to the frequency axis. So:

$$\mathcal{F}[K_{x,x}](f) = X(f)\overline{X(f)} = |X(f)|^2$$

But $|X(f)|^2$ is the *energy spectral density* as we have seen before. So, it is true, the Fourier transform of the autocorrelation is the energy spectral density.

As we have seen the Autocorrelation function gives us the intuition of how the information is distributed in time in our signal. This result shows that the concept of information is strictly related to the concept of energy.

Let's compute the energy spectral density of the signal:

$$x(t) = V_p e^{-\frac{t}{T_p}}$$

One way could be to compute the Fourier transform of the autocorrelation. We already computed the energy autocorrelation before, it is:

$$K_{x,x}(\tau) = \frac{V_p^2 T_p}{2} e^{-\frac{|\tau|}{T_p}}$$

Computing this Fourier transform, we proved that the Fourier transform of the autocorrelation is just $|X(f)|^2$. Indeed, we know that the impulse response of a single pole system with time constant T_p is an exponential. Then we can immediately say that the Fourier transform of x has the form:

$$X(f) = \frac{A}{1 + j\frac{f}{f_p}}$$

in which $f_p = \frac{1}{2\pi T_p}$. We can find the constant A using the properties of the Fourier transform. In particular the value in zero of the Fourier transform is the integral of the function in time thus directly obtaining
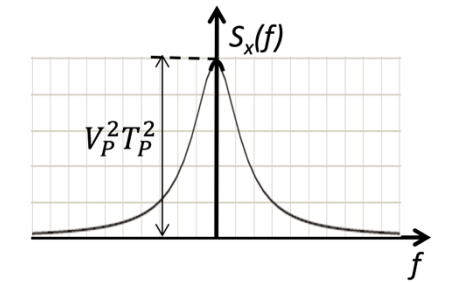
$$A = \int_0^{+\infty} V_p e^{-\frac{t}{T_p}} dt = V_p T_p$$

So we get, for the energy spectral density:

$$S_x(f) = |X(f)|^2 = \frac{(V_p T_p)^2}{1 + (\frac{f}{f_p})^2}$$

The plot is reported in figure 2.12. The reader can check that $E = \int_{-\infty}^{+\infty} S_x(f) df$ by direct computation of the integral.

Figure 2.12.

Inequalities related to Auto and Cross Correlation:
Let's consider a real signal $x(t)$, and consider its autocorrelation $K_{x,x}(\tau)$. We know that for $\tau = 0$ we obtain the energy of the signal $x(t)$. What happens for $\tau \neq 0$? Will we get a larger or lower value? Intuitively we expect to have a lower value, since the correlation operation has the intuition of giving the degree of similarity. We obviously expect that the maximum degree of similarity is obtained when we compare the waveform $x(t)$ with itself. As we consider $\tau \neq 0$ we are comparing $x(t)$ with itself shifted and so we expect a lower degree of similarity. It is possible to prove that:

$$K_{x,x}(\tau) \leq K_{x,x}(0)$$

and:

$$K_{f,g}(\tau) \leq \sqrt{K_{f,f}(0)K_{g,g}(0)}$$

2.10 Power of a signal, time and frequency domain view

So far, we considered signals with finite energy. Obviously, there are also signals with infinite energy: for instance, the simplest case is a constant signal. Is it physically possible to have a diverging energy?

Let's consider an infinite energy signal. If we look at it for a finite amount of time obviously it will have a finite energy. The idea could be to understand which is the "flow" of energy per unit time. In order to do that we can look at the signal for a finite amount of time T, compute the energy and divide by the time of observation T. We expect that, if we repeat this procedure for increasing values of T, the resulting limit will give us the intuition of energy per unit time related to the signal. In this way, we can define the *Power* for an infinite energy signal as:

$$P[x(t)] = \lim_{T \to +\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)^2 dt$$

Signals with finite power are called *power signals*. We can write the definition of power in a more compact way defining the so called truncated signal starting from $x(t)$:

$$x_T(t) = \begin{cases} x(t) & |t| < \dfrac{T}{2} \\ 0 & |t| > \dfrac{T}{2} \end{cases}$$

with this definition we can rewrite the power as :

$$P[x(t)] = \lim_{T \to +\infty} \frac{1}{T} E[x_T(t)]$$

2.11 Frequency domain interpretation of power

Let's try to give an intuitive view of power in frequency domain. The idea can be to use the concept of truncated signal we already saw. Truncated signals are always finite energy signals, so we can apply all the theory we developed for finite energy signals. If $x(t)$ is a non-periodic signal and $x_T(t)$ is the truncated signal as we previously defined, we can write:

$$E[x_T(t)] = \int_{-\infty}^{+\infty} |X_T(f)|^2 df$$

in which $X_T(f)$ is the Fourier transform of $x_T(t)$ and then $|X_T(f)|^2$ is the density of energy. Since we know that, for the signal $x(t)$ we can write:

$$P[x(t)] = \lim_{T \to +\infty} \frac{1}{T} E[x_T(t)] = \int_{-\infty}^{+\infty} \lim_{T \to +\infty} \frac{1}{T} |X_T(f)|^2 df$$

So, it makes sense to define the *power spectral density* at the frequency f as:

$$S_x(f) = \lim_{T \to +\infty} \frac{1}{T} |X_T(f)|^2$$

In this way:

$$P[x(t)] = \int_{-\infty}^{+\infty} S_x(f) df$$

The physical interpretation of the function $S_x(f)$ is the density of power at the frequency f of our initial signal. In particular multiplying $S_x(f)$ by a small frequency interval Δf we obtain the power content of our signal in the frequency interval f, $f + \Delta f$.

The procedure we followed is intuitive. Using truncated functions, we have found a succession of functions $x_T(t)$ that converges to our function $x(t)$. These functions are all finite energy signals, so the concept of energy spectral density is properly defined, and also Parseval relation holds. Since we are interested in the power content of our final signal we can define, for each of these truncated functions, the power content at frequency f simply dividing their energy spectral density by the time T of their duration. Then, making $T \to +\infty$ we obtain the concept of power spectral density for our function $x(t)$.

2.12 Extension of the Cross-correlation to Power signals

As we have seen, the concept of correlation that we have used so far is well defined only for energy signals. Is it possible to generalize it to power signals? As we have done before, we can convert the operation of integral between $-\infty$ and $+\infty$ into a time average. In this way we get a definition of cross correlation for two real power signals f and g:

$$K_{f,g}(\tau) = \lim_{T \to +\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} f(\alpha) g(t + \alpha) d\alpha$$

This can be also written in terms of the time average operation we introduced before:

$$K_{f,g}(\tau) = < f(\alpha) g(t + \alpha) >$$

We will call this operation *power cross-correlation*. If $f = g$ we will define it as *power auto-correlation*. Obviously, all the properties we showed for the energy correlation holds for the power correlation. For instance, the inequalities and the

symmetry properties still hold. We can notice that the power auto-correlation of the power signal $x(t)$ for $\tau = 0$ is the power:

$$K_{x,x}(0) = \lim_{T \to +\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} x^2(\alpha) d\alpha = P[x(t)]$$

What about the Fourier transform of the power autocorrelation? We can do a trick. We can observe that, if $x(t)$ is a power signal, its truncated signal $x_T(t)$ is an energy signal. Moreover, if we look at the power autocorrelation of $x(t)$ we can write:

$$K_{x,x}(\tau) = \lim_{T \to +\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} x(\alpha)x(\tau + \alpha) d\alpha = \lim_{T \to +\infty} \frac{1}{T} \int_{-\infty}^{+\infty} x_T(\alpha)x_T(\tau + \alpha) d\alpha$$

where we used the fact that the truncated signal is null for $|t| > \frac{T}{2}$ to extend the integral to all the real line (the reader can prove that the last equality holds only because there is the limit, while is it not true if we remove the limit.).

We notice that the last integral can be seen as the energy correlation of the signal $x_T(t)$. So if we take the Fourier transform of the previous equation, and using the fact that the Fourier transform of an energy correlation is the energy spectral density we obtain:

$$\mathcal{F}[K_{x,x}](f) = \mathcal{F}[\lim_{T \to +\infty} \frac{1}{T} \int_{-\infty}^{+\infty} x_T(\alpha)x_T(\tau + \alpha) d\alpha] =$$

$$\lim_{T \to +\infty} \frac{1}{T} \mathcal{F}[\int_{-\infty}^{+\infty} x_T(\alpha)x_T(\tau + \alpha) d\alpha](f) = \lim_{T \to +\infty} \frac{1}{T} |X_T(f)|^2$$

where $X_T(f)$ is the Fourier transform of the truncated signal. We can recognize that the term $\lim_{T \to +\infty} \frac{1}{T} |X_T(f)|^2$ is the definition of power spectral density. So we obtain that the Fourier transform of the power auto-correlation is the power spectral density $S_x(f)$:

$$\mathcal{F}[K_{x,x}](f) = S_x(f)$$

This is analogous to what we proved in the case of the energy correlation.

Noise Description

To properly carry out a measurement, we must take into account not only signal but also noise. The presence of noise limits the system resolution. Therefore, it is necessary to deeply understand what noise is and how it is generated. This chapter provides the mathematical description of noise, that is a key tool to describe, analyze and deal with this phenomenon.

3.1 Introduction

A generic measurement system can be schematically represented by a chain of blocks, each one representing a specific operation or device. In figure 3.1 we can see a typical set-up for sensor-based measurements, in which four main blocks are sketched: a sensor, a preamplifier, a filtering block and a meter.

First of all, the sensor transduces the physical variable of interest into an electrical quantity. Knowing the relationship between the variable to be measured and the electric output signal is necessary to recover the information. Typically, a linear input/output correspondence is preferable, since it allows the direct computation of the physical quantity by simply multiplying the electrical output by a transduction factor.
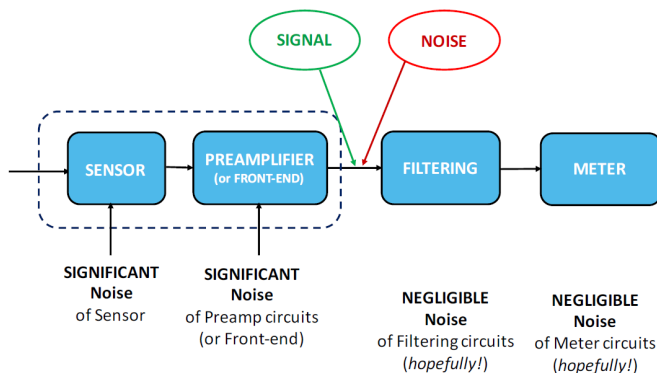
Figure 3.1 Diagram of a typical measurement set-up.

The second stage of the acquisition chain is typically an amplification block, which is needed to make the signal level sufficiently high to be read by the following stages. This block is called preamplifier, and, in some cases, it could be also the sensor front end.

As highlighted in figure. 3.1, at the preamplifier output we typically have a significant noise level superimposed to the signal, which is due to sensor intrinsic noise and to noise limitations of the gain stage. This explains why, after the preamplifier, we typically find a filtering block. The task of the filtering stage is shaping noise and signal to improve the Signal-to-Noise Ratio (SNR).

Finally, the last block is constituted by meter circuits.

It is worth noting that the noise of the two last blocks can be typically neglected in a well-designed system, while we must pay attention to noise coming out from the first two blocks.

After these practical considerations we can more easily understand noise effects on a measurement and we can also try to give some qualitative definitions: noise can be described like random disturbances which are superimposed to a useful signal and tend to obscure its information content, thus degrading the quality of the signal. These unwanted fluctuations are generated randomly within the devices themselves and they may limit the resolution of the detection system. Therefore, in a measurement system it's fundamental to reduce the noise contribution.

In conclusion, we can summarize this introductory discussion stating that acquiring signals means to recover information from noise. This explains why it's so important to properly deal with noise to perform a correct measurement.

3.2 Statistics of noise samples and probability distribution

Using an oscilloscope, it is possible to look at noise waveforms in time in a real acquisition system. Depending on what kind of noise we are dealing with, we can visualize different noise shapes on the oscilloscope screen; some examples are sketched in the fig. 3.2.
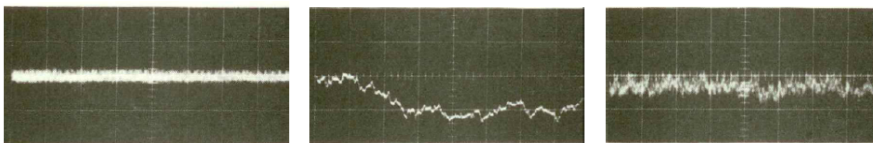
Figure 3.2 Noise waveforms on the oscilloscope screen (50μs/div). From left to right we can see a typical white noise waveform (spectrum *S=constant*), a typical random-walk noise waveform (spectrum $S=1/f^2$) and a typical flicker noise waveform (having a spectrum dependence *S=1/f*).

Looking at noise waveforms in time, we can evaluate the noise amplitude at a particular time instant t_1: we will call it $x(t_1)$.

Performing this measurement on many identical devices, as shown in figure 3.3, we can sample different noise amplitudes $x(t_1)$, ending up with a set of different values, that give us an ensemble of noise waveform amplitude.
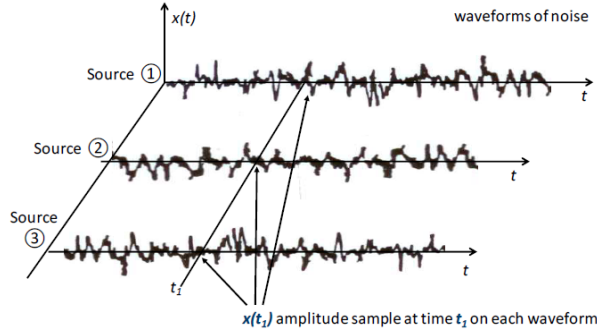

Figure 3.3 Noise waveforms ensemble coming from sampling equal noise sources at the same t_1

Once we have sampled the noise amplitude $x(t_1)$ at the particular time instant t_1, we can then compare this value to a scale of discrete values x_k, spaced by constant interval Δx and classify this sampled value at the nearest value x_k of the scale. Observing a high number of noise waveforms N and calling ΔN_k the number of sampled waveforms classified at x_k we can define the statistical frequency of the amplitude x_k as the ratio:

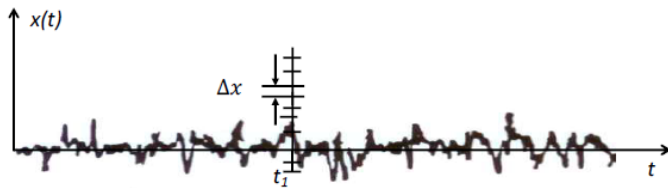$$\Delta f_k = \frac{\Delta N_k}{N}$$

Figure 3.4 Noise samples amplitude classification.

Using this procedure, we collect N values of $x(t_1)$ in N different waveforms. Then we define ΔN_0 as the number of occurrences in Δx around $x=0$, ΔN_1 the number of samples classified in the first Δx centered in x_1 and so on, up to ΔN_k (that is the number of samples classified in the k-th Δx centered in $x_k=\Delta x_k$). We can evaluate the k-th statistical frequency and build a histogram of the measured x values, as sketched on the left part of Fig. 3.5.

Moving to differentials, so $\Delta x \to dx$, we get for an infinitesimal variation dx:

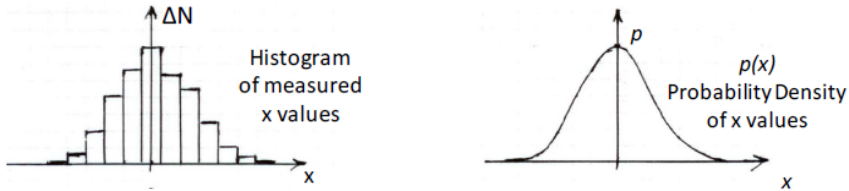$$\Delta N_k \to dN_k = n(x_k)dx$$

hence for $\Delta f_k \rightarrow df_k$ we obtain

$$df_k = \frac{dN_k}{N} = \frac{n(x_k)}{N} dx$$

For $N \rightarrow \infty$ we can explicit the last equation in this way

$$df_k = \frac{n(x_k)}{N} dx = p(x)dx$$

where the ratio $\frac{n(x_k)}{N}$ is called probability density $p(x)$ of x values.

Figure 3.5 On the left: histogram of measured x values. On the right: probability density $p(x)$ of x values.

More in general, we may have two kind of probability noise densities: it may be constant in time or it may be varying in time. In the first case we call the noise **stationary**, having $p = p(x)$, while in the second case, a probability density that varies over time gives origin to a **non-stationary** noise, having $p = p(x,t)$.

At this point of our discussion, the reader may be induced to think that knowing the probability density corresponds to having a complete mathematical description of noise behavior. However, the question is subtle and a bit misleading because probability density alone does not give us a complete description of the noise. In fact, different cases can have equal probability density.

Let's use an example to better explain this concept. Let's observe the output waveforms of two different sets of amplifiers, one composed by noisy elements (set A) and the other one composed by low-noise amplifiers (set B). Consider having a stationary noise with a probability density $p_A(x)$ in set A and a random baseline offset with probability density $p_B(x)$ in set B.
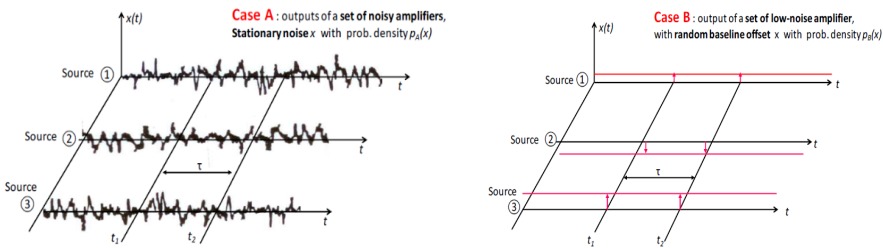
Figure 3.6 Output noise evaluation at time instants t1 and t2 on two different sets of amplifiers. On the left, case A: noisy amplifiers; on the right, case B: set of low-noise amplifiers.

In set A, if we evaluate the outputs noise amplitude at two different time instants t_1 and t_2, we end up with two random values, respectively $x(t_1)$ and $x(t_2)$ having the same probability density $p_A(x)$. Now, if we call τ the time interval between t_1 and t_2 ($\tau = t_2 - t_1$) and we study $x(t_1)$ and $x(t_2)$ as a function of τ, we end up with some properties of the sampled values, related to the interval τ. In particular, we can see that, for each source, $x(t_1)$ and $x(t_2)$ are practically identical for an ultra-slow interval τ, they are somewhat different for a short τ, while they become more and more different as we progressively increase the time interval τ.

Considering the amplifiers of set B, instead, the situation is totally different.

In this case, if we sample the output waveform at two different time instants t_1 and t_2, we still have that both $x(t_1)$ and $x(t_2)$ are random values with a probability density $p_B(x)$, but now if we compare $x(t_1)$ and $x(t_2)$ sampled on the same amplifier, we see that they are identical for any time distance τ between t_1 and t_2. In other words, the value of the amplifier output in set B is different from component to component (following a statistical fluctuation), but for each amplifier the output is constant over time.

A time-constant value at the output is called offset or baseline. Offset is completely different from noise, but they may look the same if we only study the probability density. Hence, we need other tools to fully describe noise.

3.3 Complete noise description with probability distribution

To fully describe noise in a proper way, as underlined at the end of the last paragraph, the marginal probability $p_m(x, t)dx$ of having a value x at time t is not sufficient. We need also to know, instead, the joint probability $p_j(x_1, x_2, t_1, t_2)dx_1 dx_2$ of having a value x_1 at time t_1 and a value x_2 at time t_2. Therefore, a full description of noise is obtained by knowing both the marginal probability density $p_m(x) = p_m(x, t_1)$ for every instant t_1 (that for stationary noise does not depend on time t_1, i.e. $p_m = p_m(x)$), and also the joint probability density $p_j(x_1, x_2) = p_j(x_1, x_2, t_1, t_2) = p_j(x_1, x_2, t_1, t_1 + \tau)$ for every couple of instants t_1 and $t_2 = t_1 + \tau$. We can notice that for stationary noise the joint probability p_j depends only on the time interval τ and not on the time position t_1.
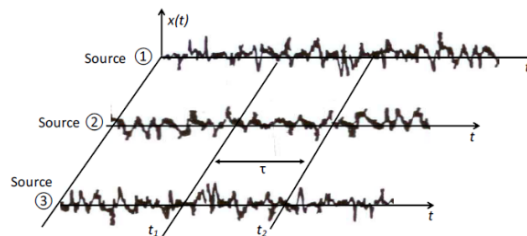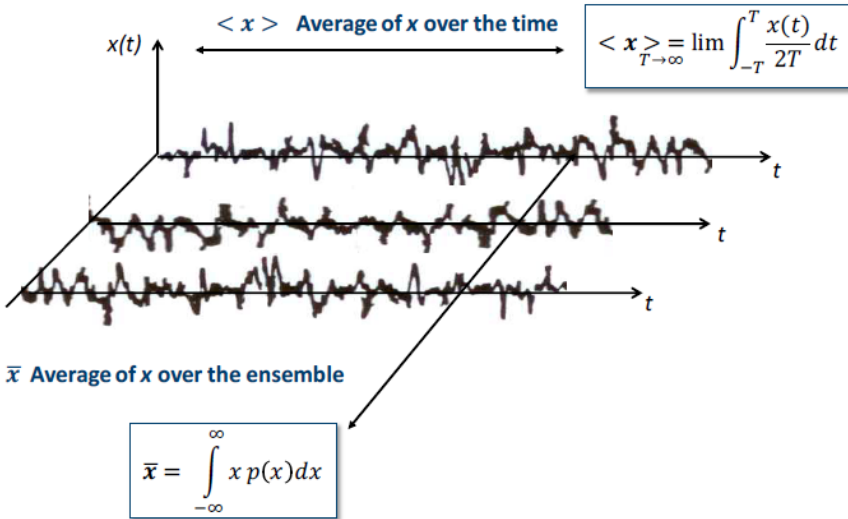

Figure 3.7 Graphic visualization of τ on a set of noise waveforms.

Before going on with this discussion, it's worth recalling some basics. First of all, averaging in time is different from averaging over the ensemble. To better explain this statement, let's consider the experiment of the previous section where we looked at noise amplitude waveforms of many sources. When we average in time, we choose a particular waveform among all the waveforms analyzed and we calculate the integral $\int_{-T}^{T} \frac{x(t)}{2T} dt$ for $T \to \infty$. The integral over time will be referred to with the following notation: $< x >$.

In this way, we obtain the mean value over time, that we can consider like the continuous component associated to this particular waveform evolving in time.

In summary:

$$< x > = \lim_{T \to \infty} \int_{-T}^{T} \frac{x(t)}{2T} dt$$

Figure 3.8 Graphic visualization of the differences between average over time and average over the ensemble on a set of noise waveforms.

On the other hand, when we average the amplitude x over the ensemble, we integrate the x variable multiplied by its probability distribution $p(x)$ from $-\infty$ to $+\infty$. We will indicate the integration result with this notation: $\bar{x}$. In formulas:

$$\bar{x} = \int_{-\infty}^{+\infty} x\, p(x) dx$$

A "graphic explanation" of these differences, maybe more straightforward and intuitive, is shown in figure 3.8.

Although this observation may seem obvious and somewhat trivial, this is a tricky

point and it's fundamental to keep clearly in mind these formulas to better and deeply understand next topics.

3.3 Noise description with 2nd order moments of PD

A fundamental way to describe noise is using the 2nd order moment of the probability density. Let's recall some statistical definitions.

Given any two statistical variables x and y, we define moments of a marginal probability distribution $p(x)$

$$m_n = \overline{x^n} = \int_{-\infty}^{+\infty} x^n p(x) dx$$

and moments of a joint probability distribution $p(x,y)$

$$m_{jk} = \overline{x^j y^k} = \int_{-\infty}^{+\infty} x^j y^k p(x,y) dx dy$$

Moments m_n and m_{jk} give the information on the features of the marginal and joint probability distributions, respectively. In particular, the higher the order n or $j+k$, the more detailed is the information.

If we consider a description of noise limited to the 2nd order moment we obtain the so-called **mean square value** or **variance**, considering the only marginal $p(x)$

$$m_2 = \overline{x^2} = \int_{-\infty}^{+\infty} x^2 p(x) dx = \sigma_x{}^2$$

In case of a joint $p(x,y)$, the 2nd order moment is called **mean product value** or **covariance** of x and y

$$m_{11} = \overline{xy} = \int_{-\infty}^{+\infty} xy p(x,y) dx dy = \sigma_{xy}{}^2$$

Concerning noise, we can evaluate the mean square value of noise amplitude waveform when sampled at a generic time instant t_1

$$\overline{x^2(t_1)} = \sigma_x{}^2(t_1)$$

which does not depend on t_1 in the case of stationary noise.

For any pair of time instants t_1 and $t_2=t_1+\tau$ (look at figure 3.6 for a graphic visualization) it's possible to calculate the mean product

$$\overline{x(t_1)x(t_2)} = \overline{x(t_1)x(t_1 + \tau)}$$

which, for stationary noise, depends only on the time interval τ and not on the time position t_1.

3.4 Noise autocorrelation function and noise power

We will now introduce a key tool for noise description: the ***noise autocorrelation function***. It is defined in this way

$$R_{xx}(t_1, t_1 + \tau) = \overline{x(t_1)x(t_1 + \tau)}$$

where we have expressed t_2 as $t_2 = t_1 + \tau$.

We can notice that the noise autocorrelation function is always function of the interval τ between the two instants t_1 and t_2. Furthermore, it is also a function of t_1 in the case of non-stationary noise.

At this point it's worth pointing some differences between signal and noise autocorrelation functions. ***The autocorrelation function of a noise x, called $R_{xx}(\tau)$, is an ensemble average***, while for a signal x ***the autocorrelation function $K_{xx}(\tau)$ is a time average***.

Another important definition is that the noise mean square value is also called noise power and it corresponds to the autocorrelation function at $\tau=0$

$$\overline{x^2(t)} = R_{xx}(t, 0)$$

For stationary noise, we see that the autocorrelation function for $\tau=0$ is constant at any t, therefore, using the same definition above introduced, it is

$$\overline{x^2} = R_{xx}(0)$$

3.5 Noise power spectrum

Finally, the noise power spectrum is defined in this chapter. We can start our discussion saying that noise has power-type waveforms, having divergent energy. Furthermore, considering the whole ensemble, we have statistical variations from waveform to waveform. We can extend the concepts of power and power spectrum that we defined for a generic signal by adding the ensemble average. In other words, the power of a noise signal is defined as the ensemble average of the power of a noise waveform:

$$P = \lim_{T \to \infty} \int_{-T}^{T} \overline{\frac{x^2(\alpha)}{2T}} d\alpha$$

In order to apply the Parseval's theorem, we must change the extremes of integration. In fact, we remind the reader that Parseval's theorem is valid only for the entire integral calculated from $-\infty$ to $+\infty$ and not for the truncated integral evaluated in $\pm T$. Hence, we introduce the truncated waveform $x_T(\alpha)$ and we change the extremes of integration

$$P = \lim_{T \to \infty} \overline{\int_{-\infty}^{+\infty} \frac{x_T{}^2(\alpha)}{2T} d\alpha}$$

Applying now Parseval's theorem we get

$$\lim_{T \to \infty} \overline{\int_{-\infty}^{+\infty} \frac{x_T{}^2(\alpha)}{2T} d\alpha} = \lim_{T \to \infty} \overline{\int_{-\infty}^{+\infty} \frac{|X_T{}^2(f)|}{2T} df}$$

Moving the limit into the integral and then average operator into the limit operator we get

$$\lim_{T \to \infty} \overline{\int_{-\infty}^{+\infty} \frac{|X_T{}^2(f)|}{2T} df} = \int_{-\infty}^{+\infty} \lim_{T \to \infty} \overline{\frac{|X_T{}^2(f)|}{2T}} df =$$
$$= \int_{-\infty}^{+\infty} \lim_{T \to \infty} \frac{\overline{|X_T{}^2(f)|}}{2T} df$$

We define the ***power spectrum of the noise*** as

$$S_x(f) = \lim_{T \to \infty} \frac{\overline{|X_T{}^2(f)|}}{2T}$$

Hence the ***noise power*** is

$$P = \int_{-\infty}^{+\infty} S_x(f) df$$

By averaging over the ensemble, we can extend to noise also the second definition of power spectrum introduced for the signals

$$S_x(f) = \overline{F[K_{xx}(\tau)]} = F[\overline{K_{xx}(\tau)}] =$$
$$= F\left[\lim_{T \to \infty} \overline{\frac{\int_{-\infty}^{\infty} x_T(\alpha) x_T(\alpha + \tau) d\alpha}{2T}}\right] =$$
$$= F\left[\lim_{T \to \infty} \overline{\frac{K_{xx,T}(\tau)}{2T}}\right] = \lim_{T \to \infty} \frac{F[\overline{K_{xx,T}(\tau)}]}{2T}$$

And so we can write the noise power as

$$P = \int_{-\infty}^{+\infty} S_x(f) df = \overline{K_{xx}(0)}$$

The last formula represents the most important result of all this chapter. Indeed, it will be largely used in studying noise and signal filtering.

Before concluding this discussion, it's worth doing some additional observations on noise power spectral density. As we can see from previous formulas, the mathematical spectral density $S_x(f)$ is defined over a frequency range that goes from $-\infty$ to $+\infty$, therefore we call it **bilateral spectral density** and we normally underline this adding a second subscript B to the notation in this way $S_{xB}(f)$. Then the noise power is calculated as

$$P = \int_{-\infty}^{+\infty} S_{xB}(f)df$$

Since $S_{xB}(f)$ is symmetrical $S_{xB}(-f) = S_{xB}(+f)$, hence we can write the noise power in this way

$$P = 2\int_{0}^{+\infty} S_{xB}(f)df = \int_{0}^{+\infty} 2S_{xB}(f)df$$

At this point we can define the **unilateral spectral density** $S_{xU}(f)$ as

$$S_{xU}(f) = 2S_{xB}(f)$$

This last spectral density definition is usually employed in engineering tasks for making computations only in the positive frequency range. Therefore, the noise power computed using the unilateral power spectral density definition is

$$P = \int_{0}^{+\infty} S_{xU}(f)df$$

Finally, a last observation remains to be pointed out. From the previous equation we have

$$S_x(f) = \overline{F[K_{xx}(\tau)]}$$

We notice that $\overline{F[K_{xx}(\tau)]}$ results from a double average: the first over the time

$$K_{xx}(\tau) = <x(\tau)x(t+\tau)>$$

and then over the ensemble.
It can be demonstrated that the order of averaging can be exchanged, as reported in the appendix section at the end of this chapter.
As a result, the power spectrum is related to the ensemble autocorrelation function in this way:

$$S_x(f) = F[<R_{xx}(t, t+\tau)>]$$

Where K_{xx} has been replaced with R_{xx}.

For non-stationary noise $S_x(f)$ can be defined with reference to the time-average of the ensemble autocorrelation function of noise.

For stationary noise instead, there is no need of time averaging, in fact it is simply

$$< R_{xx}(t, t + \tau) > \, = R_{xx}(\tau)$$

Therefore

$$S_x(f) = F[R_{xx}(\tau)]$$

3.6 Appendix

Here we will simply demonstrate that the order of time-averaging and ensemble-averaging can be exchanged in the definition of noise power spectrum.

So, let's verify that

$$\overline{K_{xx}(\tau)} = \, < R_{xx}(t, t + \tau) >$$

We have that

$$\overline{K_{xx}(\tau)} = \lim_{T \to \infty} \int_{-T}^{T} \overline{\frac{x(\alpha)x(\alpha + \tau)}{2T}} d\alpha =$$

$$= \lim_{T \to \infty} \int_{-T}^{T} \frac{\overline{x(\alpha)x(\alpha + \tau)}}{2T} d\alpha =$$

$$= \lim_{T \to \infty} \int_{-T}^{T} \frac{R_{xx}(\alpha, \alpha + \tau)}{2T} d\alpha =$$

$$= \, < R_{xx}(t, t + \tau) >$$

CHAPTER 4

Noise types and sources

We define noise the statistical contribution coming from our system that sums up with the signal, while we define as disturb any other contribution coming from outside the system. In this chapter we will focus our attention on the noise, discussing its characteristics and its contributions due to both the sensor and the electronics.

4.1 Introduction

In the previous chapter, we have seen how to analytically describe the noise in the time domain with the autocorrelation function and in the frequency domain with power spectral density.

In this chapter, we will analyze more in detail the physical origin of different noise sources. In the first part we will study the autocorrelation function and the power spectral density for the most common types of noises: the shot noise and the Johnson-Nyquist noise. Finally, in the last part we will define what is white noise.

Our final goal is to find the rms value of the noise: in fact, in the end we want to know what the minimum detectable signal of the system is, and it can be written as:

$$V_{pmin} = SNR_{min} \cdot \sigma_n$$

Normally, the minimum value of a usable SNR is considered to be 3. However, for simplicity we will consider the minimum SNR=1, in this way we are assuming that the minimum detectable signal is the one equal to the standard deviation of the noise. Clearly, this is more a convention than a practical value.

This concept will be explained better in the next chapters; at the moment we just want to underline that a critical part of our acquisition stage is to understand the variance of the noise. To do that we need to compute the power spectral density and the autocorrelation function of the noise.

4.2 Shot noise

The shot noise is a kind of noise present when we have random fluctuations of the current value due to the discrete variations of the carriers moving under the effect of the electric field.

For example, considering a reverse biased p-n junction, one carrier reaching the depletion layer will cross this region under the effect of the electric field, and it will move in an almost straight line due to the high electric field; the number of carriers crossing the junction could have a constant average value but it has also an intrinsic variation due to the discrete nature of the carriers.

We can study the total current in our device studying the behavior of each single pulse due to a single carrier. At the moment, let's use for this pulse an exponential decay. We know that the integral of the current pulse f(t) in time is the charge of the carrier, which is the elementary charge q, so we can rewrite the pulse in order to highlight h(t), a pulse with unitary area.

$f(t) = q\, h(t)$

q pulse charge

$h(t)$ normalized pulse shape: $\int_{-\infty}^{\infty} h(t)\,dt = 1$
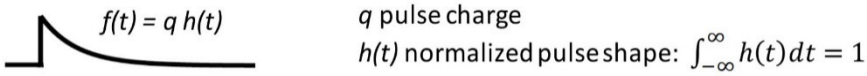
Figure 4.1 Exponential shape of the pulse

The sequence of pulses crossing the depleted region is a Poisson statistical process, since every pulse is independent from the other.

The probability of having a pulse in the interval $t+dt$ is $p \cdot dt$ where p is constant.

4.2.1 Shot noise mean

We can then calculate the current at a generic time t as given by the sum of all the contribution of the pulses started before the time t, as represented in figure 4.2.

If the pulse has started far away in time its contribution will be negligible compared to one that has just started, given the exponential shape of the pulse itself.
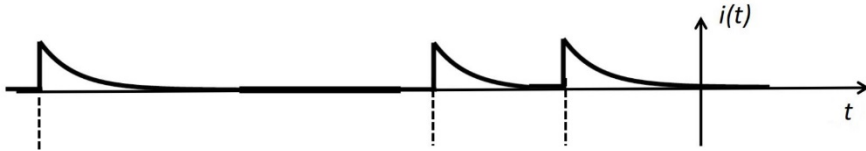
Figure 4.2 Contribution of pulses in time

To compute the average current, let's change the variable of integration: we introduce a new axis α which is flipped compared to the original one and has the zero value in correspondence of the time of measurement of the current.

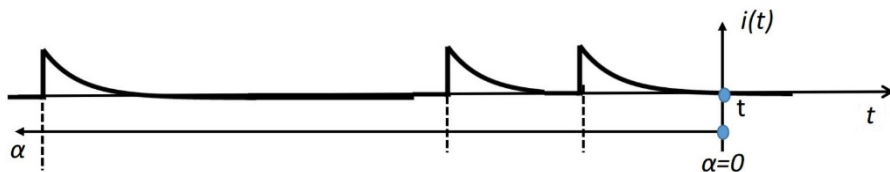**Figure 4.3** Contribution of the pulses on the axis α

A pulse which starts at time alpha contributes a current $q \cdot h(\alpha)$ at time t. The probability that a pulse starts in $\alpha \rightarrow \alpha + d\alpha$ is $p \cdot d\alpha$.

The mean current in time t is so the sum of the mean effects of the pulses.

$$\overline{\iota(t)} = I = \int_0^\infty q h(\alpha) \cdot p d\alpha = pq \int_0^\infty h(\alpha) d\alpha = pq$$

Note that the result can be understood even without computation: we have a pulse with a rate p and each one has a charge q. So, the current is the product of the average pulse rate times the charge.

4.2.2 Shot noise mean square

Our goal is to evaluate the noise power and the power spectral density of the shot noise, so the next step is to compute the square value.

The procedure is like the previous one, but we have to take in consideration two different pulses, and we have to evaluate the contribution of two pulses at the time t. The contribution of pulses to the current is:

$$\overline{\iota^2} = \left(q \cdot h(\alpha) + q \cdot h(\beta)\right)^2 =$$
$$= q^2 \cdot h^2(\alpha) + q^2 \cdot h^2(\beta) + q \cdot h(\alpha) \cdot q \cdot h(\beta) + q \cdot h(\alpha) \cdot q \cdot h(\beta)$$

Every pulse introduces two contributions: one related to itself and the other one related to the cross product.

The probability of the squared pulse contribution is the same of the probability of the pulse itself, while the probability of the cross product is the joint probability of the two pulses, because both pulses must occur at the same time. Since the pulses are independent from each other, the joint probability is just the product of the two probabilities.

Now we can do the integral to calculate the squared current:

$$\overline{\iota^2(t)} = \int_0^\infty q^2 h^2(\alpha) \cdot p d\alpha + \iint_0^\infty q h(\alpha) \, p d\alpha \, q h(\beta) \, p \, d\beta =$$
$$= pq^2 \int_0^\infty h^2(\alpha) d\alpha + pq \int_0^\infty h(\alpha) d\alpha \cdot pq \int_0^\infty h(\alpha) d\alpha =$$
$$= pq^2 \int_0^\infty h^2(\alpha) d\alpha + (pq)^2 = pq^2 \int_0^\infty h^2(\alpha) d\alpha + (\overline{\iota(t)})^2 =$$
$$= pq^2 \int_0^\infty h^2(\alpha) d\alpha + I^2$$

The result is the sum of the average value squared and another term. For our studies we are interested in the fluctuations of the current around the nominal value, so we can subtract the average squared value and we analyze the remaining term.

$$\overline{n_\iota{}^2} = \overline{\iota^2(t)} - \left(\overline{\iota(t)}\right)^2 = pq^2 \int_0^\infty h^2(\alpha)d\alpha =$$

$$\overline{n_\iota{}^2} = qI \int_0^\infty h^2(\alpha)d\alpha$$

This formula is the result of the **Campbell Theorem** and it gives us a direct relationship between the power of the noise and the shape of the delta response of the system. It is important to notice that we used the exponential decay time just to have a graphical example of pulse shape, but the obtained result is totally independent from the shape of the pulse and it can be extended to all the cases.

4.2.3 Power spectrum

Observing that the function $h(\alpha)$ is null for α negative (when α is negative, t is greater than t_m, the time in which we are evaluating the current so there cannot be any contribution), we can then rewrite the Campbell Theorem in this form:

$$\overline{n_\iota{}^2} = qI \int_0^\infty h^2(\alpha)d\alpha = qI \int_{-\infty}^\infty h^2(\alpha)d\alpha$$

we define $H(f)$ as the Fourier transform of the pulse $h(t)$

$$H(f) = F[h(t)]$$

Thanks to the Parseval Theorem, we can rewrite the formula in the frequency domain

$$\overline{n_\iota{}^2} = qI \int_{-\infty}^\infty h^2(\alpha)d\alpha = qI \int_{-\infty}^\infty |H(f)|^2 \, df$$

From the noise theory we know that

$$\overline{n_\iota{}^2} = \int_{-\infty}^\infty S_n(f)df$$

By matching the latter formula with the result of the Campbell theorem we can assume that

$$S_n(f) = qI \, |H(f)|^2$$

Pay attention, matching the two integrals is not enough to find the expression of $S_n(f)$. Indeed, there can be an infinite number of different functions that have the same integral from minus to plus infinite. To demonstrate that $S_n(f)$ is the value we found, we should need to compute the autocorrelation of the current.

Remarks on the Fourier transform.
Be careful, the variable α is not the variable t, so in principle we should compute the Fourier transform of $h(\alpha)$. Let's analyze alpha: we introduced the variable alpha as **t_m-t**, so we have a time shifting and multiplication by -1.

$\alpha = -(t\text{-}t_m)$

By the properties of the Fourier transform, we know that the effect of the time shifting is a phase shift in frequency, but we don't mind here since we are interested in the absolute value. Moreover, the multiplication by -1 gives a flipped Fourier transform, but h(t) is real so |H(f)| is pair. For these reasons we can say that the Fourier transform of $h(\alpha)$ is the same of h(t).
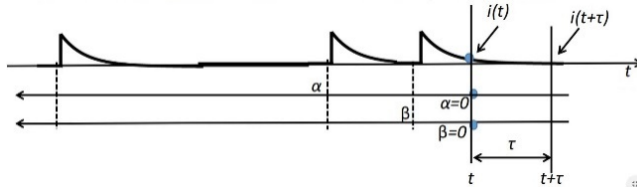

4.2.4 Autocorrelation function of the current

Figure 4.4 Contributions of more pulses at time t and t+τ

By definition the autocorrelation of the current is given by the following formula:

$$R_{ii}(\tau) = \overline{i(t)i(t+\tau)}$$

The current contributions in both time intervals are:

$$i(t) = qh(\alpha) + qh(\beta)$$
$$i(t+\tau) = qh(\alpha+\tau) + qh(\beta+\tau)$$

The procedure is similar to the one used to compute the mean square of the current. The multiplication between the current consists of two types of terms, the first one being a square contribution of each pulse individually and the second one a cross-product contribution between a pair of pulses

$$qh(\alpha) \cdot qh(\alpha + \tau) \quad \text{and} \quad qh(\alpha) \cdot qh(\beta + \tau)$$

The probability of the square contribution is the probability that the single pulse exists $p \cdot d\alpha$, while the probability of the cross-product contribution is the probability of the two pulses to exist $p \cdot d\alpha \cdot p \cdot d\beta$.

To compute the autocorrelation, we have to do the ensemble average of $i(t)i(t + \tau)$.

$$R_{ii}(\tau) = \overline{\iota(t)\iota(t + \tau)} =$$

$$= \int_{-\infty}^{\infty} q^2 h(\alpha)h(\alpha + \tau) \, pd\alpha + \iint_{-\infty}^{\infty} qh(\alpha) \, pd\alpha \cdot qh(\beta + \tau) \, p \, d\beta =$$

$$R_{ii}(\tau) = pq^2 \int_{-\infty}^{\infty} h(\alpha)h(\alpha + \tau) \, d\alpha + pq \int_{-\infty}^{\infty} h(\alpha)d\alpha \cdot pq \int_{-\infty}^{\infty} h(\beta + \tau)d\beta =$$

$$= pq^2 \int_{0}^{\infty} h(\alpha)h(\alpha + \tau) \, d\alpha + (pq)^2 =$$

$$= pq^2 \int_{0}^{\infty} h(\alpha)h(\alpha + \tau) \, d\alpha + (\overline{\iota(t)})^2 =$$

$$= qI \int_{0}^{\infty} h(\alpha)h(\alpha + \tau) \, d\alpha + I^2$$

We are interested in finding the autocorrelation of the noise, not the autocorrelation of the whole current. We know that the noise is defined as the random fluctuation around the average value.

$$n_i(t) = i(t) - \overline{\iota(t)} = i(t) - I$$

To compute its autocorrelation, we subtract the average squared value from the autocorrelation of the current and we end up with the following result

$$R_{nn}(\tau) = R_{ii}(\tau) - \left(\overline{\iota(t)}\right)^2 = R_{ii}(\tau) - I^2 =$$
$$= qI \int_{0}^{\infty} h(\alpha)h(\alpha + \tau) \, d\alpha$$

$$R_{nn}(\tau) = qI \, k_{hh}(\tau)$$

To find the power spectral density we need to apply the Fourier transform to the noise autocorrelation.

$$S_n(f) = F[R_{nn}(\tau)] \qquad S_n(f) = qI|H(f)|^2$$

4.2.5 Autocorrelation function

It has been demonstrated that the autocorrelation function is the inverse Fourier transform of the power spectral density.

$$R_{nn}(\tau) = F^{-1}\left[S_n(f)\right] = qI \cdot F^{-1}\left[\left|H[f]\right|^2\right] = qI \cdot k_{hh}(\tau)$$

We can then use this result to compute the power of the noise also in the time domain:

$$\int_{-\infty}^{\infty} S_n(f)df = \overline{n_i^2} = pq^2 \cdot k_{hh}(0) = qI \int_0^{\infty} h^2(\alpha)d\alpha$$

4.3 Noise in diodes

4.3.1 Noise in reverse biased diode

When the diode is reverse biased, we have a higher voltage drop across the depletion region. The only contribution to the current is given by the minority carriers that fall down the potential barrier and cross the junction
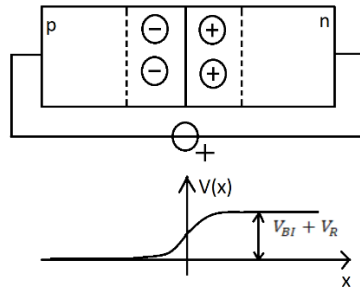
Figure 4.5 p-n junction in reverse bias and potential profile

We can then model the current as a sequence of elementary pulses: the pulse width will be related to the transit time in the junction.
The transit time of the carriers in the depletion region is the ratio between depletion length and speed. The depletion length ranges from 0.1μm to 100μm, so the transit time will range from a few ps to 1ns.
For most of our applications, we are interested in frequencies up to the MHz range, which means that the correlation times is in the order of microseconds. Since the shape of the current pulse is much shorter, it can be considered a δ for our studies.

$$h(t) \approx \delta(t) \quad \text{and} \quad |H(f)| \approx 1$$

With this approximation the noise spectrum is:

$$S_{nB}(f) = qI_s \cdot |H(f)|^2 \;\approx\; qI_s \qquad \text{bilateral power spectral density}$$

$$S_{nU}(f) = 2qI_s \cdot |H(f)|^2 \approx 2qI_s \qquad \text{unilateral power spectral density}$$

Therefore, the corresponding autocorrelation function is δ-like:

$$R_{nn}(\tau) = qI_s \cdot k_{hh}(\tau) \approx qI_s \cdot \delta(\tau)$$

4.3.2 Noise in forward biased diode

The current flowing in the diode in forward regime is composed by two contributions: the majority carriers that cross the potential barrier and the minority carriers that fall down the potential barrier.

Differently from the reverse bias condition, now we are reducing the potential barrier and the number of holes and electrons able to jump over increases exponentially.

$$I = I_S \left(e^{\frac{qV}{kT}} - 1 \right) = I_S\, e^{\frac{qV}{kT}} - I_S$$

The mean current is the difference of the two components, but the fluctuations are quadratically added.

The statistic fluctuations of the noise can be either positive or negative with the same probability and, in order to take into account the variations, we square and then we sum the contributions.

$$S_{nU}(f) = 2qI_S e^{\frac{qV}{kT}} + 2qI_S = 2q(I + I_S) + 2qI_S = 2qI + 4qI_S$$

In forward bias I>>Is because the barrier is reduced, and the spectrum is

$$S_{nU}(f) \approx 2qI$$

At zero bias I=0 the minority and the majority current are equal, so the spectrum is

$$S_{nU}(f) \approx 4qI_s$$

4.4 Noise in Resistors (Johnson-Nyquist noise)

The Johnson noise, also known as thermal noise, is generated by the Brownian motion of the carriers. It is mainly present in resistors and transistors.

If we analyse a floating resistor with both terminals open, we see that the voltage difference on the terminals is not zero, it varies from time to time. Only the average

value is equal to zero. The variation of the voltage value is due to the electrons moving randomly. In a classical model we have the electrons in the conductance band completely free to move within the conductor volume. It may happen from time to time that we have more electrons towards a terminal, making its potential to decrease, and fewer electrons onto the other terminal thus making its potential to increase. The voltage varies within a time interval t_c equal to tens of femtoseconds, which is the collision time of the carriers. Therefore, we can consider the noise to be flat up to a frequency equal to $1/t_c$, so in the order of THz.

The voltage is a statistical variable with a Gaussian probability distribution.

For our pourpose we just need to focus our attention on the power and on the spectral density of the noise. In order to model a noisy resistor, we put an equivalent voltage generator in series to a noiseless resistor, and the voltage generator power equals the power spectral density of the resistor.

The voltage spectral density is

$$S_{vB}(f) = 2kTR$$

Sometimes it is useful to deal with currents, so we can also model the noise with an equivalent current generator in parallel to the resistor, its value is simply given by

$$\frac{S_{vB}(f)}{R^2}$$

This is equivalent to move from a Thevenin equivalent circuit to a Norton equivalent circuit; the only difference is that we are dealing with power spectral density and therefore the transfer functions are squared.
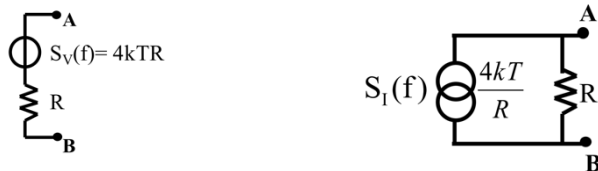

Figure 4.6 Equivalent resistors with voltage and current noise generator

4.5 White Noise

By definition, white noise is a particular kind of noise that has a constant power spectral density.

The term white is a recall to the white visible light, which is given by the sum of all the frequencies of the visible range with the same weight.

The autocorrelation function of the white noise is a δ in the origin. This is because the autocorrelation is the inverse Fourier transform of the power spectral density.
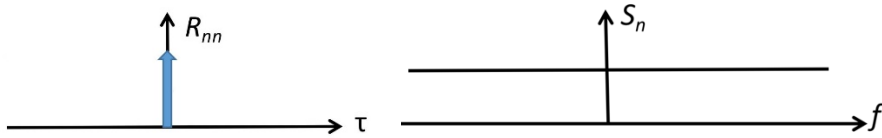
Figure 4.7 White noise autocorrelation function and power spectral density

In real world, an ideal white noise does not exist. Indeed, analysing it in the frequency domain we know that the integral of the power spectral density from minus infinite to plus infinite is the noise power, and clearly for the ideal white noise this integral does not give a finite result. This means that the power noise is infinite and therefore the noise has an infinite rms value which will give in any case an SNR equal to zero. At the same way in the time domain we know that the noise power is related to the value in zero of the autocorrelation function, but with the ideal white noise the autocorrelation is a delta, and so, again, its power results infinite.

In the real world we will never have an ideal white noise, since we will always have a high frequency cut off.

The autocorrelation time will be related to the current pulse duration. If we consider the shot noise with the pulse duration of hundreds of ps, we obtain an upper frequency of tens of GHz and we could consider our noise as white.

For the Johnson-Nyquist noise the pulse lasts tens of fs, so the upper frequency is hundreds of THz and again we can consider the noise white.

In general, we will consider *the noise white if the upper limit is much higher than the frequency we are considering in our system*, or, in the time domain, if the width of the noise autocorrelation is much shorter than the time scale we are working with.

Differences between stationary and non-stationary noise

In the previous chapter we have analysed the difference between stationary and non-stationary noise.

In the case of non-stationary white noise, we still have a delta autocorrelation function, but the power is depending on the time of measurement t_m, so the S_b value of the power spectral density varies with t_m.

For the properties of the Fourier transform the value in the origin of the power spectral density is the integral of the autocorrelation function.

It follows that the area of the autocorrelation function changes with time.

$$R_{nn}\left(t, t + \tau\right) = S_b\left(t\right) \cdot \delta\left(\tau\right)$$

Noise analysis and simulation

In this chapter, the concept of ideal white noise will be introduced and discussed both in time and frequency domain. We will see that white noise cannot exist in the real world, but it can be extremely useful to approximate very wide band noise. Finally, an example of noise filtering in the discrete-time domain is reported.

5.1 Introduction

In the previous chapters, noise has been discussed in terms of mathematical analysis and characteristics of different types of noise with their physical sources.

Now, we want to go deeper into details of a particular type of noise: the white noise. Starting from an ideal case of infinite-band white noise, we will then discuss the more realistic case of wide-band noise, describing its characteristics and filtering methods

5.2 White noise

We will divide the analysis of the ideal white noise into two steps: first of all, we will analyze the stationary, white noise and then we will extend the definition to the non-stationary noise case.

Noise is called stationary if its probability distribution is constant over time. On the other hand, the probability distribution of non-stationary noise is time-dependent.
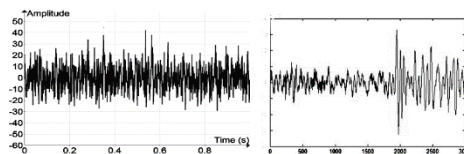
Figure 5.1 Stationary and non-stationary noise waveforms

Ideal **stationary** white noise is a concept derived from Johnson noise and shot noise that can be completely described by its essential feature: autocorrelation is nil at any time distance τ, other than zero, between two samples.

We can define it also as a random signal having equal intensity at different frequencies which means that white noise features a constant power spectral density. As a result, the autocorrelation of the noise can be described as a delta function centered in zero with area equal to S_b (bilateral spectral density)
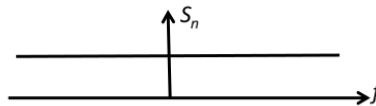
$$R_{nn(\tau)} = S_b\delta(\tau)$$
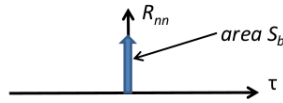
Figure 5.2 Ideal stationary white noise autocorrelation

Being the power spectrum of a stationary noise, by definition, the Fourier transform of the autocorrelation, the transform of $R_{nn}(\tau)$ yields a constant spectrum with magnitude S_b

$$S_n(f) = S_b$$

Figure 5.3 Ideal stationary white noise power spectrum

We can already see why this is an ideal case: since the spectrum is constant with no frequency limit, this noise has a divergent power. The relation between power spectrum of a signal x(t) and its power is

$$P = \int_{-\infty}^{+\infty} S_x(f)dx$$

That's why constant unlimited spectral density means divergent power. To extend the concept to the more general case of **non-stationary** noise, we have to go back to the definition of ideal white noise: it is a type of noise featuring no correlation between any two samples spaced by any finite amount of time. In this case, the main difference with respect to stationary noise relies in the noise intensity, that is not constant anymore, but it depends on time t. As a result, the autocorrelation function of ideal white non-stationary noise is still a delta, but its area is time-dependent:

$$R_{nn}(t, t + \tau) = S_b(t)\delta(\tau)$$

5.3 Wide band noise

In the real case, noise cannot have infinite power and thus it cannot have a constant power spectrum; in order to properly describe real white noise, we need to remove some idealities from our previous model.

Real white noise features very ***small width of autocorrelation*** and so a ***very wide (but not infinite) band*** with constant spectral density S_b. Practically speaking, noise is well approximated by ideal white one when it is correlated only at a time distance shorter than the minimum time interval of interest in the system.

It is worth noting that in real cases, an upper bandwidth limit always exists. Indeed, bandwidth limitations arise from physical constraints of the system or due to low pass filtering enforced by the circuitry.

Insight "Resistor bandwidth limitation"

To understand the intrinsic bandwidth limitations that affect any real system, let's consider one of the most ideal components we know: the resistor.

Ideally, a resistor shouldn't change its electrical behavior no matter the frequency of the signal we are applying. However, this is just an ideal scenario since different types of real resistors (and thus different construction structures and techniques) lead to different frequency responses due to parasitic capacitances and/or inductances. A few examples are reported in the following to give the reader an insight into the causes of bandwidth limitations.

- Carbon-composition resistors

A tubular case is filled with a mixture of carbon granules and a filler material; the ratio between these two components defines the resistive values. They act as pure resistors at frequencies up to the Megahertz (MHz) range. Such a high bandwidth extension poses a limitation to their value, that has to be at most equal to a few tens of kilo Ohm.

- Film-type resistors

Made as rods of a ceramic material (insulator) coated with a metal, the resistance value of film-type resistors is controlled through the thickness of the metal deposit. Most of these resistors have a spiral structure that makes them behave as inductors at frequencies in the MHz range. Some film-type resistors do not have a spiral structure, such as surface mount resistors: the provide a purely resistive behavior up to hundreds of MHz.

- Wire-wound resistors

They are a coil of wire coated with insulator and the resistance value is controlled acting on the length of the wire itself. Due to its structure, the parasitic inductance is

pretty high. This is confirmed by the frequency response that shows deviation from ideality already around hundreds of Hz.

The **Lorenzian spectrum** is a typical real case of wide-band noise; its spectrum is limited by a single pole with time constant T_p (pole frequency $f_p=1/(2\pi T_p)$).
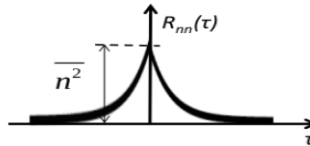
$$R_{nn}(\tau) = \bar{n}^2\, e^{-|\tau|/T_p}$$

Figure 5.5 Lorenzian autocorrelation
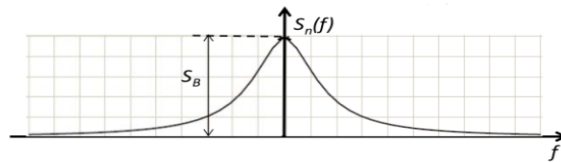
$$S_n(f) = \frac{S_b}{1 + \left(2\pi f T_p\right)^2}$$

Figure 5.6 Lorenzian power spectrum

5.4 Basic parameters of white band noise

In many cases, we can use a simplified description of a wide-band noise retaining its main features in order to make the calculations easier. This comes in handy especially when the filter's autocorrelation time width is comparable to the noise one. With a Lorentzian noise spectrum featuring a non-negligible autocorrelation time width, for example, we need an approximation that leads us to a result close to the real one to make a first order estimation. In this case, we cannot consider the noise as ideal while the correct calculation with the exponential shape can make the computation time quite long.

The two important characteristics of the real case that we want to preserve in the approximated versions of the autocorrelation $R_{nn}(\tau)$ and of the spectrum $S_n(f)$ are the mean square n^2 and the spectral density S_b.

In the time domain we want a small autocorrelation width, but instead of using a delta function we can approximate $R_{nn}(\tau)$ as a triangular function with width at half maximum equal to $2T_n$. The following constraints must be preserved:

- Same mean square noise: $R_{nn}(0) = \bar{n}^2$

- Same spectral density: in time domain, it means having the area of Rnn(τ) equal to Sb. This can be expressed as $2T_n * \bar{n}^2 = S_b$

We can thus derive the autocorrelation width as $\Delta\tau = 2T_n$.
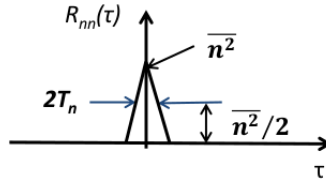

Figure 5.7 Approximated autocorrelation

In the frequency domain, we wish to retain a spectrum having broad bandwidth and a constant amplitude. We can thus approximate $S_n(f)$ with a rectangular function, featuring a bandwidth limitation at f_n, and bound to the following constraints:

- Same spectral density of the real case: $S_n(0) = S_b$.
- Same mean square noise: the area of $S_n(f)$ in the frequency domain must be equal to $\bar{n}^2$. This can be expressed as $S_b * f_n = \bar{n}^2$.

We can thus derive the noise bandwidth as $\Delta f = 2f_n$.
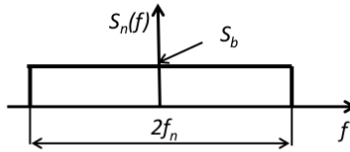
Figure 5.8 Approximated power spectrum

The approximation is consistent since $\Delta\tau\Delta f = 1$ as the expression that relates the power spectrum and the autocorrelation function requires $S_n(f) = F[R_{nn}(\tau)]$

Following this discussion, we can derive a **simplified description of the Lorentzian spectrum** with triangular autocorrelation or a rectangular spectrum (depending on the final use).

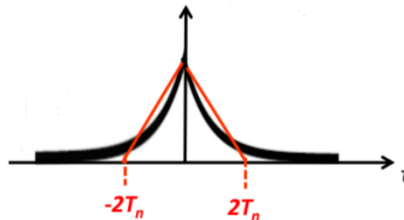The autocorrelation should have $\Delta\tau = 2T_p$ and thus $T_p = T_n$.

Figure 5.9 Approximated Lorenzian autocorrelation (in red) and real (in black)

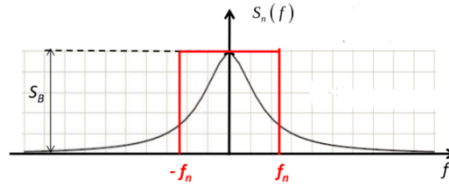In the frequency domain, the spectrum has a constant amplitude S_b up to $f_n = 1/4T_p$

Figure 5.10 Approximated Lorenzian spectrum (in red) and real (in black)

5.5 Foundation of white noise filtering

As seen in the previous chapters, noise can be seen as the superposition of elementary pulses, each one characterized by its own waveform (and therefore its Fourier transform). The type of noise elementary pulses contributes to determine the noise autocorrelation function, its spectrum and thus the type of noise.

The filtering action of a linear constant-parameter filter modifies the elementary pulse shape and so it can modify the noise properties too.

Noise filtering can be understood and evaluated by studying the filtering action on the elementary pulses.

Let's consider as an **example** the noise of a photodiode connected to an amplifier as shown in figure below.
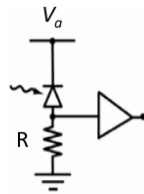

Figure 5.11 Circuit example schematics

A photodiode is basically a pn-junction whose current flow depends on the illumination of the device. It is typically biased applying a positive voltage between cathode and anode (i.e. reverse bias operation). In this condition, the light that is absorbed in the depleted region can release an electron-hole pair which contributes to the current flow through the device.

The device model usually includes a capacitor, which is physically justified by the depletion capacitance of a pn-junction (and the input capacitance of the amplifier) whose value can be approximated as

$$C = \varepsilon_0 \varepsilon_{Si} \frac{A}{W_d}$$

where A is the cross-sectional area of the device and W_d is the width of the depletion layer.

We can model the system as a current generator in parallel with the resistance R and the equivalent capacitance of the sensor as shown in figure 5.12
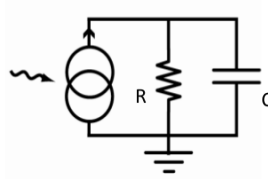

Figure 5.12 Equivalent circuit model

The current generator represents the diode current, which can be decomposed in the superposition of elementary current pulses $p_i(t)$. If the characteristic time T_h of the elementary pulses is small enough (with respect to the RC time constant), they can be approximated to delta-pulses.

$$qp_i(t) \simeq q\delta(t)$$
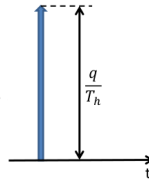
Figure 5.13 Elementary input pulse

Given a stream of current noise delta-pulses, the current flowing through the resistor R can be decomposed into elementary pulses with exponential shape, resulting from the presence of the RC low pass filter. The elementary output pulse is shown in figure 5.14.

$$f(t) = \frac{1}{T_f} e^{-\frac{t}{T_f}} 1(t)$$
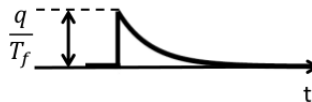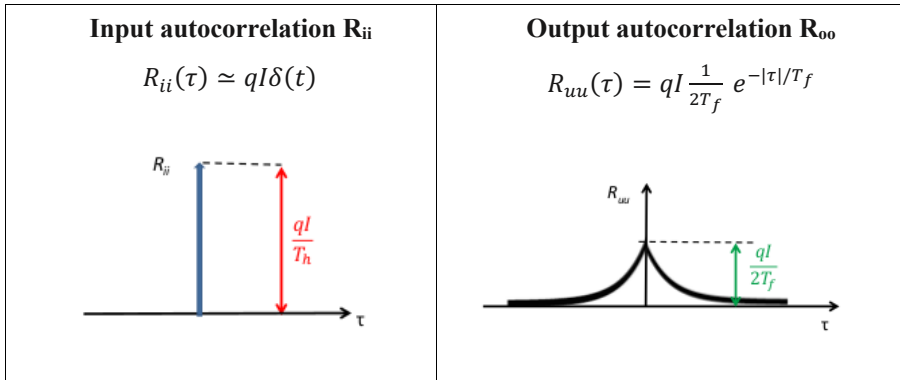
$$F(f) = \frac{1}{1 + j2\pi f\, T_f}$$

Figure 5.14 Elementary output pulse

In the time domain, in order to evaluate the mean square value of noise before and after the filtering action we need to consider the central values of the autocorrelation functions

Input autocorrelation R_{ii}	Output autocorrelation R_{oo}
$R_{ii}(\tau) \simeq qI\delta(t)$	$R_{uu}(\tau) = qI\dfrac{1}{2T_f}\,e^{-\|\tau\|/T_f}$

Figure 5.15 Input and output autocorrelation of the example

In the frequency domain, in order to evaluate the mean square value of noise before and after the filtering action we need to consider the areas of input and output spectral density

Input spectral density $S_i(f)$	Output spectral density $S_o(f)$
$S_i(f) = S_b$	$S_u(f) = qI\dfrac{1}{1+\left(2\pi f\,T_f\right)^2}$
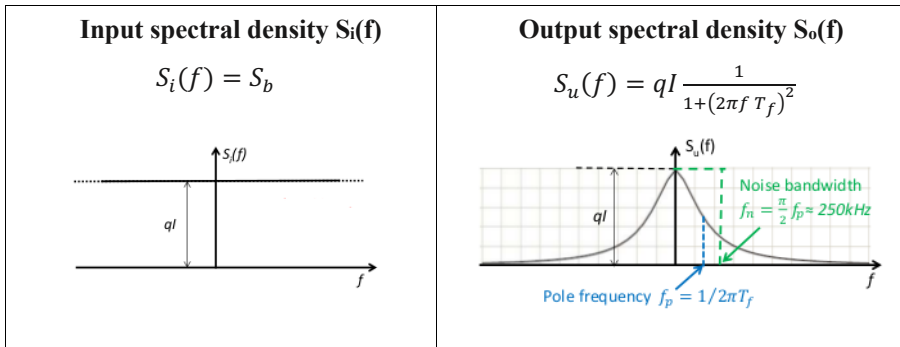
Figure 5.16 Input and output spectrum of the example

As expected, the presence of a single-pole low-pass filter limits the ideal input noise in frequency and produces an output noise featuring a Lorentzian spectrum.

05.6 Filtering white noise in discrete time

We now consider filtering in the digital domain, that is, we assume the time interval of interest to be discretized.
Calling T_s the sampling time, i.e. the time distance between two consecutive samples, we have a discrete time scale where

$$t_n = t_{n-1} + T_s$$

We will refer to the noise sample at time n as n_n.
Filtering consists into multiplying each sample of the noise by a given weight. The output is the sum of the weighted noise samples:

$$n_f = w_1 n_1 + w_2 n_2 + w_3 n_3 \ldots = \sum_N w_k n_k$$

Where N is the total number of collected samples.
We want to extract the output noise mean square value that can be written as

$$\overline{n_f^2} = \overline{w_1^2 n_1^2 + w_2^2 n_2^2 + \ldots + w_1 n_1 \cdot w_2 n_2 + w_1 n_1 \cdot w_3 n_3 + \cdots} =$$

$$= w_1^2 \overline{n_1^2} + w_2^2 \overline{n_2^2} + \ldots w_1 w_2 \overline{n_1 n_2} + w_1 w_3 \overline{n_1 n_3} + \ldots$$

If we consider the noise as ideal and uncorrelated (so with a delta autocorrelation) all rectangular terms are nil no matter how small the sampling time is.
Taking into account a real noise the rectangular terms can be eliminated only if the sampling time T_s is reasonably larger than the noise autocorrelation.
If a Lorentzian spectrum is considered, the sampling time T_s should be at least five times larger than the exponential decay time, in order to consider the samples uncorrelated; if an approximated Lorentzian spectrum is considered instead, T_s should be at least two times larger than the exponential decay time.

The cancellation of the square values leaves us just with

$$\overline{n_f^2} = \sum w_k^2 \, \overline{n_k^2}$$

If the noise is stationary the mean square noise in each time point is equal (we have no dependence on time) so the expression can be further simplified as

$$\overline{n_f^2} = \bar{n}^2 \sum w_k^2$$

As we will see in the following chapters, a similar simplification holds also for continuous time filtering.

CHAPTER 6

6

Filtering Signals

The presence of noise can make it difficult or even prevent the acquisition of a signal: in these cases, filtering is mandatory. A filter typically acts on both the signal and the noise. In this chapter, we will discuss how a filter can have an impact on the signal. Filters can be modeled by linear or non-linear equations (for the first kind superposition of effects can be used) and their parameters can be constant or time-variant. In this chapter, we will only deal with linear filters, both constant-parameter and time-variant ones.

6.1 Introduction

If you want to measure with a certain precision a signal buried in noise, you need a system able to have a different effect on signal and noise, enhancing the first with respect to the second and thus increasing the signal-to-noise ratio. In this chapter, we will focus on the filtering effect on signals, analyzing simple examples that are useful to understand how to build the weighting function of a filter, that is, as its name suggests, the set of weights which must be applied to the signal in every time instant to determine the output.

6.2 Discrete-time and Continuous-time Signal Filtering

To better understand the concept of filtering, we can use digital/discrete systems which allow a simple approach that can be extended to the analog/continuous ones. So, let's consider this acquisition scheme in figure 6.1 where $y(t_m)$ is the acquired signal at a specific time t_m.
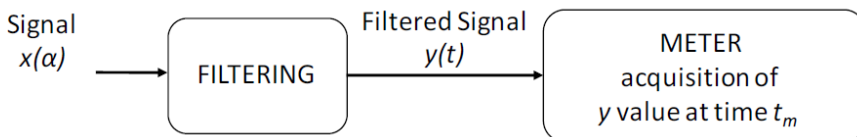
**Figure 6.1** Generic acquisition chain.

If the filter is linear, i.e. the elements inside of it are described by linear equations, the output of a linear combination of the inputs is the linear combination of the output

obtained feeding to the filter the same inputs singularly. In other words, the output is a weighted sum of input values x taken at various time t with weights that do not depend on the input x. In discrete time filtering this sum is:

$$y(t_m) = w_1 x(\alpha_1) + w_2 x(\alpha_2) + w_3 x(\alpha_3) + \cdots + w_n x(\alpha_n) = \sum_{k=0}^{n} w_k x(\alpha_k)$$

$$y(t_m) = \sum_{k=0}^{n} w_k x_k$$

On the other hand, if we consider time-variant filters, the weights can change over time and so will the shape of the filter. Note that also in this case they aren't a function of the input signal, otherwise we would talk about an *adaptive filter*, which is a completely different matter.
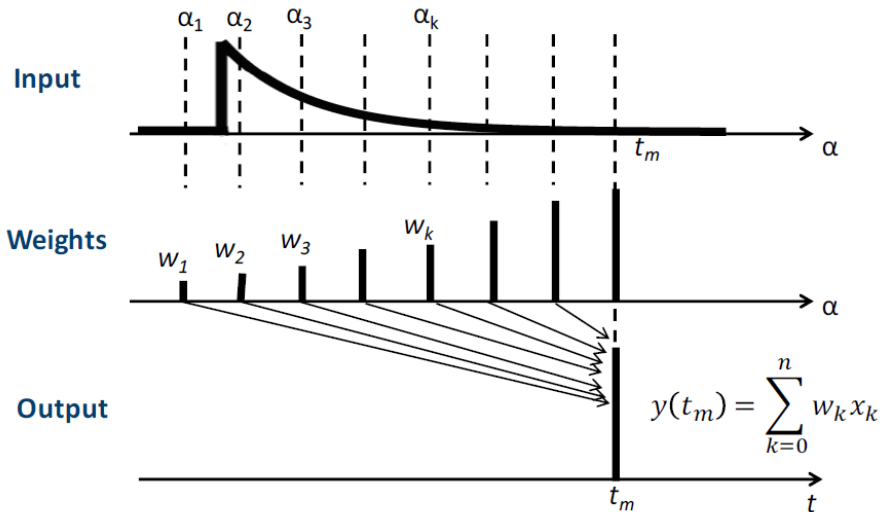
Figure 6.2 Weights for every sample of the input and calculation of the output for a discrete-time system.

To understand the meaning of "weight", we can use the example in figure 6.2, where the signal is decaying exponentially and the weights are as shown (as we can imagine, this is not a good weighting function because it'd be better, intuitively, to have a higher weight where the signal is high and a lower weight where the signal is low, as will be clearer later in this book). To obtain the desired output value, we can multiply the samples with the weight and add them together. Now, if we want to extend the same concept to the analog case with a continuous-time weighting function, instead of a sum, we have an integral. This is necessary because we have an infinite number of samples and weights as shown in figure 6.3.
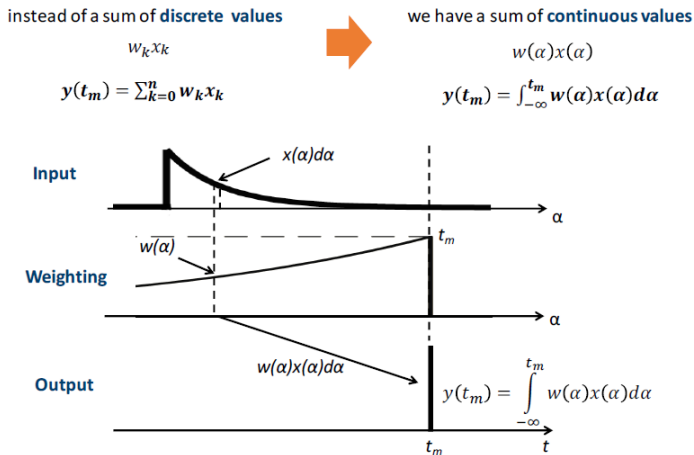
instead of a sum of **discrete values**

$$w_k x_k$$

$$y(t_m) = \sum_{k=0}^{n} w_k x_k$$

we have a sum of **continuous values**

$$w(\alpha)x(\alpha)$$

$$y(t_m) = \int_{-\infty}^{t_m} w(\alpha)x(\alpha)d\alpha$$

Input

$x(\alpha)d\alpha$

α

Weighting

$w(\alpha)$

t_m

α

Output

$w(\alpha)x(\alpha)d\alpha$

$$y(t_m) = \int_{-\infty}^{t_m} w(\alpha)x(\alpha)d\alpha$$

t_m

t_m t

Figure 6.3 Weighting function and evaluation of the output for a continuous-time system.

The **weighting function** (also called memory function) of a filter assigns a weight $w(\alpha)$ to each element $x(\alpha)d\alpha$. Why is it called memory function?
Because the output y depends on every element of the input signal that entered the filter: the system has memory of the past. We will see in the following paragraphs how to build this function starting from the impulse response of the filter.

6.3 Constant Parameter Filters

As we said, constant parameters filters have a response that is invariant with respect to a translation along the time axis: if the same signal x is applied to the filter at two different time instants, t_m and $t_m + \alpha$, the output is the same but translated in time by the same quantity ($\Delta t = \alpha$).
We will now discuss how to calculate the weighting function of a constant parameter filter. First of all, it is necessary to compute the impulse response of the filter.
In particular, we need to know the effect of a delta applied at a generic instant α_1 on the output measured at t_m. As shown in figure 6.4, time instants α_1 and t_m are both referred to the same arbitrary origin on the time axis, but the relevant parameter is the time distance between them. (A more complete expression of the weighting function should be w(α, t), where both α and t can vary).
How can we derive the weighting function? If we look at the top of figure 6.4 we see that a delta applied at time α_1 gives a contribution to the output which is equal to its delta response at t_m, that is $h(t_m - \alpha_1)$. Well, it is quite clear now that $h(t_m - \alpha_1)$ corresponds to the weight that the filter gives to a delta applied at α_1. The same holds for any delta applied at different time distances with respect to t_m.
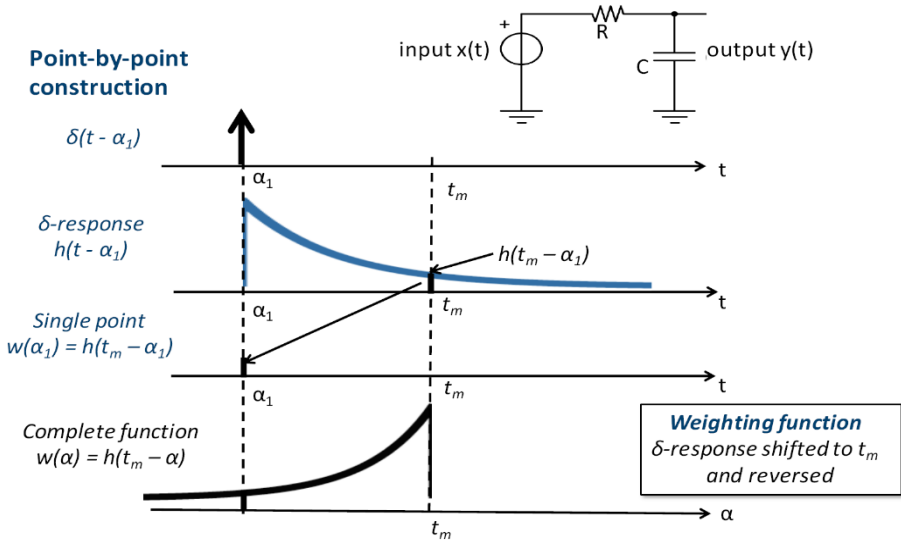We have seen in chapter 2 that any input signal can be written as the superposition of deltas.

Figure 6.4 Graphic representation of the weighting function construction

Therefore, if we know the delta response of the filter, we can calculate the weights that the filter applies to any delta that composes the input signal on the basis of the time distance between the input delta itself and the measurement time. As a result, the complete weighting function of a constant parameter filter is its delta response, reversed in time and shifted to t_m.

Once t_m is fixed, the weighting function is unique; it has always the same shape and the same position with respect to t_m. In other words, when t_m is changed, w(α) changes in a very simple way: *"it walks with t_m just like a tethered dog follows his boss"*. This is shown in figure 6.5.
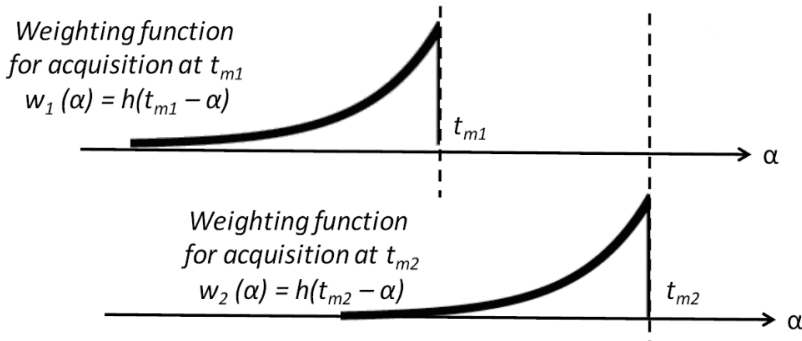
Figure 6.5 Weighting function of a constant parameter filter at two different acquisition times.

So, for constant-parameter filters (but not for time-variant filters):

$$w(\alpha, t_m) = h(t_m - \alpha) \quad \forall t_m$$

The result is confirmed analytically. The weighting function is defined by:

$$y(t_m) = \int_{-\infty}^{t_m} x(\alpha)w(\alpha, t_m)d\alpha \quad \forall t_m$$

But for constant parameter filter we also know that the output at time t_m is the convolution between the input $x(t)$ and the impulse response $h(t)$ evaluated in t_m:

$$y(t_m) = x(t) * h(t)|_{t_m} = \int_{-\infty}^{\infty} x(\alpha)h(t_m - \alpha)d\alpha \quad \forall t_m$$

Since h(t-α) is defined only for α<tm the integral can be rewritten as:

$$y(t_m) = \int_{-\infty}^{t_m} x(\alpha)h(t_m - \alpha)d\alpha \quad \forall t_m$$

From the equations above, therefore, it is:

$$w(\alpha, t_m) \equiv h(t_m - \alpha)$$

As a result, constant-parameter filters are completely characterized in the time domain by the impulse response to the Dirac delta $h(t)$.
It is important to remember that this result is not valid for time-variant filters.

6.3 Time-variant Linear Filter

A time-variant filter is a filter whose parameters are not constant over time. The easiest way to change the filter parameter over time is having a switch in the circuit. We can consider a simple RC circuit with a switch, that is controlled by an external signal, as an example. By changing the switch position (DOWN=switch closed, UP=switch open), we are changing the configuration of the filter over time and the delta-response of the filter will also change. For example, a delta applied at the input when the switch is open will have no effect on the output.
Now, we will derive the weighting function using the same approach that we used for constant parameter filters. Let's consider a time t_m and analyse the effect of a delta on the output. Two different cases have to be distinguished: when the switch is open, the input signal cannot affect the circuit output, which remains constant; when the switch is closed the circuit behaves as a constant parameter filter featuring the
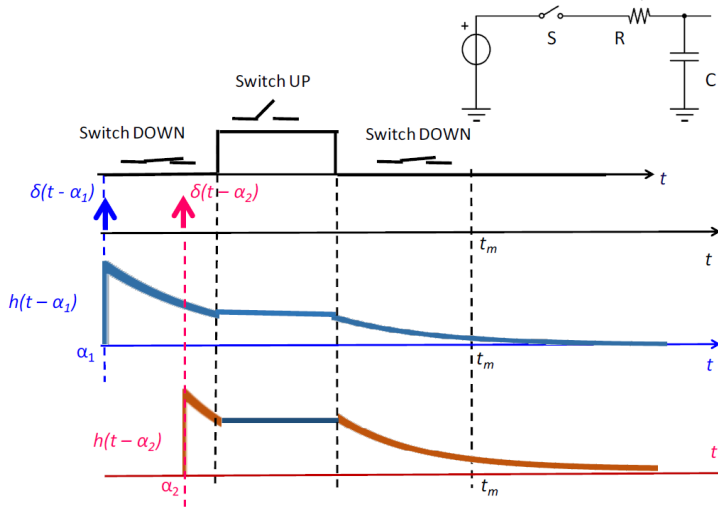
Figure 6.6 δ-response of time-variant filter consisting in a RC network and a switch in series to the input. The δ-response is not unique, but it strongly depends on the circuit configuration over time.

same delta response that we have seen in the previous paragraph (see figure 6.4). Figure 6.6 shows the response to a delta that is applied when the switch is closed and a double switching transition (DOWN-UP-DOWN) occurs afterwards.

Starting from this result, we can calculate the weighting function, which strongly depends on the measurement time. Two graphical examples showing how to derive the weighting functions of time-variant filters are reported in figure 6.7 and 6.8. Note that when the switch is UP, $w(\alpha)$ is zero.
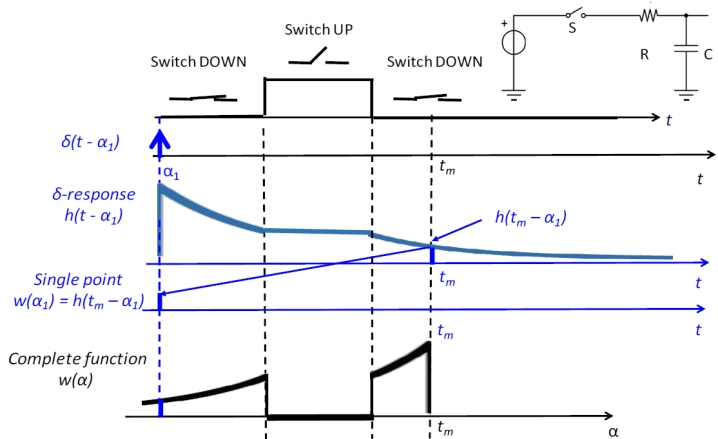
Figure 6.7 Weighting function of a time-variant filter at a certain measurement time following a DOWN-UP-DOWN sequence of the switch.
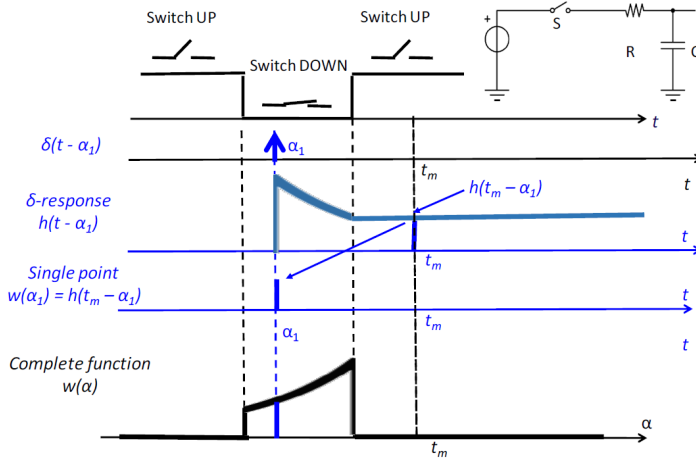
Figure 6.8 Weighting function of a time-variant filter at another measurement time following a UP-DOWN-UP sequence of the switch.

6.7 Weighting in the Frequency Domain

Sometimes it can be convenient to change from time domain to frequency domain to make calculations easier. This is valid also for the weighting functions. The idea of considering $y(t)$ as a weighted sum of components can be extended to the frequency domain. Let's consider Parseval's theorem:

$$\int_{-\infty}^{+\infty} a^2(t)dt = \int_{-\infty}^{+\infty} A(f)A^*(f)df = \int_{-\infty}^{+\infty} A(f)A(-f)df$$

Which can be extended to the product of two functions $a(t)$ and $b(t)$

$$\int_{-\infty}^{+\infty} a(t)b(t)dt = \int_{-\infty}^{+\infty} A(f)B^*(f)df = \int_{-\infty}^{+\infty} A(f)B(-f)df$$

Denoting by $W(f) = F[w(t)]$ we have:

$$y(t_m) = \int_{-\infty}^{+\infty} x(\alpha)w(\alpha)d\alpha = \int_{-\infty}^{+\infty} X(f)W(-f)df$$

In general, the value y acquired at the time t_m at the filtering system output can be considered either as a weighted sum of instantaneous input values $x(t)$ with weights $w(t)$ or as a weighted sum of Fourier components $X(f)$ of the input signal $x(t)$ with weights $W(-f) = F[w(-t)]$.

6.8 Summary

- For constant-parameter filters (and only for them) the weighting function is simply the δ-response function reversed and shifted in time.
- That's NOT true for time-variant linear filters, which do not have a unique δ-response. The shape of the δ-response depends on when the δ-function is applied to the input during the evolution in time of the filter.
- The weighting function in linear time-variant filters may be difficult to compute, but it always exists.
- For filters that vary in time with a simple law, it is fairly simple to compute the weighting function at a given measurement time, in particular for switched-parameter filters.
- Switched-parameter filters undergo abrupt changes at the transition from a time interval to the next one, but within each interval, their parameters stay constant.
- The values of electrical variables (voltages, currents) before and after switching must be carefully checked because they can be discontinuous, i.e. they may exhibit abrupt variations at the switching time.

Filtering noise

In the previous chapters we acquired some basic tools to study signals and noise, such as the autocorrelation function and the spectral density. In chapter 6 we studied how to filter the signal; now let's see how to apply these tools to calculate the output noise of a generic filter.

7.1 Introduction

The aim of this chapter is the calculation of the output noise from a system; this will allow us to study also the output noise from any filter.

The dissertation can be extremely complex; therefore, after an introductive dissertation of the noise output of a generic system, we will simplify the general case to get closer to more practical situations. We will then focus on stationary noise and we will see how the result is modified when the noise is also white.

Finally, we will see how to further simplify the treatment in the case of systems/filters with constant parameters.
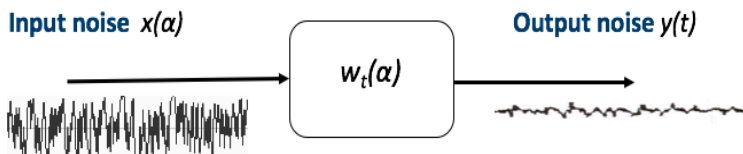

Input noise $x(\alpha)$ $w_t(\alpha)$ **Output noise $y(t)$**

Figure 7.1

Let's start from the characteristics of noise. We have seen how the autocorrelation function is a key tool to obtain important information such as noise density and so its total power. Referring to figure 7.1 and following the definition, we can write the autocorrelation function of both the input and the output noise of our system.

$$R_{xx}(\alpha, \alpha + \gamma) = \overline{x(\alpha)x(\alpha + \gamma)}$$

$$R_{yy}(t, t + \tau) = \overline{y(t)y(t + \tau)}$$

However, recalling the results obtained in the last chapter, we can write the autocorrelation of the output as a function of the input, obtaining:

$$R_{yy}(t_1, t_2) = \overline{y(t_1)y(t_2)} =$$

$$= \overline{\int_{-\infty}^{\infty} x(\alpha)w_1(\alpha)d\alpha \cdot \int_{-\infty}^{\infty} x(\beta)w_2(\beta)d\beta} = \iint_{-\infty}^{\infty} \overline{x(\alpha)x(\beta)} \cdot w_1(\alpha)w_2(\beta)d\alpha\,d\beta =$$

$$= \iint_{-\infty}^{\infty} R_{xx}(\alpha, \beta)\,w_1(\alpha)w_2(\beta)d\alpha\,d\beta$$

by setting in evidence the intervals of autocorrelation at the input

$$\gamma = \beta - \alpha$$

and at the output

$$\tau = t2 - t1$$

the autocorrelation function can be expressed as

$$R_{yy}(t_1, t_1 + \tau) = \iint_{-\infty}^{\infty} R_{xx}(\alpha, \alpha + \gamma)\,w_1(\alpha)w_2(\alpha + \gamma)d\alpha\,d\gamma$$

An extremely important parameter is the power of noise, so we can write the mean square noise at time t_1:

$$\overline{y^2(t_1)} = R_{yy}(t_1, t_1) = \iint_{-\infty}^{\infty} R_{xx}(\alpha, \alpha + \gamma)\,w_1(\alpha)\,w_1(\alpha + \gamma)d\alpha\,d\gamma$$

These equations are valid for **all cases** of noise and linear filtering, that is, also for non-stationary input noise and for time-variant filters.

7.2 Filtering stationary noise

The equations we just found change and can be simplified in case of stationary noise, that is an extremely frequent condition.
When the noise is stationary, its autocorrelation function at the input only depends on the time interval

$$\gamma = \beta - \alpha$$

We can then simplify the input autocorrelation:

$$R_{xx}(\alpha, \alpha + \gamma) = R_{xx}(\gamma)$$

The output autocorrelation function can be rewritten as well

$$R_{yy}(t_1, t_1 + \tau) = \iint_{-\infty}^{\infty} R_{xx}(\gamma)\, w_1(\alpha) w_2(\alpha + \gamma) d\alpha\, d\gamma =$$

$$= \int_{-\infty}^{\infty} R_{xx}(\gamma) \int_{-\infty}^{\infty} w_1(\alpha) w_2(\alpha + \gamma) d\alpha\, d\gamma$$

It is important to point out that we set, as condition, a stationary *input* noise. In this condition:
   a)   a constant parameter filter produces stationary output noise.
   b)   a time-variant filter can produce a non-stationary output noise!

Starting from the last equation we can try to further simplify it

$$R_{yy}(t_1, t_1 + \tau) = \int_{-\infty}^{\infty} R_{xx}(\gamma) \int_{-\infty}^{\infty} w_1(\alpha) w_2(\alpha + \gamma) d\alpha\, d\gamma$$

we can see how the last part of the equation is the definition of the crosscorrelation of the functions $w_1(\alpha)$ and $w_2(\alpha)$, which we will denote by $k_{12w}(\gamma)$

$$k_{12w}(\gamma) = \int_{-\infty}^{\infty} w_1(\alpha) w_2(\alpha + \gamma) d\alpha$$

As a result, we can derive two important formulas to study of the output noise, which are the output autocorrelation function:

$$R_{yy}(t_1, t_1 + \tau) = \int_{-\infty}^{\infty} R_{xx}(\gamma) \cdot k_{12w}(\gamma)\, d\gamma$$

and the mean square noise, where we just have to consider the autocorrelation of $w_1(\alpha)$

$$\overline{y^2(t_1)} = R_{yy}(t_1, t_1) = \int_{-\infty}^{\infty} R_{xx}(\gamma) \cdot k_{11w}(\gamma)\, d\gamma$$

The last equation gives us the mean square output noise in case of stationary input noise and for any linear filter, both constant-parameter and time-variant filters. Recalling that

$$F[k_{11w}(\gamma)] = |W_1(f)|^2$$

using the Parseval theorem extension, we can calculate the mean square output noise in the frequency domain obtaining

$$\overline{y^2(t_1)} = \int_{-\infty}^{\infty} S_x(f) \cdot |W_1(f)|^2 \, df$$

This is an extremely useful result, since it gives us, in an intuitive way, the effect of a filter on the total output noise power. In the integral we can in fact recognize the input noise spectral density and $|W_1(f)|^2$. If our goal is to reduce the total noise power, then we will have to design a filter aimed at minimizing this product.

7.3 Filtering white noise

We can further simplify the study of the output noise if we have a white stationary noise, i.e. with a constant intensity (power).
As definition of white noise, we know that its autocorrelation function can be written as

$$R_{xx}(\alpha, \alpha + \gamma) = R_{xx}(\gamma) = S_b \cdot \delta(\gamma)$$

Therefore, the equation of the output autocorrelation becomes

$$R_{yy}(t_1, t_1 + \tau) = S_b \int_{-\infty}^{\infty} w_1(\alpha) w_2(\alpha) d\alpha = \boldsymbol{S_b \cdot k_{12w}(0)}$$

and the output mean square value

$$\overline{y^2(t_1)} = S_b \cdot k_{11w}(0) = S_b \int_{-\infty}^{\infty} w_1^2(\alpha) d\alpha$$

As we can see from the last equation, the calculation of the autocorrelation of the weighting function, and in particular its value in zero, it is extremely important for the calculation of the output noise power.
As in the previous paragraph, we can obtain the mean square value also in the frequency domain exploiting the Parseval theorem.

We can then write

$$\overline{y^2(t_1)} = S_b \int\limits_{-\infty}^{\infty} |W_1(f)|^2 \, df$$

We could have obtained the same results exploiting also one of the Fourier transform properties that we recalled in chapter 2:
"the value in zero of the Fourier transform is the integral of the function in time, and the value in zero of the function in time is the integral of the Fourier transform"

$$x(0) = \int_{-\infty}^{+\infty} X(f) df$$

so we can obtain the same result about the mean square value recalling that

$$F[k_{11w}(\gamma)] = |W_1(f)|^2$$

7.4 Filtering noise with constant-parameter filters

A constant-parameter filter is completely characterized by the δ-response *h(t)* in the time domain and by the transfer function   *H(f) = F[h(t)]* in the frequency domain. In the previous chapters, we understood that the weighting function for an acquisition time *t_m* and the δ-response are simply related by:

$$w_m(\alpha) = h(t - \alpha)$$

and therefore, in frequency

$$|W_m(f)|^2 = |H(f)|^2$$

Let's now apply these two relations to the autocorrelation function of the output noise:

$$R_{yy}(t_1, t_2) = \iint\limits_{-\infty}^{\infty} R_{xx}(\alpha, \beta) \, w_1(\alpha) w_2(\beta) d\alpha \, d\beta$$

$$= \iint\limits_{-\infty}^{\infty} R_{xx}(\alpha, \beta) \cdot h(t_1 - \alpha) h(t_2 - \beta) d\alpha \, d\beta =$$

$$= \int\limits_{-\infty}^{\infty} h(t_1 - \alpha) d\alpha \int\limits_{-\infty}^{\infty} R_{xx}(\alpha, \beta) \cdot h(t_2 - \beta) d\beta =$$

$$= R_{xx}(\alpha, \beta) * h(\beta) * h(\alpha)$$

Taking into account that:

- the stationary input autocorrelation depends only on the interval $\gamma = \beta - \alpha$
- $d\beta = d\gamma$
- $d\alpha = -d\gamma$

the output autocorrelation is also stationary and depends only on the interval τ

$$R_{yy}(\tau) = R_{xx}(\gamma) * h(\gamma) * h(-\gamma) = R_{xx}(\gamma) * k_{hh}(\gamma)$$

and therefore

$$S_y(f) = S_x(f) \cdot |H(f)|^2$$

that give us the output noise spectral density as a function of the input noise spectral density and the Fourier transform of the delta response of our system.
As usual we can calculate the output mean square value both in the time domain and in the frequency domain, computing the autocorrelation of the noise:

$$\overline{y^2} = R_{yy}(0) = \int_{-\infty}^{\infty} R_{xx}(\gamma) \, k_{hh}(\gamma) d\gamma$$

and using the Parseval theorem or just integrating the output noise spectral density

$$\overline{y^2} = \int_{-\infty}^{\infty} S_x(f)|H(f)|^2 df$$

If the input noise is white, we finally have

$$R_{xx}(\gamma) = S_b \delta(\gamma)$$

and therefore

$$\overline{y^2} = S_b \, k_{hh}(0)$$

$$\overline{y^2} = S_b \int_{-\infty}^{\infty} |H(f)|^2 df$$

Constant-Parameter Low Pass Filters

As discussed in the previous chapters, the information of interest is typically accompanied by noise. In some cases, it is possible to select a filter that reduces the contribution of the noise without having a substantial impact on the signal, hence improving the signal-to-noise ratio. Low Pass Filters are a class of filters that can be used to achieve this goal. In this chapter, we will present and discuss some constant-parameters filters: the passive RC-integrator, the active RC-integrator and the Mobile Mean filter.

8.1 Introduction

The information of interest is carried out by the signal. Unfortunately, superimposed noise, which is typically due to the sensor physics and the following readout electronics, might be non-negligible causing a degradation or even an obscuration of the information itself. To mitigate the impact of noise, we need a filter able to exploit at best the differences between signal and noise, taking well into account what kind of information is to be recovered.

It's a quite common situation to have a low-frequency signal accompanied by a wide-band noise. In these cases, it is worth it to simply add a filter that cancels out everything above the maximum bandwidth of the signal, leaving untouched the rest of it. The filtering weight is concentrated in a relatively narrow frequency band from zero to a limited frequency, while above the band-limit it falls to negligible value, as depicted in figure 8.1. These kinds of filters are called "Low Pass Filters" (LPF).

Correspondingly, in the time domain the weighting function has relatively wide time-width, as well as its autocorrelation. So, in the time domain, the action of the filter is to produce a (weighted) time-average of the input over a certain time interval, delimited by the width of the weighting function.

8.2 The RC integrator

The easiest way to implement a constant-parameter low pass filter is to use a resistor in series to our signal followed by a capacitor towards ground as in the figure 8.2. The voltage across the capacitor will be a low pass filtered version of the input signal.
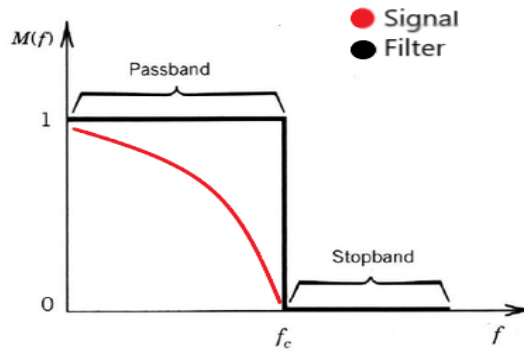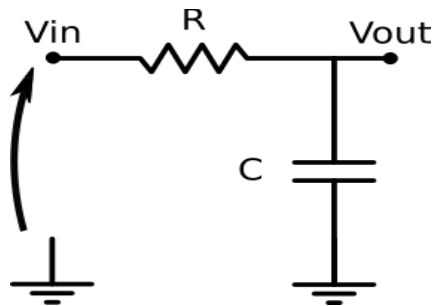
Figure 8.1 Ideal Low Pass filter.

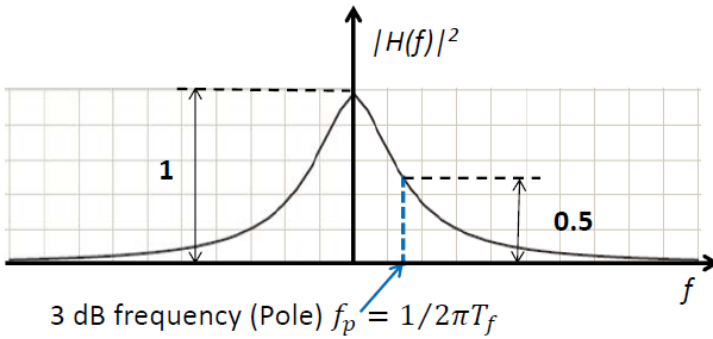**Figure 8.2** Basic scheme of a RC integrator.

In the Laplace domain, the output voltage can be calculated as follows:

$$V_{out} = V_{in} \frac{\frac{1}{SC}}{R + \frac{1}{SC}} = V_{in} \frac{1}{1 + SCR}$$
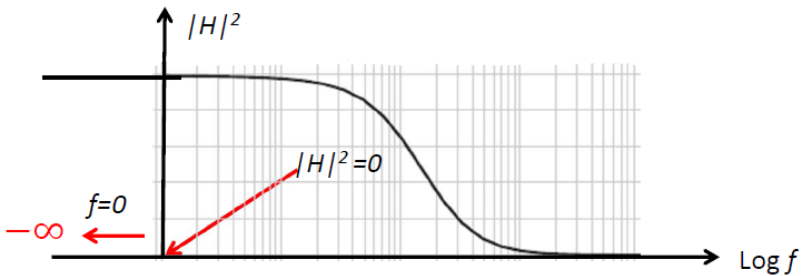
This result shows us that the transfer function is
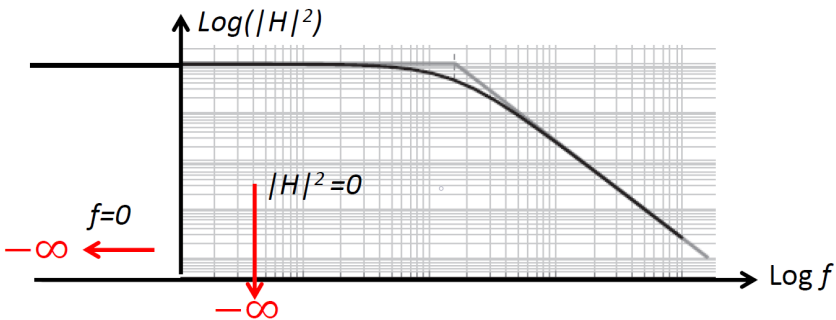
$$H(f) = \frac{1}{1 + j2\pi f T_f}$$

Figure 8.3 Plot of the squared module of the transfer function. Both the axes are in a linear scale. The pole occurs when the module is $1/\sqrt{2}$ of its original value, so when the square of the module is halved.

Figure 8.4 Plot of the squared module of the transfer function. Only the frequency axis is in logarithmic scale.

Figure 8.5 Plot of the squared module of the transfer function. Both the axes are in logarithmic scale. The grey one is the approximate version, the black one is the real one. After the pole, the function goes down with a slope of -20 dB/decade.

The transfer function confirms that a RC series network implements a low pass filter (at infinite frequency the transfer function is equal to zero) with a pole having a time constant T_f

$$T_f = RC$$

for $2\pi f T_f \ll 1$, i.e. for $f \ll f_p = \frac{1}{2\pi T_f}$, we have $H(f) \cong 1$, and then $V_{out} \cong V_{in}$.

for $2\pi f T_f \gg 1$, i.e. for $f \gg f_p = \frac{1}{2\pi T_f}$, we have $H(f) \cong 0$, and then $V_{out} \cong 0$.

It is worth noting that a RC LPF does not feature a sharp transition between passband and stopband, but a relatively smooth one. To have an idea, the module of $H(f)$ at the pole frequency $(f = f_p)$ is $|H(f)| = \frac{1}{\sqrt{2}}$.

To understand the effect of this circuit on noise, we can easily plot $|H(f)|^2$ and compare it with the ideal one of a LPF. In figures 8.3, 8.4 and 8.5, the transfer function is plotted for different horizontal and vertical axes. Normally the student is used to plot this function in a Bode diagram, while in the next chapters it will be useful to be familiar with all of these representations.

In the time domain, the circuit behavior can be analyzed solving the differential equation of the circuit and finding the δ-response.

Being the current in the resistor equal to the one in the capacitor for the current Kirchhoff's law, and being

$$I_c = C \frac{dV_{out}(t)}{dt}$$

we get

$$\frac{V_{in}(t) - V_{out}(t)}{R} = C \frac{dV_{out}(t)}{dt}$$

that is

$$\frac{dV_{out}(t)}{dt} + \frac{V_{out}(t)}{RC} = \frac{V_{in}(t)}{RC}$$

To solve it we can use the Lagrange formula:

$$V_{out}(t) = e^{-\frac{t}{RC}} \left[V_{out}(t_0) + \int_{t_0}^{t} \frac{V_{in}(t)}{RC} e^{\frac{t}{RC}} dt \right]$$

So, setting the starting condition as $V_{out}(t_0) = 0\,V$ (which means that at $t_0 = 0$ the capacitor is discharged), we have a Cauchy problem with a unique solution.

To calculate the δ-response, we set $V_{in}(t) = \delta(t)$.
Since

$$\int_{-\infty}^{t} \delta(t)\, e^{\frac{t}{RC}}\, dt = \int_{-\infty}^{t} \delta(t) = 1(t)$$

we get, overall, the well-known result:

$$V_{out}(t) = h(t) = \frac{1}{RC}\, e^{-\frac{t}{RC}}\, 1(t)$$
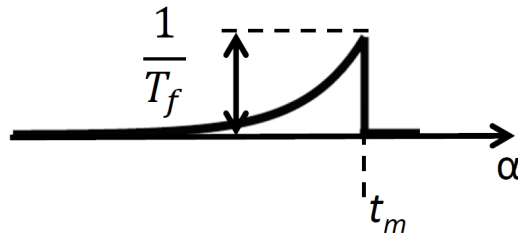

Figure 8.6 δ-response of the RC integrator.

It is worth noting a couple of properties of result:

• if the area of the input delta is unitary, also the area of the response is unitary. This is obvious reasoning in term of charges;

• V_{out} is the voltage across the capacitor and it experiences a sudden rise. Since the current in the capacitor is the derivative of the voltage across it, we expect "infinite current" into it, so "infinite current" into the resistor as well. This is not strange: the input signal has infinite amplitude as well, so it's theoretically able to provide all the current needed by the circuit. This mathematical model suggests that we should be careful whenever we feed to our circuit a finite amount of charge in a very small time, since we could burn some of the components. In real cases though, the capability of the sensor to provide charges is limited by the technology, so we can usually forget about it.

In conclusion, the δ-response is

$$h(t) = \frac{1}{RC}\, e^{-\frac{t}{RC}}\, 1(t)$$

The weighting function $w(t)$ of the filter is the flipped version of the δ-response, starting from t_m (i.e. the observation time).

Figure 8.7 Weighting function of the RC integrator.

It's quite clear now that the output will be an average of the input, where the weights decrease exponentially. It's also clear that it makes sense neglecting the value of the input signals that are too distant in time from the observation time. A good approximation is to consider the output as an average of the input over a time interval $\approx 2T_f$ preceding t_m (we will see later in this chapter why this approximation is a good one).

The step response of the circuit can be easily calculated by setting $V_{in}(t) = 1(t)$ in the Lagrange formula above. Then we'd get

$$Vout(t) = e^{-\frac{t}{RC}}\left[e^{\frac{t}{RC}} - 1\right] = 1 - e^{-\frac{t}{RC}}$$
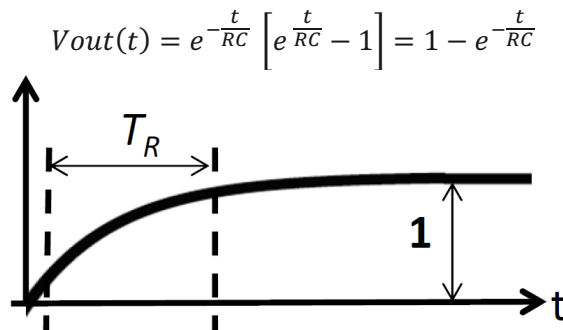
Figure 8.8 Step response of the RC integrator.

Here, the comments we made about the δ-response are no longer valid. In this case, we can point out that the asymptotic value of the output voltage is the amplitude of the input step. This was expected, since we have unitary DC gain.
This suggests that the step response could be also found by inspection. In fact, the asymptotic value is the input one, having unitary DC gain; at the step the capacitor prevents an abrupt variation of the voltage across its terminals (it acts as a short circuit on the signal), so we start from the initial value, i.e. 0V; eventually, we reach our target value with an exponential rise with a time constant that is the one of the pole. No need to solve the differential equation for that then.

We can also give an estimation of the output rise time (10-90% of its variation) that could useful to compare step responses of different filters.
We have to compute

$$1 - e^{-\frac{t}{RC}} = \frac{1}{10} \implies t = T_f \ln\left(\frac{10}{9}\right) \approx 0,1 \, T_f$$

and also

$$1 - e^{-\frac{t}{RC}} = \frac{9}{10} \implies t = T_f \ln(10) \approx 2,3 \, T_f$$

Therefore, the risetime will be

$$T_r = 2,3 \, T_f - 0,1 \, T_f = 2,2 \, T_f \approx \frac{1}{3 f_p}$$

Lastly, we could check how the transfer function that we found is indeed the Laplace transform of the δ-response.
In fact, we have

$$H(s) = \mathcal{L}\{h(t)\} = \int_0^{+\infty} \frac{1}{RC} \, e^{-\frac{t}{RC}} e^{-st} dt = -\frac{1}{RC} \frac{1}{s+\frac{1}{RC}} \left[e^{-\left(s+\frac{1}{RC}\right)t}\right]_0^{+\infty} = \frac{1}{1+SCR}$$

as expected.

Now we can evaluate the effect of the circuit filtering action on the noise, studying the circuit both in the time and in the frequency domain.
In the time domain, we first need the autocorrelation of the weighting function.
The autocorrelation will be

$$K_{ww}(t) = K_{hh}(t)$$

$$= \int_{-\infty}^{+\infty} h(\tau)h(t+\tau)d\tau = \frac{1}{(RC)^2} \int_{-\infty}^{+\infty} e^{-\frac{\tau}{RC}} sca(\tau) \, e^{-\frac{\tau+t}{RC}} sca(\tau+t)d\tau$$

$$= \frac{1}{(RC)^2} e^{\frac{t}{RC}} \int_{-\infty}^{+\infty} e^{-\frac{2\tau}{RC}} sca(\tau)sca(\tau+t)d\tau$$

For $t > 0$ the integral goes from zero to infinite.
So, we get

$$K_{ww}(t) = \frac{1}{(RC)^2} e^{\frac{t}{RC}} \int_0^{+\infty} e^{-\frac{2\tau}{RC}} d\tau = \frac{1}{2RC} e^{\frac{t}{RC}}$$
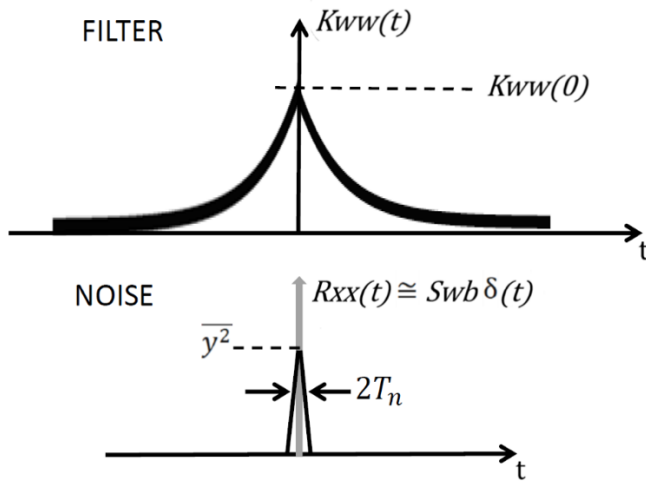
For t<0 instead it goes from t to infinite.
So, we get

$$K_{ww}(t) = \frac{1}{(RC)^2} e^{\frac{t}{RC}} \int_t^{+\infty} e^{-\frac{2\tau}{RC}} d\tau = \frac{1}{2RC} e^{\frac{t}{RC}} e^{-\frac{2t}{RC}} = \frac{1}{2RC} e^{\frac{-t}{RC}}$$

Putting everything together we can write

$$K_{ww}(t) = \frac{1}{2RC} e^{\frac{-|t|}{RC}}$$

In figure 8.9 (top) we can see the plot of this result.

Figure 8.9 Autocorrelation of the weighting function of the filter (top), compared with the autocorrelation of the noise (bottom).

Now, to find the RMS value of the noise at the output we have to compute

$$\overline{y^2} = \int_{-\infty}^{+\infty} K_{ww}(t) R_{xx}(t) dt$$

where $R_{xx}(t)$ is the autocorrelation of the noise.
Assuming wide-band white noise with bilateral spectral density S_{wb}, we'll have a time constant $T_n \ll T_f = RC$. This means that the autocorrelation of the noise can be considered delta-like. So, to compute the effect of the filter on it we only need the value of the autocorrelation of the filter in zero.

This approximation corresponds to saying that, in the frequencies of interest, the noise spectral density is substantially flat.
Therefore, we can write

$$\overline{y^2} \cong S_{wb} K_{ww}(0) = S_{wb} \frac{1}{2T_f}$$

In order to compare different filters from the noise point of view, it is extremely useful to introduce the definition of **noise equivalent bandwidth**: it is defined with reference to a white noise input S_b as the bandwidth value to be employed for computing **simply by a multiplication** the output mean square noise. We want that

$$\overline{y^2} = S_{wb} 2f_n$$

and so

$$S_{wb} 2f_n = S_{wb} K_{ww}(0) \implies f_n = \frac{1}{4T_f} = \frac{\pi}{2} f_p$$

This result can be verified in the frequency domain. In fact, it is

$$\overline{y^2} = \int_{-\infty}^{+\infty} S_{wb,out} df$$

$$= \int_{-\infty}^{+\infty} S_{wb} |H(f)|^2 df$$

$$\cong S_{wb} \int_{-\infty}^{+\infty} \left| \frac{1}{1 + j2\pi f T_f} \right|^2 df = S_{wb} \int_{-\infty}^{+\infty} \frac{1}{1 + (2\pi f T_f)^2} df$$

$$= S_{wb} \frac{1}{2\pi T_f} \left[arctg(2\pi f T_f) \right]_{-\infty}^{+\infty} = S_{wb} \frac{1}{2\pi T_f} \left[\frac{\pi}{2} - \left(-\frac{\pi}{2} \right) \right]$$

$$= S_{wb} \frac{1}{2T_f}$$

that gives the same result once it's equated with the definition of equivalent bandwidth provided above.
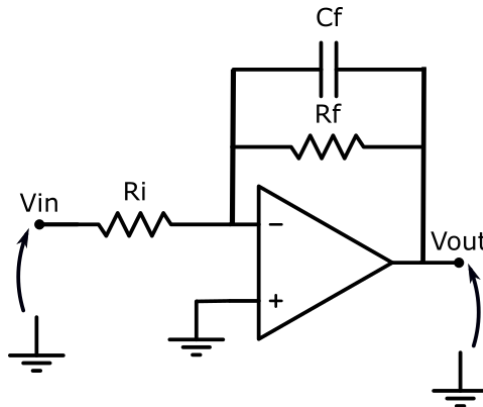
8.3 The active RC integrator

As already pointed out, the RC integrator seen so far has a unitary DC gain. This could be an issue whenever we want to combine the filtering action with a gain stage. This might be the case when we are in the early stages of the readout chain, where we want to have some gain in order to make the noise of the following stages completely negligible.
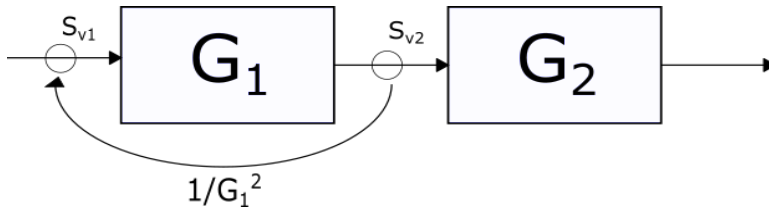
Figure 8.10 Block scheme of a generic readout chain. The input generator of a given stage can be brought back to the input dividing it by the overall gain of the previous stages. It's clear that, the higher the gain, the more negligible this noise generator will be.

N.B. The gain of a stage doesn't change the S/N of it though. In fact, both the signal and the RMS value of the noise will be multiplied by the same factor G. Overall, the S/N would remain untouched.

A stage with a certain gain can be achieved with active components, such as the operational amplifier. A low pass filter can be done easily adding a capacitor in the feedback network of an inverting stage, as shown in figure 8.11.

Figure 8.11 Active RC integrator.

All the rigorous steps done for the simple RC can be repeated here, with very few differences. Instead of repeating it, we'd rather discuss the circuit with an asymptotic analysis. The outcome will be the same.

At DC, that is an inverting stage with a gain

$$G(0) = -\frac{R_f}{R_i}$$

At high frquencies, when the C acts pretty much as a short circuit, we have the inverting input of the amplifier connected to the output in a buffer configuration. The non-inverting pin is instead grounded, so we expect the output voltage to be stuck at zero as well for every input we have. So, it's

$$G(\infty) = 0$$

It's clear that we have a single pole, while the zero is at infinite frequency. To find the pole we should find the equivalent resistance seen by the capacitor. We could simply consider that the feedback fixes the inverting pin at virtual ground, so no current can flow there. So, the current can only flow inside the resistor in parallel, R_f. It's

$$T_f = R_f C_f => f_p = \frac{1}{2\pi R_f C_f}$$

Overall, we have the same transfer function of the simple RC integrator, with an extra gain at DC.

$$H(s) = -\frac{R_f}{R_i} \frac{1}{1 + sT_f}$$

N.B. That's an inverting stage, hence the minus. It doesn't really change anything, provided we're still in the dynamic range of the circuits.

The δ-response will be the same as above, with a value in zero magnified by a factor equal to the DC gain, i.e.
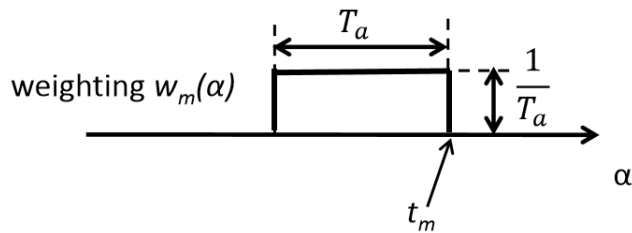
$$h(t) = -\frac{R_f}{R_i} e^{-\frac{t}{R_f C_f}} 1(t)$$

Figure 8.12 δ-response of the active RC integrator. If we consider the voltage axis positive going up, this waveform should actually be flipped.

8.4 The mobile-mean low-pass filter

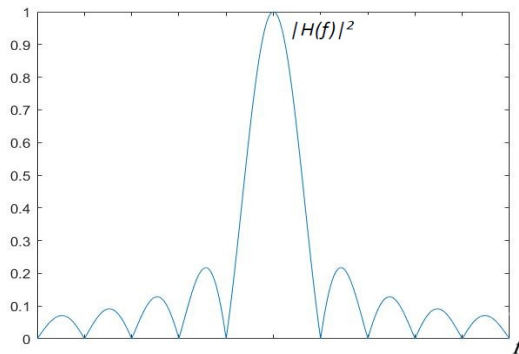A mobile-mean filter (MMF) is a filter with a weighting function like the one depicted in figure 8.13

Figure 8.13 Weighting function of the mobile-mean filter (MMF).

In principle, this is the easiest filter one could think of: we're only averaging for a finite amount of time the value of our input signal, with a weight constant for all the duration of the averaging operation. Instead, for the RC integrator, the weights decrease exponentially.

It's worth pointing out that this is, once again, a constant parameter filter, since it has the same weighting function for every t_m chosen.

The easiest way to confirm its low-pass filtering action is to see at its transfer function, i.e. at the Fourier transform of the δ-response. Since its δ-response (that is the weighting function flipped) is a rectangular function, we know its frequency domain representation is a cardinal sine, or sinc function, as represented in figure 8.14.

**Figure 8.14** Weighting function of the mobile-mean filter (MMF).

So, at low frequency we leave the signal unaltered, while at higher frequency it's always more and more attenuated, with some ripples.

It's worth noting that there are some *zeros* in this transfer function. This means that there are frequencies whose effect on the output will be zero. This feature might be exploited to get rid of some disturbances at certain well-known frequencies, for example.

The existence of the zeros of the MMF transfer function can be also explained in the time domain. In fact, one way to see the weight of each frequency component is to

check the area of the sine at that frequency that is integrated by the rectangle. It's clear that there are certain frequencies, multiples of the fundamental one, that are exactly integrated over a certain number of periods. So, their contribution will be nihil. This is better clarified in figure 8.15.
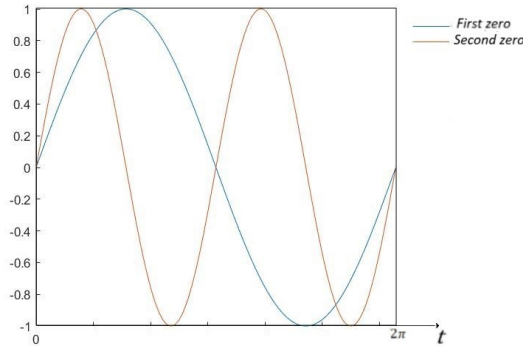
Figure 8.15 Explanation of the existence of zeros in the transfer function of the MMF.

This filter is practically implemented following the idea depicted in figure 8.16. The biggest problem in the practical realization of these kind of filters comes from the analog delay line, implemented with very long transmission lines with limited performance. Even realizing delay lines of more than few tens of nanoseconds is impractical. This implies that the first zero, that gives us an idea of bandwidth of the filter, can be up to $\sim \frac{1}{20 \, ns} = 50 \, MHz$. If the bandwidth of the signal is narrower than this value, we should look for other solutions.
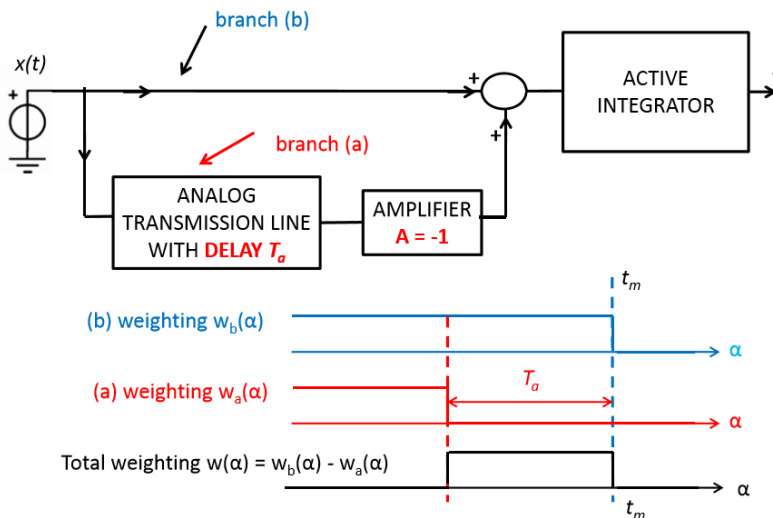
Figure 8.16 Conceptual scheme for the practical implementation of an MMF.

Let's try to compare this filter with the passive RC-integrator. For the signal we can easily notice that they both have the same (unitary) DC gain.
For the noise, let's suppose we have a wide-band noise, as we've done with the RC-integrator. To have the same performances in terms of RMS noise at the output, we want that the result of

$$\overline{y^2} \cong S_{wb} K_{ww}(0)$$

to be the same for both the filters. So, we need the same value of the autocorrelation in zero for both of them. The autocorrelation of a rectangular signal is simply a triangular signal, with a peak equal to $\frac{1}{T_a}$.

For the passive RC integrator instead, we found $K_{ww}(0) = \frac{1}{2T_f}$.

So, to have the same RMS noise we simply need to choose

$$T_a = 2T_f$$

This result confirms our previous intuitive approximation of the RC-integrator as an average over a period equal to 2 time constants.
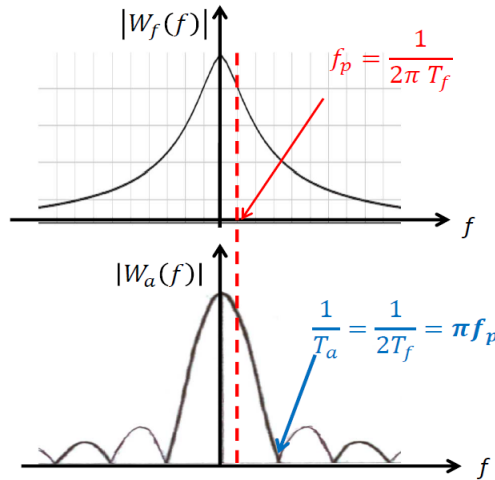
Figure 8.17 Comparison between the transfer function of a RC integrator (top) and the one of an MMF (bottom).

8.5 Final considerations

8.5.1 Ideal behavior

If our signal has frequency components squeezed at low frequencies, we'd like a "rectangular" filter in the frequency domain: this filter simply does not exist. In fact,

its equivalent in time would be a cardinal sin, that has two problems: it has an infinite duration of the response to the stimulus, and moreover is anti-causal: that would imply that the system is reacting to a delta input before its application, that is clearly not physically meaningful.
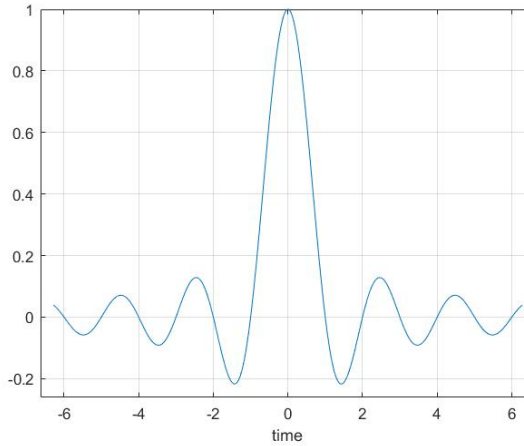
Figure 8.18 The cardinal sine has an infinite duration, and also it's not zero for t<0. It clearly cannot model the δ-response of any physical system.

However, this delta response can be implemented in the digital domain, where the exploitation of memories makes the anti-causality of a filtering system feasible.
It's worth noticing that it is possible to realize a rectangular in the time domain. In fact, the MMF does that, but its practical implementation is often limited, as discussed above.

8.5.2 Cascading elementary cells
We can easily obtain high-order LPF by cascading the simple configurations discussed above.
In general, for LPF with real poles, it's better to compute the equivalent noise bandwidth in the time domain, since it's easier to unravel the integrals to get there.
For instance, for a second order RC, we'd have a δ-response

$$h(t) = \frac{t}{T_f^{\,2}} e^{-\frac{t}{T_f}} 1(t)$$

and an RMS value at the output

$$\overline{\sigma_{out}^{\,2}} = S_{Wb} k_{hh}(0) = S_{Wb} \int_0^{+\infty} \left(\frac{t}{T_f^{\,2}} e^{-\frac{t}{T_f}} \right)^2 dt$$

To solve the integral, we should integrate by parts finding the value of the noise equivalent bandwidth:

$$\int_0^{+\infty} \left(\frac{t}{T_f^2} e^{-\frac{t}{T_f}} \right)^2 dt = \left[-\frac{t^2}{T_f^4} \frac{T_f}{2} e^{-\frac{2t}{T_f}} \right]_0^{+\infty} - \int_0^{+\infty} -\frac{2t}{T_f^4} \frac{T_f}{2} e^{-\frac{2t}{T_f}} dt$$

$$= \left[-\frac{t}{T_f^3} \frac{T_f}{2} e^{-\frac{2t}{T_f}} \right]_0^{+\infty} - \int_0^{+\infty} -\frac{1}{T_f^3} \frac{T_f}{2} e^{-\frac{2t}{T_f}} dt = \frac{1}{4T_f}$$

8.6 A practical example

Let's suppose to have a rectangular signal we want to low-pass filter. Suppose we can only use a simple RC integrator. Given that the duration of our signal is T_s and the amplitude is V_s, how should we set the pole of our filter to obtain the best result? So, how long should our exponential tail last? The signal is

$$y_{out} = \int_{-\infty}^{t_m} V_s \, rect \left(\frac{\tau - \frac{T_s}{2}}{T_s} \right) \frac{1}{T_f} e^{-\frac{t_m - \tau}{T_f}} d\tau$$

$$= \frac{V_s}{T_f} e^{-\frac{t_m}{T_f}} \int_0^{t_m} e^{\frac{\tau}{T_f}} d\tau = V_s e^{-\frac{t_m}{T_f}} \left[e^{\frac{\tau}{T_f}} \right]_0^{t_m} = V_s \left(1 - e^{-\frac{t_m}{T_f}} \right)$$

N.B. That's the way we write a rectangular signal with duration T_s that starts at $t=0$.

For the noise instead, as already pointed out, we have

$$\overline{\sigma_{out}^2} = S_{Wb} k_{hh}(0) = S_{Wb} \frac{1}{2T_f}$$

So, the S/N will be

$$\frac{S}{N} = \frac{V_s \left(1 - e^{-\frac{t}{T_f}} \right)}{\sqrt{S_{Wb} \frac{1}{2T_f}}} = \frac{V_s \sqrt{2T_f} \left(1 - e^{-\frac{t}{T_f}} \right)}{\sqrt{S_{Wb}}}$$

Now, we want to find the best possible T_f, i.e. the T_f that would maximize that S/N. Let's first verify that there is a maximum.

- For $T_f \to 0$, the S/N $\to 0$.
- For $T_f \to \infty$, it is that $\frac{S}{N} \sim \sqrt{T_f} \frac{t}{T_f} = \frac{t}{\sqrt{T_f}} \to 0$.
- There is at least one nonzero value in the middle.

So, since that function is continuous, we expect to have a maximum somewhere. To find it, we should set

$$\frac{d}{dT_f}\left(\frac{S}{N}\right) = 0$$

So, it's, with $t_m = T_s$,

$$\frac{d}{dT_f}\left(\sqrt{T_f}\left(1 - e^{-\frac{T_s}{T_f}}\right)\right) = \frac{1}{2\sqrt{T_f}}\left(1 - e^{-\frac{T_s}{T_f}}\right) - \sqrt{T_f}\, e^{-\frac{T_s}{T_f}}\,\frac{T_s}{(T_f)^2} = 0$$

and so

$$\left(1 - e^{-\frac{T_s}{T_f}}\right)T_f - 2\,T_s\, e^{-\frac{T_s}{T_f}} = 0$$

$$e^{-\frac{T_s}{T_f}}\left[T_f\left(e^{\frac{T_s}{T_f}} - 1\right) - 2\,T_s\right] = 0$$

$$e^{\frac{T_s}{T_f}} = \frac{2\,T_s + T_f}{T_f}$$

This can be graphically solved finding:

$$T_f = 0{,}8\,T_s$$

CHAPTER 9

Switched-parameter Low Pass Filter

In this chapter we will introduce the switched parameter filter. When some basic properties of the signal waveform of interest are known such as its arrival time, duration or shape, the filter can be switched on only when the signal is present. This can be achieved by employing a switch activated by a synchronization signal. It will be shown that this different approach can be used in different applications.

9.1 Introduction

Under the hypothesis of linearity, the output y of any filter can be computed as a weighted sum of input values x taken at various time α. The weights of the filter represent the effect of the input at time α on the output at time t_m. In continuous-time signal filtering this operation consists in an integral:

$$y(t_m) = \int_{-\infty}^{t_m} w(\alpha)x(\alpha)d\alpha$$

Let's consider the basic circuit shown in figure 9.1:
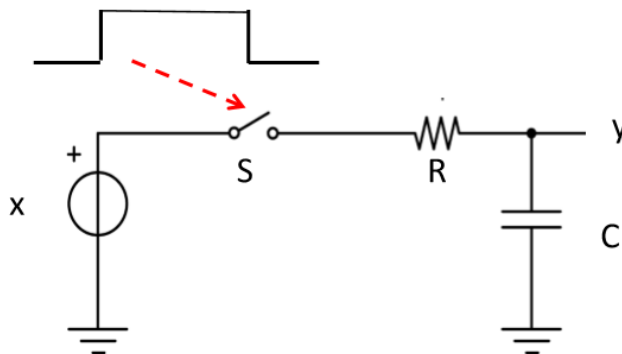
Figure 9.1 Switched-Parameter LPF

- If the switch S is closed the circuit behaves like a constant-parameters low pass filter. For a constant-parameters filter we saw that the weighting function is the delta response of the circuit shifted by t_m and reversed.

$$y(t_m) = \int_{-\infty}^{t_m} w(\alpha)x(\alpha)d\alpha = \int_{-\infty}^{t_m} h(t_m - \alpha)x(\alpha)d\alpha$$

- If the switch S is open the circuit is in hold and no current can flow, so the input can no longer affect the output node.

Usually, the opening of the switch is related to a *sync signal* that tells us when the signal we want to analyze is present.

Depending on the choice of the time constant, the filter can have different behaviors. In figure 9.2 the behavior of the filter for different values of the time constant is shown.
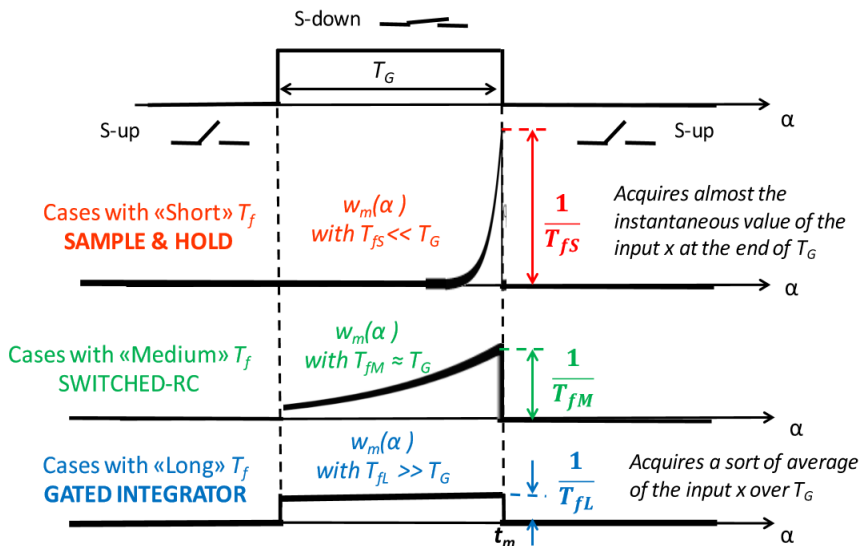

Figure 9.2 Different behaviors depending on RC

When the time constant of the low pass filter T_f is much lower than T_G (where T_G is the time for which the switch is closed), the circuit behaves like a **Sample and Hold**. In this case the circuit tracks almost instantaneously the input voltage when the switch is closed.

If the time constant T_f is of the same order of T_G, the circuit behaves like a **switched RC low pass filter**, performing an exponential average of the input signal.

The last case is when the time constant T_f is much higher than T_G. In this situation a true integration of the input signal is made. This type of filter is called **Gated Integrator**.
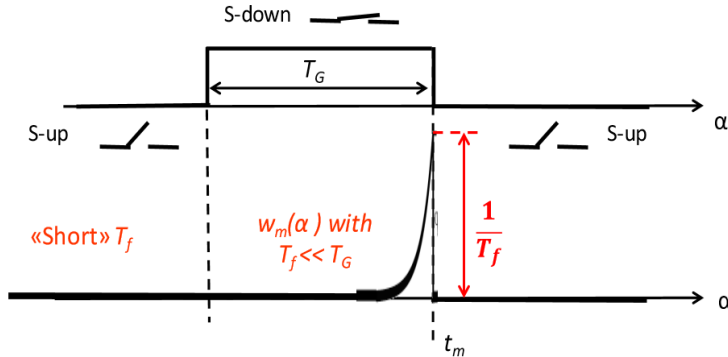
9.2 Sample & Hold

Figure 9.3 Weighting function of S&H

Sample & Hold circuits are a fundamental block in data acquisition systems, mostly because they are widely used in analog to digital conversion. In this case the goal is very simple: they have to sample the analog input signal and hold it at least for the time needed by the ADC to convert it into a digital signal.

The weighting function is identical to a constant parameter RC filter, thus presenting a unity DC gain

$$W_m(0) = \int_0^\infty w_m(\alpha)d\alpha = 1$$

In real world applications, the width of the Sample & Hold weighting function cannot assume any value, but it is limited by the finite value of the time constant T_f, which depends on real components present inside the system, mainly the resistance of fast switching devices and the capacitance used to store the information. One of the main challenges in the design of a Sample and Hold is the minimization of the acquisition time, that is linked with T_f. Today the exploitation of discrete components allows the design of systems featuring a T_f down to few nanoseconds, while integrated-circuits technologies can provide a T_f of only some picoseconds, thanks to lower parasitic capacitances.

The drawback of a very narrow weighting function in the time domain is a wide response in the frequency domain, and therefore a limited effect on noise filtering.

If a wide band noise is present, featuring a bilateral spectral density S_b=2kTR due to the resistor, the Sample & Hold will play the role of an RC integrator. Under the

hypothesis of white noise, the noise autocorrelation function is shaped as a delta pulse centered in 0. The noise power present at the output can be computed as follows:

$$y_R{}^2 = \int_{-\infty}^{\infty} R_{nn}(\tau)k_{ww}(\tau)d\tau = S_b k_{ww}(0)$$

In case of a S&H, the weighting function of the filter is:

$$w(\alpha) = \frac{1}{T_f} e^{-\frac{(t_m - \alpha)}{T_f}} 1(t_m - \alpha)$$

From which the autocorrelation is derived

$$k_{ww}(\tau) = \frac{1}{2T_f} e^{-\frac{|\tau|}{T_f}}$$

$$y_R{}^2 = 2kTR \cdot \frac{1}{2T_f} = 2kTR \cdot \frac{1}{2RC}$$

$$y_R{}^2 = \frac{kT}{C}$$

As soon as we reduce the value of T_f we increase the value of the autocorrelation in zero, and so also the output total noise.

Computations show also an important result: noise power present at the output is independent on the value of R and gets higher as the capacitance C gets smaller.

9.3 Gated integrator

Let's analyze now the last case of the figure 9.2, that is a switched-parameters RC network featuring a time constant of the RC network T_f much higher than T_G.
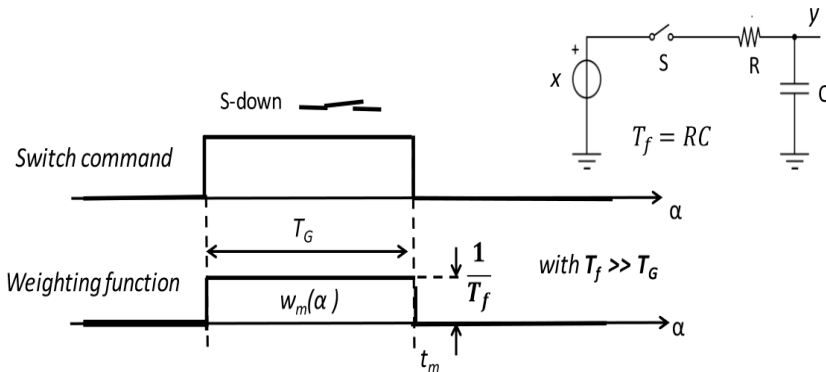

Figure 9.5 Weighting function of GI

In this case, the circuit performs a true integration of the input signal. This filter is particularly useful when a relatively slow signal (so that it is constant over the integration window) must be recovered in presence of strong wide band noise. The name of this filter is **Gated Integrator** (GI).

The DC gain of a Gated Integrator is much lower than one

$$G = W_m(0) = \int_0^\infty w_m(\alpha)d\alpha = \frac{T_G}{T_f} \ll 1$$

The filter weighting function can be approximated with

$$w_m(\alpha) = \frac{1}{T_f} \cdot rect_{T_g}(\alpha - t_m + \frac{T_g}{2})$$

Where the rect signal is defined as

$$rect_{T_g}(t) = \begin{cases} 1 \ for -\frac{T_g}{2} \le t \le \frac{T_g}{2} \\ 0 \ for \ |t| > \frac{T_g}{2} \end{cases}$$

Computing the Fourier transform we have

$$F\left(rect_{T_g}(t)\right) = \int_{-\frac{T_g}{2}}^{\frac{T_g}{2}} e^{-j2\pi ft}dt = \frac{1}{-j2\pi f}[e^{-j2\pi ft}]_{-\frac{T_g}{2}}^{\frac{T_g}{2}} = \frac{e^{j\pi fT_g} - e^{-j\pi fT_g}}{2j} \cdot \frac{T_g}{\pi fT_g}$$

$$= T_g \cdot \frac{\sin \pi fT_g}{\pi fT_g}$$

The ratio between a sine wave and its argument is the well-known cardinal sine

$$sinc_{T_g}(f) = \frac{\sin \pi fT_g}{\pi fT_g}$$

The modulus of a Gated Integrator transfer function then is

$$|W_m(f)| = \frac{T_g}{T_f} \cdot sinc_{T_g}(f)$$

Exploiting the symmetry of the rect shape, the autocorrelation can be computed convolving two identical rect signals, thus obtaining a triangular shape

$$k_{ffw} = w_m(\alpha) * w_m{}^*(-\alpha) = \frac{1}{T_f} \cdot rect_{T_g}(t) * \frac{1}{T_f} \cdot rect_{T_g}(t) = T_g \cdot \left(\frac{1}{T_f}\right)^2 \cdot tri_{T_g}(t)$$

The tri function is defined as

$$tri_{T_g}(t) = \begin{cases} (\dfrac{t}{T_g} + 1) & for - T_g < t \le 0 \\[3mm] (-\dfrac{t}{T_g} + 1) & for \ 0 < t < T_g \end{cases}$$

The same approach we used to calculate the autocorrelation in time can be used to extract its Fourier transform. Since in time the autocorrelation is the convolution of two rect, in the frequency domain we will have the product of their Fourier transform, that is a square sinc as shown in the next formula:

$$F\big(w_m(\alpha) * w_m{}^*(-\alpha)\big) = |W_m(f)|^2 = \left(\frac{T_g}{T_f}\right)^2 \cdot sinc_{T_g}{}^2(f)$$
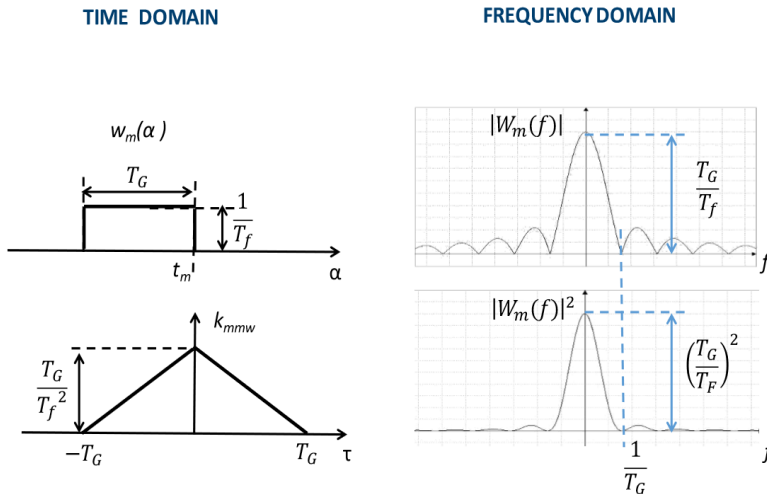

Figure 9.6 Time Vs. frequency domain

9.3 Gated Integrator S/N improvement

Differently from the Sample&Hold, the Gated Integrator can have a quite broad weighting function and so a narrow response in the frequency domain; therefore, a GI can be an effective option to reduce white noise. Let's consider a constant input signal over the integration time T_G and a broadband input noise power limited by a single

pole (for example the one of the preamplifier) with a time constant much slower of T_G. Under these hypotheses, we can consider the noise as white.

$$x_n{}^2 = S_b 2f_n = \frac{S_b}{2T_n}$$

The signal at the output of the filter is then:

$$y_S = x_S \cdot \frac{T_G}{T_f} = x_S G \qquad \text{i.e. with gain} \qquad G = \frac{T_G}{T_f} = 1$$

The output noise power, under the condition of white input noise, is

$$y_n{}^2 = \int_{-\infty}^{\infty} R_{nn}(\tau) k_{ffw}(\tau) d\tau = S_b k_{ffw}(0)$$

$$y_n{}^2 = S_b \frac{T_G}{T_f{}^2} = \frac{S_b}{T_G} (\frac{T_G}{T_f})^2 = \frac{S_b}{T_G} G^2 = \frac{S_b}{2T_n} \frac{2T_n}{T_G} G^2 = x_n{}^2 \frac{2T_n}{T_g} G^2$$

It is possible now to calculate the improvement due to the Gated Integrator on the signal to noise ratio:

$$(\frac{S}{N})_y = \frac{y_S}{\sqrt{y_n{}^2}} = \frac{x_S}{\sqrt{x_n{}^2}} \sqrt{\frac{T_G}{2T_n}} = (\frac{S}{N})_x \sqrt{\frac{T_G}{2T_n}}$$

The output signal increase as T_G and the noise as $\sqrt{T_G}$, therefore the Signal to Noise ratio increases as the square root of the integration time $\sqrt{T_G}$.
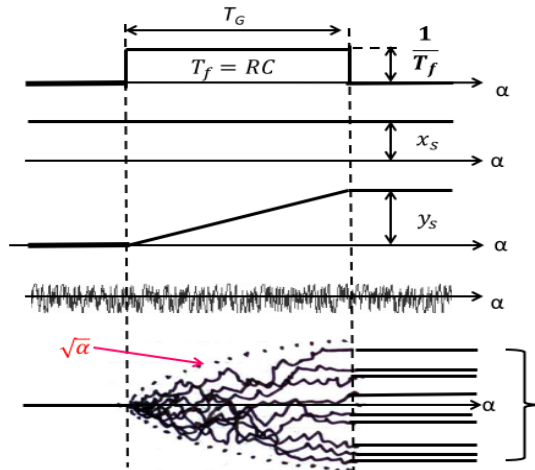

Figure 9.7 Effect of a GI on a constant signal and on white noise

It is also possible to see that the larger is the difference between the filter time constant and the autocorrelation width, the larger is the gain. This is intuitive from the frequency point of view: we are increasing the part of the noise spectrum we are removing with the filter.

9.3 Gated integrator optimization

In this paragraph we will discuss how to properly size a gated integrator to reach an optimal Signal-to-Noise ratio, given two of the most common input signal shapes.

9.3.1 Rect pulse signal
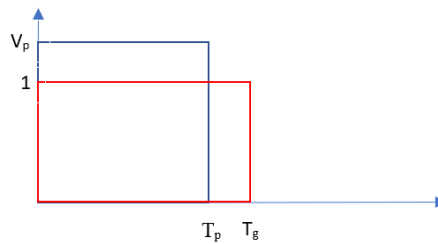
Figure 9.8 rect signal (blue) with duration Tp and GI (red) with duration Tg

Let's consider a rect signal with amplitude V_p and duration T_p. In order to evaluate the signal at the output of the GI filter, the following integral must be computed. It's a straightforward conclusion that its maximum value is for $T_g=T_p$.

$$y_s = \int_{-\infty}^{t_m} V_p dt = \int_0^{T_g} V_p dt = V_p T_g$$
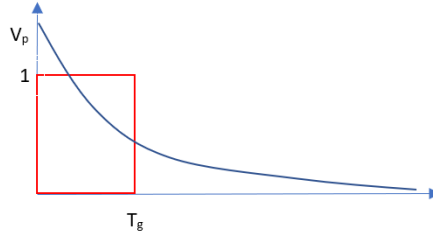
For what concerns the noise contribution:

$$\overline{n_y^2} = S_b \cdot k_{ww}(0) = S_b \cdot T_g$$

The signal to noise ratio in this case is a monotone function and it is maximized when all the signal energy is acquired (Tg=Tp).

$$SNR = \frac{V_p \cdot T_g}{\sqrt{T_g \cdot \frac{S_{nu}}{2}}} = \frac{V_p}{\sqrt{\frac{S_{nu}}{2}}} \cdot \sqrt{T_g}$$

9.4.1 Exponential decay pulse signal

Figure 9.9 Exponential signal (blue) with time constant Tp and GI (red) with duration Tg

In this case the optimal choice of T_g is not intuitive. Let's compute the SNR with a similar approach

$$s_y = \int_0^{T_g} A \cdot V_p e^{-\frac{t}{T_p}} dt = A \cdot V_p \cdot T_p \cdot (1 - e^{-\frac{T_g}{T_p}})$$

$$SNR = \frac{V_p \cdot T_p}{\sqrt{\frac{S_{nu}}{2} \cdot T_p}} \cdot \frac{(1 - e^{-\frac{T_g}{T_p}})}{\sqrt{\frac{T_g}{T_p}}}$$

This functional dependence is not monotone and has a maximum. Imposing

$$x = \frac{T_g}{T_p}$$

Computing the derivative and equating it to zero

$$max\left(\frac{(1 - e^{-x})}{\sqrt{x}}\right) \quad for \ x = 1{,}25$$

The Signal-to-Noise ratio is thus maximized for $T_g = 1.25 T_p$

It is worth noting that, in both these cases, we started to integrate the signal as soon as the signal is present avoiding taking in consideration if this is the best starting point for our weighting function. In both the cases this assumption is correct since the maximum of the signal is at its beginning, we will see in the future this is not always the case.

9.4 Gated integrator vs. Mobile Mean filter comparison

In the previous chapter, we have studied the Mobile Mean Filter. At first sight, we could state that a MMF and a GI feature the same weighting function, which translates into an identical Signal-to-Noise ratio improvement.

However, it is extremely important to point out some key differences:

- the mobile mean filter is a constant parameter filter. This means that its output is the running average of the input signal for every time instant. The Gated Integrator, instead, is a not constant parameter filter: its output it is not a waveform in time but just a value representing the signal integral over the filter window.

- the shape of the weighting function of the gated integrator does not depend on a delay line, so it is possible to modify it by simply changing the activation time of the switch. This gives the designer the freedom to create a GI filter with any time width (taking in consideration the technological issues connected to the turning on and off of the switch).

9.5 GI compared to other LPF

Let's try now to make a fair comparison between a gated integrator and a constant parameter RC filter. First of all, we have to define a way to compare the results: one typical way is to fix a unity DC gain for both filters and then compare the noise outputs of them.
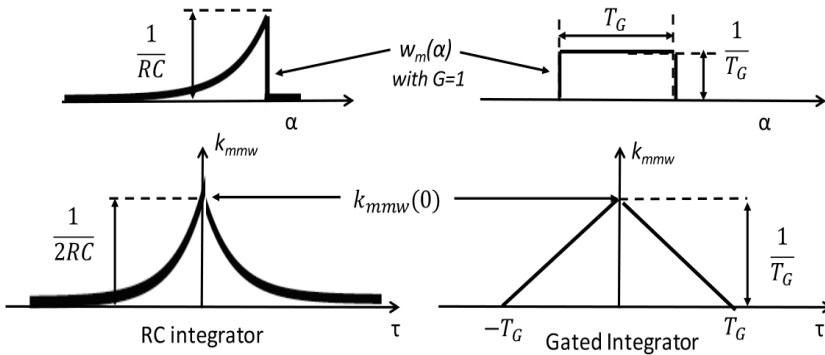

Figure 9.10 RC Vs. GI integrator (time domain)

In case of unity DC gain the two noise output powers are:
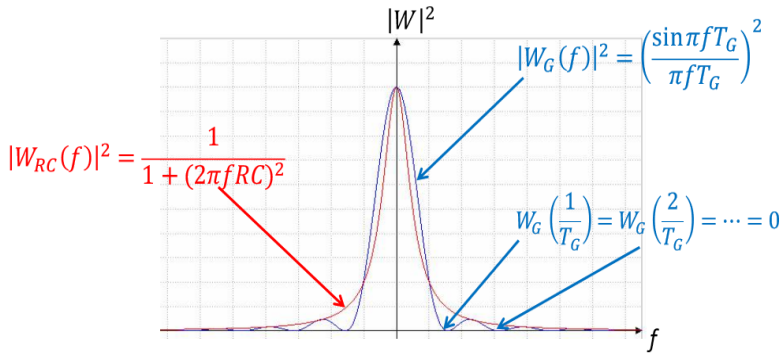
$$n_{GI}^2 = S_b \cdot \frac{1}{T_g}$$

$$n_{RC}^2 = S_b \cdot \frac{1}{2RC}$$

Therefore, they produce the same output noise power, if the following condition is verified:

$$T_G = 2RC$$

For this reason, we are used to say the RC low pass filter in the time domain is equivalent to an average on two-time constants, as discussed with the MMF.

Figure with labels:

$|W|^2$

$|W_{RC}(f)|^2 = \dfrac{1}{1 + (2\pi f RC)^2}$

$|W_G(f)|^2 = \left(\dfrac{\sin \pi f T_G}{\pi f T_G}\right)^2$

$W_G\left(\dfrac{1}{T_G}\right) = W_G\left(\dfrac{2}{T_G}\right) = \cdots = 0$

f

Figure 9.11 RC Vs. GI integrator (frequency)

As shown in figure 9.11, the transfer function of the gated integrator goes to zero at frequencies $f_k = \dfrac{k}{T_G}$ .This property can be used to cancel out specific disturbances at frequencies $f_D = f_k$ (for example radio frequency disturbs or 50Hz disturbs due to the standard power supply).

CHAPTER 10

Filtering: Switched-Parameter Averaging Filters

The scope of this chapter is to continue and conclude the discussion about low pass filters. Two examples of digital filters will be presented, focusing on their operating principles, signal to noise ratio performance and making a comparison with their corresponding analogue version.
Then, two new analogue filters will be introduced: they are particularly useful for repetitive signals, showing that a proper sampling can improve the signal to noise ratio.

10.1 Introduction

Nowadays, the fast development of more and more advanced digital circuits has been leading to the digital implementation of filters in many cases. Digital filters, indeed, have many advantages over their analog counterparts: they do not have any internal noise sources, they do not suffer from zero-level drifting, they do not suffer from variable saturation (in analog filters the signal cannot normally be larger than the power supply), they do not have any intrinsic nonlinearities and they do not suffer from long holding-times leakage (leakage usually becomes a problem in analog filters for holding times longer than 1s).

On the other hand, analog filters can often perform real-time processing of waveforms on faster time scales than their digital counterpart, and thus they are still the solution of choice in several applications.

In the following paragraphs, we will compare analog and digital filters, showing that analog filters always outperform their digital counterparts.

10.2 Discrete Time Integrator (DTI)

If we are dealing with a sampled signal, we can choose to use the Discrete-Time Integrator (DTI), which can be seen as the discrete equivalent to the Gated Integrator (GI).

As we have seen in the previous chapter, the GI basically computes an average of the input signal over a certain time window. The analog filtering action can be mimicked using a digital approach, that is taking N samples with a constant weight, sum them and divide the result by a given "value".
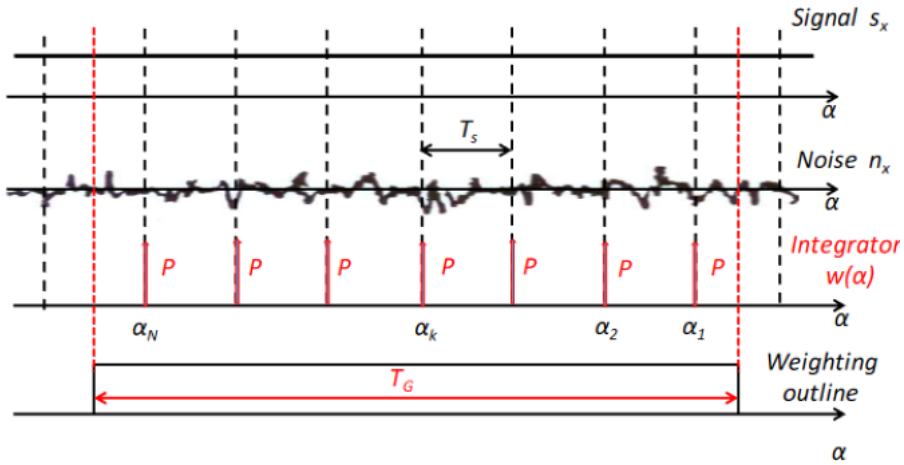
Figure 10.1 constant signal, white noise and weighting function of a Discrete-Time Integrator

For simplicity, we start to analyse a constant DC signal s_x and a wide band stationary noise n_x with autocorrelation width $2T_n$, like the one in figure 10.1. Each sample is multiplied by the weight P and summed, up to a total of N samples. Choosing a T_G interval and a sampling frequency $f_S = 1/T_S$ we find a number of samples equal to $N = T_G/T_S$.

In these conditions, the DC Gain of the DTI is equal to $N \cdot P$.

Thus, unitary gain can be achieved by choosing $P = \frac{1}{N}$. As we will see in a few lines, the choice of P will not have any impact on the final SNR.

We can similarly calculate the output due to noise at the input: the output will be equal to the sum of all the "noisy samples"

$$n_y = \sum_{k=1}^{N} P \cdot n_{x,k}$$

Thus, the mean square value will be

$$\overline{n_y}^2 = P^2 \overline{\left(n_{x,1}{}^2 + n_{x,2}{}^2 + \cdots + n_{x,1}n_{x,2} + \cdots\right)}$$
$$= P^2 \left(\overline{n_{x,1}}^2 + \overline{n_{x,2}}^2 + \cdots + \overline{n_{x,1}n_{x,2}} + \cdots\right)$$

If we consider the noise as white (or more realistically, we hypothesize that the noise autocorrelation time T_N is much smaller that the sampling time T_S), the noise samples will be uncorrelated with each other, and the cross products will be null $\left(\overline{n_{x,1}n_{x,2}} = \overline{n_{x,2}n_{x,3}} = 0\right)$. If we also consider the noise as stationary (such that that $\overline{n_{x,1}}^2 = \overline{n_{x,2}}^2 = \overline{n_{x,k}}^2 = \overline{n_x}^2$) the mean square noise becomes

$$\overline{n_y}^2 = P^2\left(\overline{n_{x,1}}^2 + \overline{n_{x,2}}^2 + \cdots + \overline{n_{x,1}n_{x,2}} + \cdots\right) = N \cdot P^2 \cdot \overline{n_x}^2$$

As with the time-continuous GI, the output noise increases linearly with the integrating time T$_G$ (remember that $N = T_G/T_S$).
We can now calculate the SNR:

$$\left(\frac{S}{N}\right)_y = \frac{N \cdot P \cdot S_x}{\sqrt{N \cdot P^2 \cdot \overline{n_x}^2}} = \sqrt{N} \cdot \left(\frac{S}{N}\right)_x$$

As expected, we see that the choice of P does not affect the final SNR. The reason becomes obvious if we consider how both the signal and the noise are similarly amplified by a factor P.

10.2 Discrete-Time Exponential Averager

In the previous paragraph, we obtained the Discrete-Time integrator by mimicking the weighting function of the Gated-integrator. Similarly, we can define the weighting function of the Discrete-Time Averager by mimicking the weighting function of a time-continuous Exponential filter.
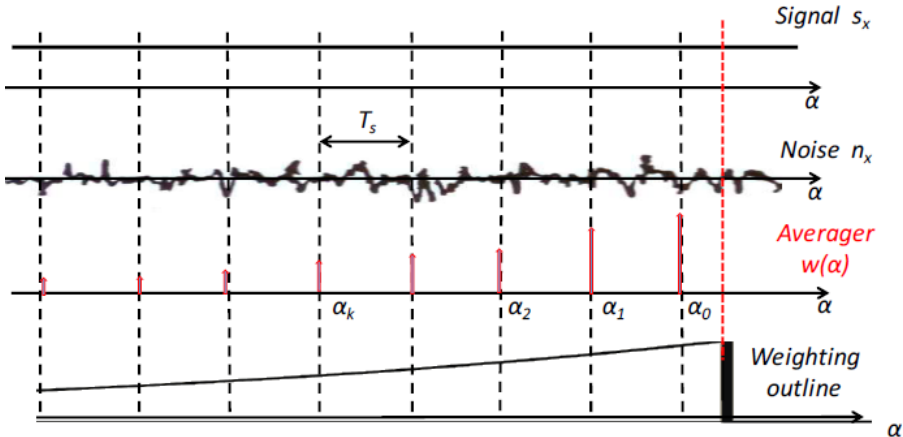
Figure 10.2 constant signal, white noise and weighting function of a Discrete-Time Averager

We expect the weighting function to be $w_k = P \cdot r^k$, with $r < 1$

If we now apply a constant input signal S$_x$, the signal at the output will be

$$S_y = \sum_{k=0}^{\infty} P \cdot S_x \cdot r^k = P \cdot S_x \cdot \frac{1}{1-r}$$

where the second step is obtained taking into account the geometric series.
A geometric series:

$$\sum_{k=0}^{\infty} x^k$$

for a value of $|x| < 1$ it has a convergent behavior, so the series sum exists and its value is

$$\sum_{k=0}^{\infty} x^k = \lim_{n\to\infty} \sum_{k=0}^{n} x^k = \lim_{n\to\infty} \frac{1 - x^{n+1}}{1 - x} = \frac{1}{1 - x}$$

The DC Gain will thus be $\frac{P}{1-r}$.

Regarding noise, if we consider a white input noise (such that all cross products are null), the output noise will be

$$n_y = \sum_{k=0}^{\infty} P \cdot n_{x,k} \cdot r^k$$

The mean square output noise will be

$$\overline{n_y}^2 = P^2\left(\overline{n_{x,1}}^2 + \overline{n_{x,2}}^2 \cdot r^2 + \cdots + \overline{n_{x,k}}^2 \cdot r^{2k} + \cdots\right)$$

If we also consider the input noise as **stationary**

$$\overline{n_y}^2 = P^2 \cdot \overline{n_x}^2 (1 + r^2 + \cdots + r^{2k} + \cdots) = (\text{Geom. series}) = P^2 \cdot \overline{n_x}^2 \cdot \frac{1}{1-r^2}$$

We can thus calculate the SNR:

$$\left(\frac{S}{N}\right)_y = \frac{P \cdot S_x \cdot \frac{1}{1-r}}{\sqrt{P^2 \cdot \overline{n_x}^2 \cdot \frac{1}{1-r^2}}} = \frac{\sqrt{1-r^2}}{1-r} \cdot \left(\frac{S}{N}\right)_x = \left(\frac{S}{N}\right)_x \sqrt{\frac{1+r}{1-r}}$$

If we consider the attenuation factor close to 1 such that $(1 - r) \ll 1$ (which means choosing the sampling period T_S much shorter than the decay time), then $(1 + r) \simeq 2$. Thus

$$\left(\frac{S}{N}\right)_y \simeq \left(\frac{S}{N}\right)_x \sqrt{\frac{2}{1-r}}$$

Question: *What is the SNR at the output of the Discrete Exponential-Averager if an exponential signal is applied at the input?*

Let's take a signal $S_x = V_P e^{-t/T_P}$

For this example, we choose the filter's time constant equal to the signal's time constant, such that $w_k = P \cdot e^{-k \cdot T_S/T_P}$. In order to simplify the calculations, we set $\alpha = e^{-T_S/T_P}$.

We can now proceed by calculating the S/N:

$$S_y = \sum_{k=0}^{\infty} S_x \cdot P\, e^{-k \cdot T_S/T_P} = \sum_{k=0}^{\infty} V_P\, e^{-k \cdot T_S/T_P} \cdot P\, e^{-k \cdot T_S/T_P}$$

$$= \sum_{k=0}^{\infty} V_P\, P\, e^{-2k \cdot T_S/T_P} = V_P\, P \cdot \sum_{k=0}^{\infty} \alpha^{2k}$$

if we now apply the Geometric series

$$S_y = \frac{V_P \cdot P}{1 - \alpha^2}$$

Now for the noise. In case of stationary white noise at the input:

$$\overline{n_y}^2 = \sum_{k=0}^{\infty} \left(\overline{n_x} \cdot P\, e^{-k \cdot T_S/T_P} \right)^2 = \overline{n_x}^2 \cdot P^2 \sum_{k=0}^{\infty} \alpha^{2k} = geom.\, series = \frac{\overline{n_x}^2 \cdot P^2}{1 - \alpha^2}$$

Thus

$$\left(\frac{S}{N}\right)_y = \frac{\dfrac{V_P \cdot P}{1 - \alpha^2}}{\sqrt{\dfrac{\overline{n_x}^2 \cdot P^2}{1 - \alpha^2}}} = \left(\frac{S}{N}\right)_x \cdot \frac{1}{\sqrt{1 - \alpha^2}}$$

We can further manipulate this expression to obtain a simpler result. The key is to consider that in a well-engineered system, $T_S \ll T_P$ (intuitively, the sampling time should be much smaller than the time constant to obtain many samples).
Thus

$$(1 - \alpha) = \left(1 - e^{-T_S/T_P}\right) \simeq T_S/T_P \text{ , and } (1 + \alpha) = \left(1 + e^{-T_S/T_P}\right) \simeq 2.$$

Therefore if $T_S \ll T_P$

$$\left(\frac{S}{N}\right)_y = \left(\frac{S}{N}\right)_x \cdot \frac{1}{\sqrt{1 - \alpha^2}} = \left(\frac{S}{N}\right)_x \cdot \frac{1}{\sqrt{(1 + \alpha)(1 - \alpha)}} \simeq \left(\frac{S}{N}\right)_x \cdot \sqrt{\frac{T_P}{2T_S}}$$

Until now we have analyzed digital filters under the assumption of total lack of correlation between noise samples; in this scenario, it seems that by increasing number of samples we can outperform analogue filters. What does really happen if we increase the number of samples within a fixed time window? To answer this question, let's go back to the DTI and compare it with its analogue version.

10.3 Discrete Time Integrator versus Gated Integrator

For our analysis, we will now consider a time window T_G in which the Discrete Time Integrator weighting function (in red in figure 10.5) has N deltas, with $N = T_G/T_S$, with an amplitude equal to $1/N$, and the Gated Integrator (in blue) has a rectangular shape with constant value $1/T_G$.
To make a simple comparison we use two filters, both with a unity DC Gain

$$S_x = S_y$$

so, we have to compare only their output noise.
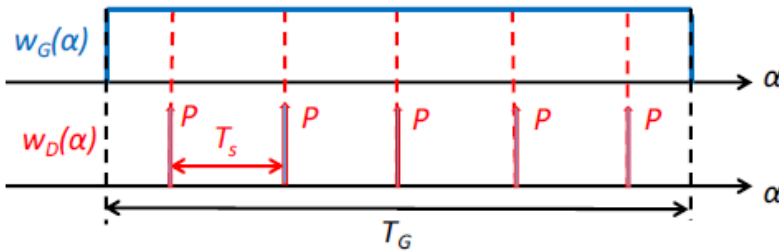

Figure 10.5 weighting function of a DTI

Under the assumption of DC signal s_x and wide-band noise S_B with correlation width $2T_n \ll T_S$, the improvement of the signal to noise ratio in these two situations is

Gated Integrator: $$\left(\frac{S}{N}\right)_y = \sqrt{\frac{T_G}{2T_N}} \cdot \left(\frac{S}{N}\right)_x$$

Discrete-Integrator: $$\left(\frac{S}{N}\right)_y = \sqrt{N} \cdot \left(\frac{S}{N}\right)_x$$

It seems that by increasing the number of samples N, which means reducing T_S, it is possible to obtain a better SNR improvement with a DTI rather than with a GI.
However, we have to pay attention at a key point: increasing the number of samples requires the reduction of T_S and this may violate our assumption of uncorrelated samples. When the assumption is no more valid, the improvement factor is no more equal to the square root of N.

At this point we have to answer the following question: what is the minimum acceptable sampling time for the DTI?

Of course until $T_S \gg 2T_n$ the noise samples are uncorrelated, but when T_S becomes equal to $2T_n$ we are playing around the limit and we have to remember that the real shape of the noise autocorrelation is an exponential, not a triangle, so there is some sort of correlation. This becomes evident if we take a look at the problem in the frequency domain. The spectral density of the DTI can be easily calculated as the sampled spectrum of the GI.

The final spectrum will have replicas of the GI's spectrum spaced $1/T_S$ from each other, as sketched in figure 10.6.
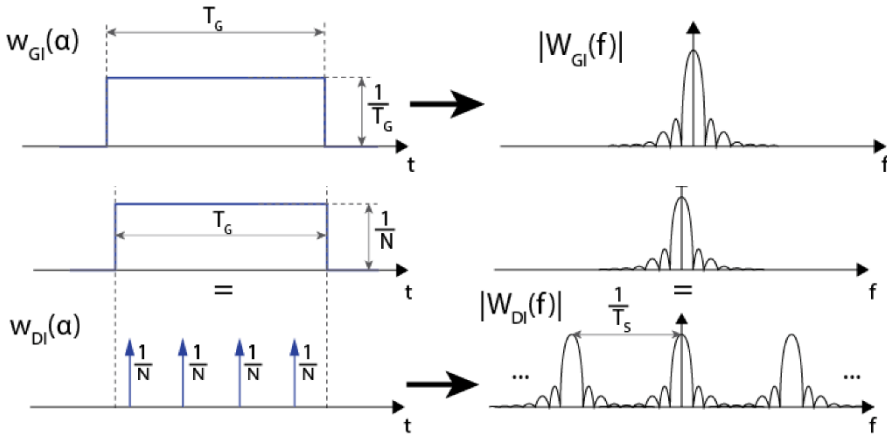
Figure 10.6 weighting function of a GI and a DTI in the frequency domain

Note that the output noise is proportional to the area of the filter spectrum. This means that the more replicas we have before the noise cutoff frequency, the higher the output noise will be.

Thus, with a smaller and smaller sampling time T_S, the filter replicas move away from each other resulting into the less and less noise at the output.
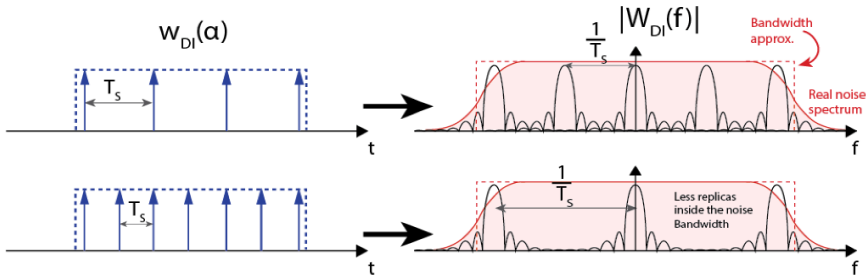

Figure 10.7 weighting function of a DTI in the frequency domain with different values of Ts

If we keep increasing the number of samples N, we will eventually reach a point where only a single replica counts to determine the output noise.
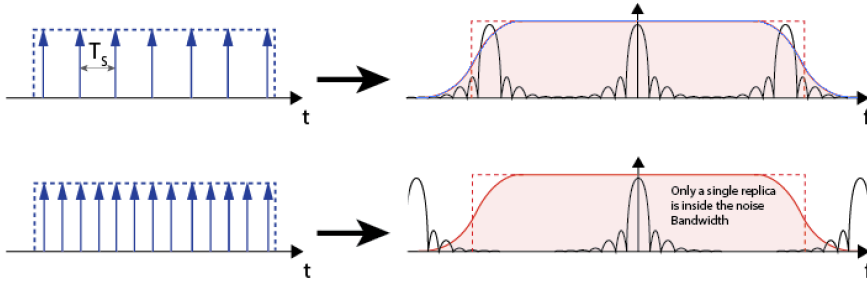This situation is the same as that of the Gated-Integrator

Figure 10.8 weighting function of a DTI in the frequency domain with different values of Ts

In conclusion we can always improve the digital signal to noise ratio, increasing the number of samples N, but the upper limit will be the performance obtained with an analogue filter due to the correlation between the noise samples.
Let's now analyse the difference between the Gate Integrator and the Discrete Time integrator in *time domain*.
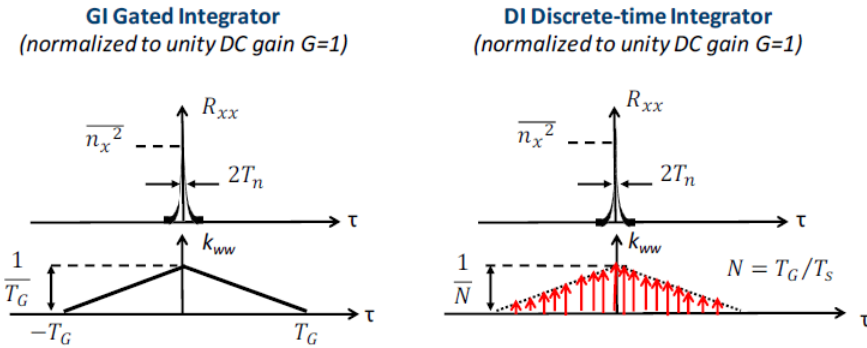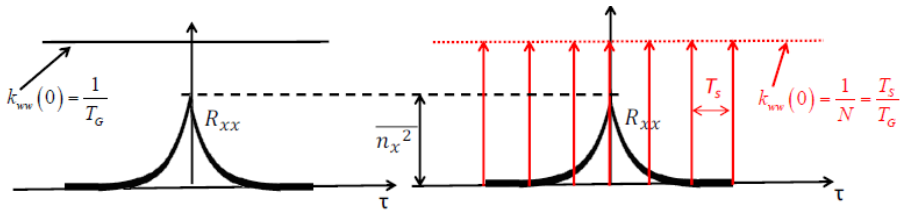

Figure 10.9 GI vs DTI in time domain

These figures show the autocorrelation of wide-band stationary noise and the weighting function of the two filters; from the previous lectures we know that under this condition it is important to know the value of these two autocorrelations in the origin to calculate the output noise. We zoom around $\tau = 0$

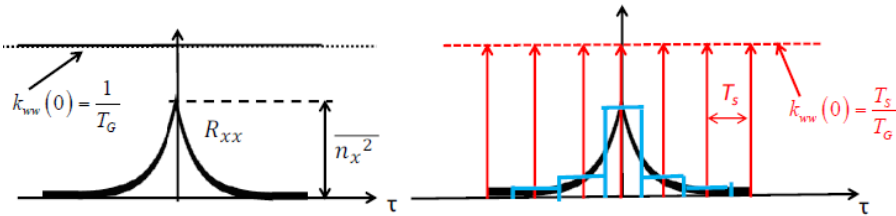**Figure 10.10** GI vs DTI in time domain

The output noise is the product of the input noise and the value of the autocorrelation of the filter's weighting function.

$$\overline{n_y}^2 = R_{YY}(0) = \int_{-\infty}^{\infty} R_{XX}(\tau) \cdot k_{ww}(\tau)d\tau$$

For the gated integrator, if we consider $T_G \gg T_N$, the output noise can be simply approximated as

$$\overline{n_y}^2 = \frac{1}{T_G} \cdot (\text{Area of } R_{XX}).$$

For the Discrete-Integrator instead, the autocorrelation function $k_{ww,DI}(\tau)$ is sampled. For this reason, the integral can be calculated by sampling R_{XX} with the deltas of $k_{ww,DI}$, and adding up all the samples

Figure 10.11 GI vs DTI in the time domain

$$R_{YY}(0) = \int_{-\infty}^{\infty} R_{XX}(\tau) \cdot k_{ww,DI}(\tau)d\tau$$

$$\simeq \int_{-\infty}^{\infty} R_{XX}(\tau) \cdot k_{ww,DI}(0) \sum_{k} \delta(\tau - k \cdot T_S)\, d\tau$$

$$= k_{ww,DI}(0) \sum_{k} R_{XX}(k \cdot T_S) = \frac{T_S}{T_G} \sum_{k} R_{XX}(k \cdot T_S)$$

In the last formula

$$R_{YY}(0) \simeq \frac{T_S}{T_G} \sum_k R_{XX}(k \cdot T_S)$$

we can consider

$$T_S \cdot \sum_k R_{XX}(k \cdot T_S)$$

as the area of the scaloid that approximates R_{XX}, as sketched in figure 10.12. Its area will be greater than that of R_{XX}, and thus $\overline{n_y}^2$ will be greater than that of the Gated-Integrator.
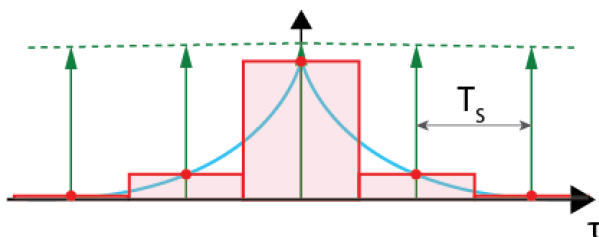

Figure 10.12 DTI: approximation of the integral in the time domain

If we decrease T_S, the scaloid will better approximate R_{XX}, to the point where its area will become, for $N \to \infty$, equal to the area of R_{XX} that is the result that we obtain with a GI.
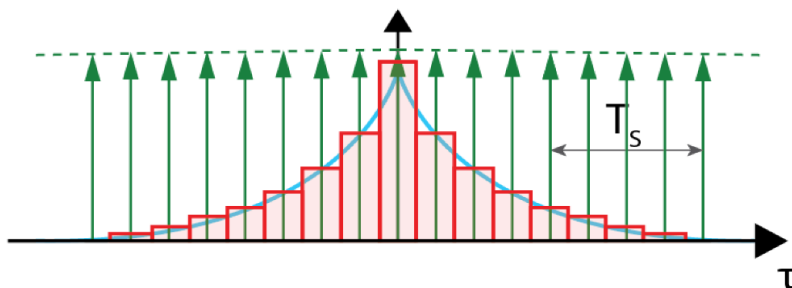

Figure 10.13 DTI: approximation of the integral in the time domain

10.4 Boxcar Integrator (BI)

The Boxcar Integrator is a filter having parameters that vary over time; it is particularly suitable to measure the amplitude of repetitive signals buried in white noise. We will show that a BI can be seen as some sort of expansion of the Gated

Integrator. First, let's step back to the GI to discuss one important limit of this filter. In order to improve the SNR obtained with a GI, we can't simply enlarge indefinitely the time window T_G of the filter itself. In fact, enlarging the integration window can lead to a SNR improvement only if we collect the signal all over the new time window. It is quite intuitive, for example, that the integration should not occur when there is only noise, otherwise the SNR would be impaired.

Nevertheless, we could consider the integration of multiple occurrences of the signal, if more than one sample is available. This approach is similar to increasing the number of samples with digital filters.
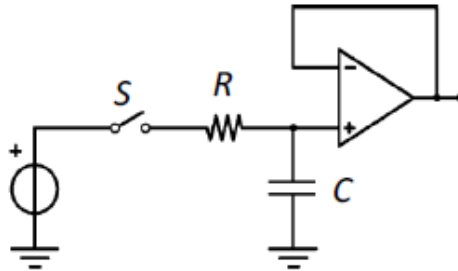

Figure 10.14 Boxcar Integrator scheme

The schematic of a Boxcar Integrator is shown in figure 10.14. It consists of a switch and a RC series network followed by a buffer; the switch is closed multiple times during the acquisition, without any reset of the capacitor.

With a proper selection of the filter parameters, the circuit implements the analog combination of two functions:

1) Sampled Acquisition by Gated Integrator
2) Exponential averaging of samples

To this aim, a basic hypothesis must hold, that is $T_F = RC \gg T_G$. In this scenario, we can derive the weighting function of this filter, starting from a key difference with a GI: in a classical GI, the capacitor is always reset between two consecutive acquisitions. In other words, the amount of charge on the capacitor of a GI at the beginning of the measure, i.e. when a pulse arrives, is zero. On the contrary, when the switch of a BI is open, that is in the time interval indicated with T_A in figure 10.15, the capacitor can't discharge and there is a buffer that copies the value at the output. As a result, when the switch is closed (time interval T$_G$) not only the incoming signal is acquired, but also the capacitor can discharge, meaning that the previously stored value is reduced by a factor r equal to:
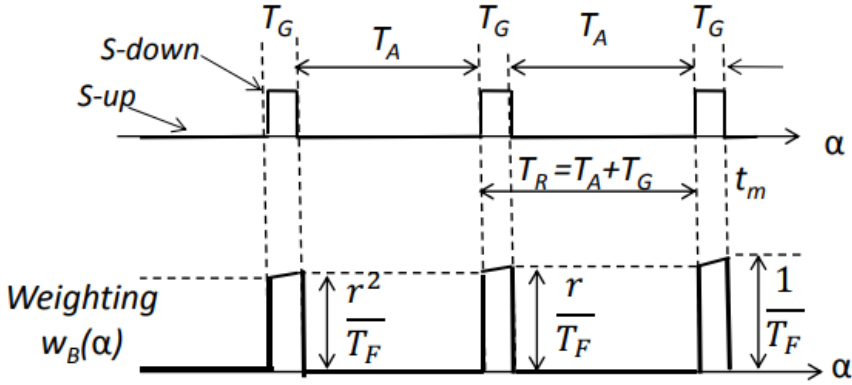
$$r = e^{-T_G/T_F}$$

Figure 10.15 control signal of the BI switch (top) and resulting weighting function (bottom)

The DC gain of this filter can be calculated observing that the weighting function $w_B(\alpha)$ of the BI can be obtained by splitting the weighting function $w_{RC}(\alpha)$ of the RC Integrator (RCI) into "slices" of width T_G which are placed over time when the signal is present. As a result, the final DC gain is equal to one.

The autocorrelation functions of BI and RCI are totally different but they feature the same value in zero.

$$K_{wwB}(0) = K_{wwRC}(0) = \frac{1}{2RC} = \frac{1}{2T_F}$$

Recall that the autocorrelation in zero is a key value to calculate the total amount of output noise; assuming a wide-band noise S_b with bandwidth $2f_n$ and autocorrelation time $2T_n$ we obtain

$$\overline{n_y^2} = S_B \cdot K_{wwB}(0) = S_b \cdot \frac{1}{2T_F} = \overline{n_x^2} \cdot \frac{T_n}{T_F}$$

The SNR at the output of the system can be written as the product of the input SNR and the improvement factor given by the filter, as follows:

$$\left(\frac{S}{N}\right)_y = \left(\frac{S}{N}\right)_x \cdot \sqrt{\frac{T_F}{T_N}}$$

It is worth noting that the SNR enhancement does not depend on the rate of the samples; indeed, it is obtained by averaging over a given number of samples and not over a given time interval. The average is updated every time a new pulse arrives, otherwise the switch is open, and the stored value is kept constant.

The Boxcar Integrator can be seen as the cascade of two filter stages:

a) A Gated Integrator featuring the same T_G and T_F of the BI; it gives a SNR improvement factor of:

$$\sqrt{\frac{T_G}{2T_n}}$$

b) An Exponential Averaging of the samples with attenuation ratio

$$r = e^{-\frac{T_G}{T_F}} \cong 1 - \frac{T_G}{T_F}$$

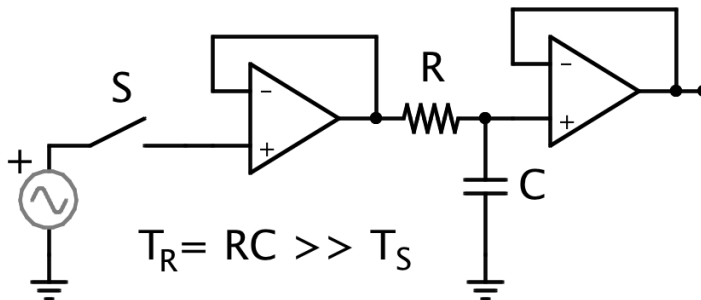that provides an additional improvement factor equal to:

$$\sqrt{\frac{1+r}{1-r}} \cong \sqrt{\frac{2}{1-r}} = \sqrt{\frac{2T_F}{T_G}}$$

The overall improvement factor can be computed as the product of the two:

$$\left(\frac{S}{N}\right)_y = \left(\frac{S}{N}\right)_x * \sqrt{\frac{T_G}{2T_n}} * \sqrt{\frac{2T_F}{T_G}} = \left(\frac{S}{N}\right)_x * \sqrt{\frac{T_F}{T_n}}$$

10.5 Ratemeter Integrator (RI)

In figure 10.15, the altered scheme of a Boxcar Integrator is reported. In particular, a buffer has been inserted between the switch S and the RC network.

Figure 10.15 Scheme of a Ratemeter Integrator

This simple modification radically changes the functionality of the circuit: a new exponential averager called **Ratemeter Integrator** (RI) is obtained.
First of all, keeping

$$T_R = RC \gg T_G$$

results into a DC gain lower than one in this case. Indeed, the low impedance at the buffer output allows the capacitor to discharge itself also when the switch is open and no signal is integrated, i.e. during the time interval T_A.
The signal replicas, that are acquired during T_G, have a constantly decreasing weight over time that now is equal to:

$$r = e^{-(T_G+T_A)/T_R} = e^{-T_S/T_R}$$

In this formula T_S is the sampling period and T_R is the decay time constant of the filter; as we can see r now depends on the sample rate.
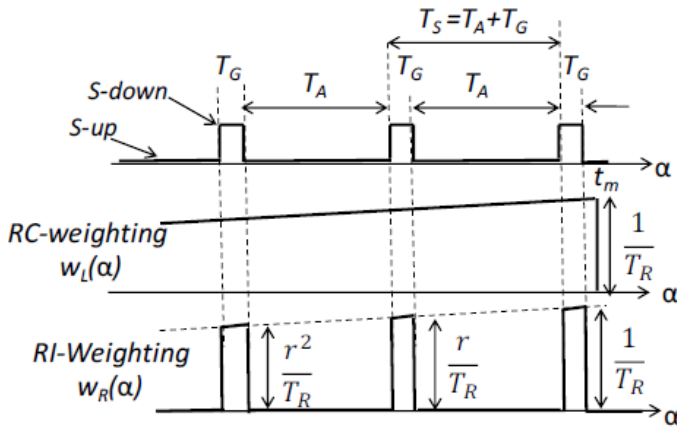

Figure 10.16 weighting function or a Ratemeter Integrator

The weighting function of the RI decreases over time even when the signal is not present; as a result, the area of the signal weighting function is for sure less than one, and in particular it is proportional to the sampling frequency

$$f_S = 1/T_S$$

The higher is the sampling frequency, the higher is the gain.

$$G = \int_{-\infty}^{+\infty} w_R(\alpha)\, d\alpha \cong \frac{T_G}{T_S} \cdot \int_{-\infty}^{+\infty} w_{RL}(\alpha)\, d\alpha \cong \frac{T_G}{T_S} = f_S \cdot T_G$$

In fact, for a given T_R the higher the sampling frequency, the higher is the number of replicas that are within the non-null weighting interval. Nevertheless, the sampling frequency depends on the signal rate, since sampling should occur only when the signal is present.

At this point it should be clear that the ratio between the weights of two samples depends now on the sample rate.

The Ratemeter Integrator is used in several applications. For example, consider a system providing a signal whose amplitude x_S is known and constant, while the repetition rate f_S of the signal is to be measured. In this scenario, an analogue Ratemeter Integrator would produce a quasi-DC output proportional to the repetition rate allowing the recovery of the information of interest. The described scenario is quite common in automotive applications to measure the time-varying rotation velocity of a car wheel, for example. In this case, a mark is placed on the wheel to generate a signal that depends on the mark itself (constant contribution) and on the speed of the wheel; a Ratemeter can be used to integrate the mark signal only when it is seen by an external sensor: in this way the output signal is proportional to wheel speed, that is the signal of interest.

As for the BI, we could analyse the Ratemeter Integrator as the cascade of two blocks: a Gated Integrator, with same T_G and T_R of the RI, which enhances the SNR by the factor

$$\sqrt{\frac{T_G}{2T_N}}$$

And an exponential averager of the samples with the attenuation ratio

$$r = e^{-T_S/T_R} \cong 1 - T_S/T_R$$

that depends now on the sample rate and enhances the SNR by

$$\sqrt{\frac{1+r}{1-r}} \cong \sqrt{\frac{2}{1-r}} = \sqrt{\frac{2T_R}{T_S}} = \sqrt{2T_R f_S}$$

In conclusion, we can see again that the overall SNR depends on the sample rate f_S

$$\left(\frac{S}{N}\right)_y = \left(\frac{S}{N}\right)_x \cdot \sqrt{\frac{T_G}{2T_N}} \cdot \sqrt{\frac{2T_R}{T_S}} = \left(\frac{S}{N}\right)_x \cdot \sqrt{f_S T_G \frac{T_R}{T_N}}$$

10.5 Boxcar Integrator versus Ratemeter Integrator

At the end of this chapter, the Ratemeter and the Boxcar Integrator are compared. In figure 10.17 a passive implementation of the two circuits (except for the buffers) is shown.
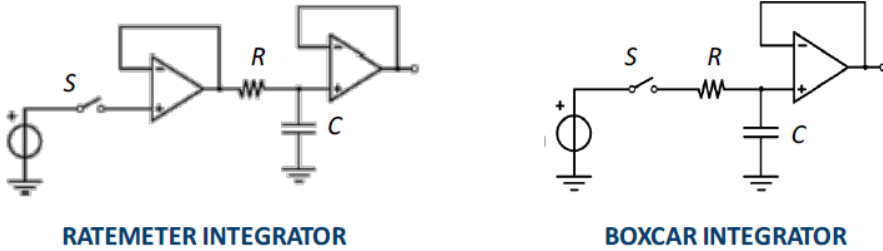
RATEMETER INTEGRATOR   BOXCAR INTEGRATOR

Figure 10.17 Scheme of passive RI and BI

In both filters the switch acts as gate on the input source, but in the RI it doesn't affect the RC integrator. RI does not have any hold state because there is a buffer to decouple, so the average is done with constant parameters for a given time.
On the contrary in the BI the switch acts also on the RC integrator: for this reason, the time constant of the integrator switches between $\tau = RC$, during the sampling phase (when the switch is close) and infinite in the hold phase.
The sample average is done on a given number of samples, defined by T_F/T_G.

If we consider active circuits, as those reported in figure 10.18, we have a DC gain $G = R_F/R_i$ that can be higher than 1. The active implementation can be useful especially for a Ratemeter Integrator that otherwise (passive implementation) features a gain much lower than 1.
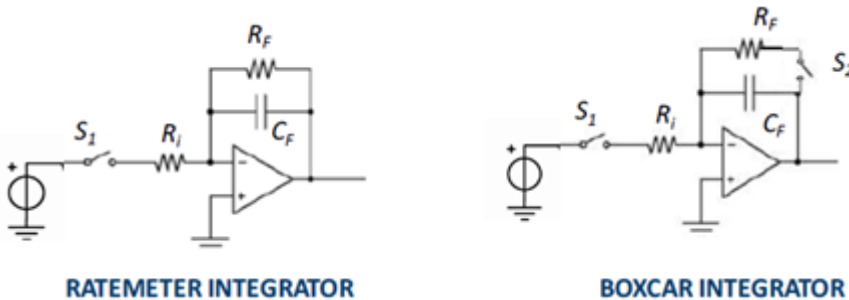

RATEMETER INTEGRATOR   BOXCAR INTEGRATOR

Figure 10.18 Scheme of active RI and BI

In the active version, the RI has the switch S_1 decoupled from the RC network thanks to the virtual ground of the OpAmp; therefore, the average is made with constant parameters for a given time defined by $\tau = R_F C_F$.
Also in the active BI the switch S_1 is decoupled from the RC integrator, while the second switch S_2 changes the time costant from $R_F C_F$, in the sampling phase, to infinite in the hold phase, thus providing a BI transfer function.

Optimum Filter

The Optimum Filter is a theoretical result: it basically describes the weighting function of the filter that ensures the highest possible signal to noise ratio for a given signal and noise. The study of the optimum filter strongly depends on signal and noise taken into account from time to time. Therefore, the theory is varied and complex and it dictates the upper limit of the achievable SNR.

11.1 Introduction

In the previous chapters we have analyzed many filtering techniques to improve the signal to noise ratio. Now, we want to discuss if a maximum SNR exists and how we can possibly design a filter to obtain it.

The general filtering theory answers positively to the first point: a maximum theoretical value for the signal to noise ratio always exists and it can be obtained by means of linear filters technique. However, we will see that the problem of the design of the optimum filter must be tackled and solved separately in each case. In the next paragraphs we will not expose deeply the optimum filtering theory, which is extremely wide and complex, but we will only determine the optimum filter for one particular case: the measure of the amplitude *A* of a signal *s(t)* buried in stochastic noise *n(t)*. First of all, we will consider the simplest type of signal: the pulse signal; we will introduce the optimum filter theory considering the basic case of pulse signals accompanied by stationary white noise, and then, that of a general signal in stationary white noise. Afterwards, we will extend the discussion to a general signal shape in stationary non-white noise. Finally, we will briefly consider the real case of a signal coming with a shot noise.

11.2 Optimum Filtering with White Stationary Noise

Pulse signals are a convenient starting point to expose optimum filter theory: they can carry information that is typically contained in the pulse amplitude, not in its shape or other parameters (e.g. pulse rise time, duration, etc.). This feature makes it quite simple to apply the optimum filtering theory to pulse signals. Moreover, pulse

signals are quite common in real applications, making this first part of the discussion particularly useful.

Here are two examples that better illustrate the importance of the pulse amplitude detection:

Automated analysis of biological cells:

Fluorescence-based analysis tools are often employed in biology: a cell in diluted solution, for example, can be labelled with a fluorescent dye that specifically binds to a particular component of the cell itself; this property can be exploited to infer the quantity of that particular component from the measurement of the optical signal. After labelling, the cells are conveyed by a laminar stream in a small duct and they cross a laser beam that excite fluorescence: the higher the amount of the labeled component, the higher will be the intensity of the luminous signal re-emitted by the cell. By measuring and classifying many pulses, the distribution of the component in the cell population can be obtained.

Ionizing radiation spectrometry:

Radiation detectors generate a current signal with charge proportional to the impinging radiation energy. Therefore, measuring the charge of each pulse and collecting the histogram of measurements provides the radiation distribution in energy, that is the energy spectrum. In this way it is possible to identify radionuclides in the source, to measure their quantity, to monitor radiation doses, etc.

The ideal case of pulse signals accompanied by stationary white noise is a good approximation for real cases where pulse signals are accompanied by wide-band noise, i.e. noise with:

- Narrow autocorrelation, i.e. **width much smaller than the signal duration**
- Wideband uniform spectrum, i.e. **upper band limit much higher than that of the signal**

To recover the pulse amplitude, a filter has to collect most of the signal and reject most of the noise. It's intuitive that its action should be:

- in the time domain, some sort of average of the signal and the white noise over the time interval occupied by the signal
- in the frequency domain, the preservation of the low frequency components in the range occupied by the signal and a cut-off of the high-frequency range where only white noise is present

This means that it a **low-pass filter (LPF) tailored to the signal** is a good candidate to achieve a high SNR. Here is an example of a Gated Integrator used to filter the pulse signal
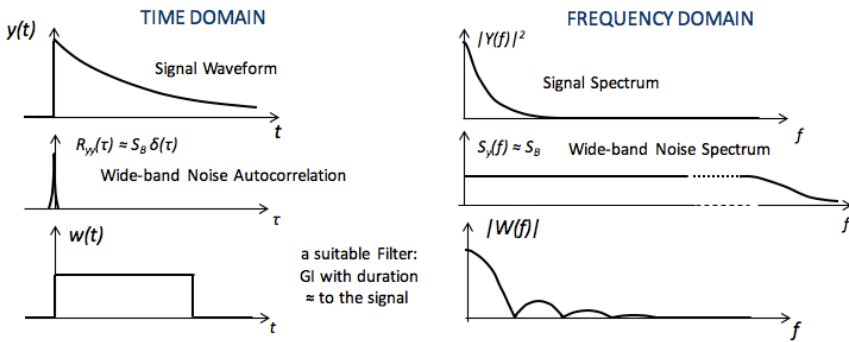
Figure 11. 1 Comparison between pulse signal and Gated Integrator Filter in Time Domain and Frequency Domain

We intuitively understood that a low pass filter can provide a high SNR for a pulse signal having a spectrum mainly concentrated on a low frequency range accompanied by wide-band noise. Nevertheless, we now want to find the optimum filter among the various low pass filters that we can select and implement. The issue is to find out **the optimal weighting function**.

11.2.1 Optimum Filter

The results that we will obtain from now on are valid for a generic signal, but during this analysis we maintain the example of the pulse signal for simplicity, keeping in mind that the dissertation can be immediately extended to other signal shapes.
The signal can be expressed pointing out the signal area A and the normalized waveform $b(t)$:

$$y(t) = A \cdot b(t)$$

with

$$\int_{-\infty}^{\infty} b(t) = 1$$

This expression of the signal means that we are supposed to know the shape of the signal, while the amplitude of it is the unknown parameter. In order to find the best result obtainable by low pass filtering a pulse signal, we analyze the SNR at the output of this simple filtering chain:
Both signal and noise enter the filter, which is fully characterized by the weighting function $w_m(\alpha)$ that we are trying to optimize. Then they are acquired at the output of the filter at the time instant t_m.
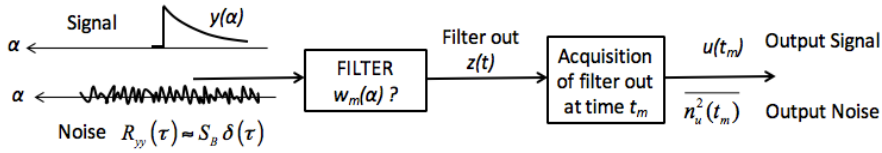
Figure 11. 2 Filtering of a signal in white stationary noise

The *signal* $u(t)$ acquired in the measurement is:

$$u(t_m) = \int_{-\infty}^{\infty} y(\alpha)w_m(\alpha)d\alpha = A \cdot \int_{-\infty}^{\infty} b(\alpha)w_m(\alpha)d\alpha = A \cdot k_{bw}(0)$$

The *noise* $n_u^2(t_m)$ acquired:

$$\overline{n_u^2(t_m)} = \int_{-\infty}^{\infty} R_{yy}(\alpha)k_{ww}(\alpha) = S_B \cdot \int_{-\infty}^{\infty} w_m^2(\alpha)d\alpha$$

Therefore, the signal to noise ratio results:

$$\left(\frac{S}{N}\right)^2 = \frac{u^2(t_m)}{n_u^2(t_m)} = \frac{A^2}{S_B} \cdot \frac{k_{bw}^2(0)}{k_{ww}(0)}$$

The $w_m(\alpha)$ that optimizes S/N for a given shape $b(\alpha)$ is found by exploiting the known property of correlation functions based on Schwartz's inequality:

$$k_{bw}^2(0) \leq k_{bb}(0) \cdot k_{ww}(0)$$

that is $\quad \dfrac{k_{bw}^2(0)}{k_{ww}(0)} \leq k_{bb}(0)$

The last inequality lead to the conclusion that:

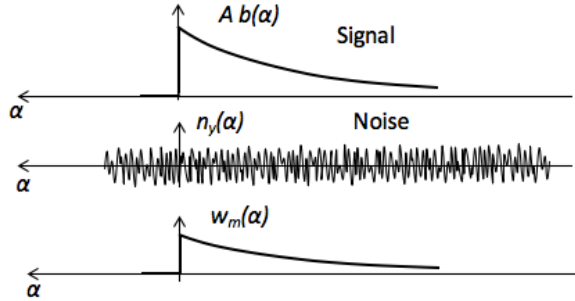$$\text{Max}\left[\frac{k_{bw}^2(0)}{k_{ww}(0)}\right] = k_{bb}(0)$$

The highest cross-correlation between the signal and the weighting function is obtained when the two have the same shape. Therefore, the maximum is achieved with **filter weighting function proportional to the signal shape**

$$w_m(\alpha) \propto b(\alpha)$$

Which normalized to unit area is:

$$w_m(\alpha) = b(\alpha)$$

11.3 Matched Filter

Figure 11. 3 Pulse Signal in white stationary noise, and matched filter which follows the shape of the signal

The best result in measurements of the amplitude of signal pulses accompanied by *stationary white noise*, which is represented in figure 11.3, is obtained with weighting function equal to the signal shape.

This conclusion is intuitive: since the noise is uncorrelated, the output noise power is the weighted sum of the noise instantaneous power at all times; since this power is equal at all times, it is convenient to give higher weight when the signal is higher, and a smaller one when the signal decreases; in other words, we need to follow the path of the signal in time.

The filter having weighting function $w_m(\alpha)$ matched to the signal shape $b(\alpha)$ is called **matched filter**.

The second stage in figure 11.2, shows that the signal is acquired by the system at a certain time t_m , therefore the optimal weighting function must have the same shape of the input signal, truncated in $t = t_m$.

Recalling that constant-parameters linear filters feature the following relation between weighting function and δ-response:

$$w(t, \tau) = h(t - \tau)$$

We can derive that the optimum filter with constant parameters in time is characterized by a delta δ-response like:

$$h_{opt} = w(t_m - t)1(t)$$

The h_{opt} has the same shape of the input signal, reversed with respect to the time axis and shifted by the quantity t_m.

11.4 Signal to Noise Ratio of the matched filter

The optimum SNR^2 provided by the *matched filter* is:

$$\left(\frac{S}{N}\right)^2 = \frac{A^2}{S_B} \cdot k_{bb}(0) = \frac{A^2}{S_B} \cdot \int_{-\infty}^{\infty} b^2(\alpha)d\alpha$$

This expression gives us the highest SNR value that we can obtain in case of a known shape signal $b(t)$ buried in a white noise having power spectrum S_B, and applying the optimum filtering technique we have just discussed. It is important to remark that the limit of the SNR is due to the nature of the specific signal and of the noise, and that it cannot be exceeded, not even from a theoretical point of view, or by developing alternative theories. In real cases, it is practically impossible to obtain the optimum value of the signal to noise ratio given by that formula, because it is obtained filtering the signal for an infinite time duration. Nevertheless, knowing the upper SNR boundary is a very useful tool and benchmark. In fact, we can use this result to determine the performance of a feasible filter.

Now we will discuss various interpretations of the signal to noise ratio of the optimum filter.

11.4.1 First interpretation:

We demonstrated that the optimal weighting function is the one proportional to the function in time of the signal we want to acquire, since it maximizes the output signal to noise ratio.

Now we write the signal to noise ratio in terms of the optimum filter:

$$\left(\frac{S}{N}\right)^2 = \frac{A^2}{S_B} \cdot k_{bb}(0) = \frac{A^2}{S_B} \cdot \int_{-\infty}^{\infty} b^2(\alpha)d\alpha$$

Where $k_{bb}(0)$ derived from the previous computations, and the integral from the known relation between the function in time and its autocorrelation in zero.

Now let's reflect on the term of this definition of SNR.

The energy E_y of a signal $A \cdot b(t)$ is:

$$E_y = A^2 \int_{-\infty}^{\infty} b^2(\alpha)d\alpha = A^2 \int_{-\infty}^{\infty} B^2(f)df$$

Where we relied on the Parseval Relation to pass from the time domain to frequency in the last equality. So, we see that $(S/N)^2$ is simply:

$$\left(\frac{S}{N}\right)^2_{opt} = \frac{E_y}{S_B} = \frac{signal\ energy}{noise\ power\ density\ (bilateral)}$$

The energy of the normalized signal $b(t)$

$$E_b = \int_{-\infty}^{\infty} b^2(\alpha)\, da = \int_{-\infty}^{\infty} B^2(f)\, df$$

Hence

$$E_y = A^2 \cdot E_b$$

We rewrite the signal to noise ratio by setting in evidence the amplitude of the input signal:

$$\left(\frac{S}{N}\right)_{opt}^2 = A^2 \cdot \frac{E_b}{S_B}$$

The minimum measurable amplitude A_{min} is defined as the amplitude that gives

$$\left(\frac{S}{N}\right)^2 = 1$$

Therefore:

$$A_{min} = \frac{\sqrt{S_B}}{\sqrt{E_b}} = \frac{\sqrt{S_B}}{\sqrt{\int_{-\infty}^{\infty} b^2(\alpha) d\alpha}} = \frac{\sqrt{S_B}}{\sqrt{\int_{-\infty}^{\infty} B^2(f) df}}$$

According to this expression, the minimum detectable signal is given by the ratio of the spectral density of the noise and the energy of the normalized signal.

11.4.2 Second Interpretation:

A second interpretation of $(S/N)_{opt}$ is found by taking in account that the input noise is only approximately white: in fact, it has finite power $\overline{n_y^2}$ and finite width $2T_n$ of the autocorrelation.
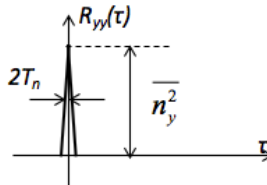

Figure 11. 4 Approximation of a real wide band noise

$$S_B = \overline{n_y^2} \cdot 2T_n$$

We can thus rewrite the expression we found in the first interpretation of the signal to noise ratio as:

$$\left(\frac{S}{N}\right)_{opt}^2 = \frac{E_y}{S_B} = \frac{E_y/2T_n}{\overline{n_y^2}}$$

Which highlights that the energy of the signal is concentrated in the interval $2T_n$.

11.3 Optimum Filter of Signals with any Stationary Noise

Optimum filters to measure the amplitude of pulse signals accompanied by any stationary noise can be obtained by *extending the result achieved for white noise.*
Before starting the discussion on how to extend the optimum filter theory to any stationary noise, here are some useful considerations to keep in mind:

1. The operations carried out by linear filters with constant parameters are always reversible
2. In principle, it is always possible to convert any given generic noise into white noise using a *whitening filter* with constant parameters, characterized by a particular transfer function

$$|H_b(f)| \propto 1/S_x(f)$$

3. It is particularly easy to evaluate the effects of a filter on the white noise

These considerations lead us to simplify the problem of the design of the optimum filter dividing it in two phases, corresponding to two cascaded stages.
The *first stage* is a constant parameter filter which aims at transforming the stationary input noise (characterized by the autocorrelation function $R_{yy}(\tau)$), into a white noise, characterized by the well-known behavior of the autocorrelation in time and the frequency spectrum (see chapter 3 about white noise).
The *second stage* is a linear filter characterized by the weighting function $w_m(\;)$ which has the role of optimizing the (S/N). The reversibility is an essential property: under this hypothesis nothing is lost in the transformation and whatever is done by the whitening filter can be reversed by the following filters. We can then proceed towards the optimum filter design, since we know what to do in the situation at the output of the whitening filter: we have pulse signal (referring to the pulse signal example) accompanied by the white stationary noise and we know that a matched filter performs the optimum filtering.
To sum up, the optimum filter in the case of any stationary noise will consist of:

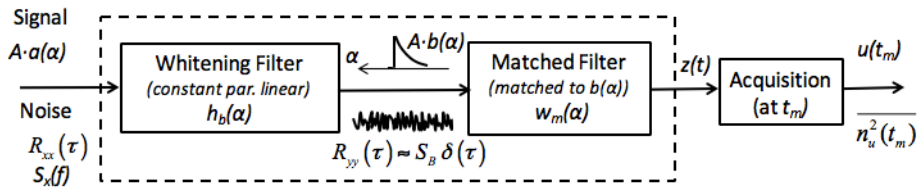*OPTIMUM FILTER= WHITENING FILTER + MATCHED FILTER*

Figure 11. 5 Filtering of a signal in any stationary noise

In the figure 11.5 the cascaded structure we are going to study.
Some observations about this cascaded structure:

- **The whitening filter modifies the waveform of the signal**; hence the following filter is matched to this modified signal, not to the original input signal.
- The subdivision of the optimum filter in whitening filter and matched filter is a useful theoretical approach to analyze the problem and find the overall optimal weighting, but it is not the necessary structure of the optimum filter. This means that in principle, we can find the optimal weighting function by combining these two stages, then we can implement this function by employing any kind of linear filter. This degree of freedom is precious because many times it is quite difficult to implement the noise-whitening filter.
- For a given noise with spectrum $S_x(f)$ the *whitening* filter is a constant-parameter linear filter that has transfer function $H_b(f)$ such that $|H_b(f)|^2 \propto 1/S_x(f)$ (in time domain: filter autocorrelation function $k_{bb}(\tau)$ such that the convolution with the noise autocorrelation $R_{xx}(\tau)$ produces a δ-like autocorrelation $R_{yy}(\tau) \propto \delta(\tau)$)
- The action of the whitening filter is more evident in cases where the actual noise results from white noise filtered by some circuit. For example, consider the Johnson noise of a resistor passed through an amplifier with upper band-limit set by a simple pole at low frequency. The whitening filter simply reverts the filtering by the amplifier with a transformation that cancels the low-pass pole.

11.4 Optimum Filter with Shot Noise

To investigate the optimum filter with shot noise, we should focus on the real cases in which shot current is dominant over all other noise contributions.
To illustrate that situation, the output of photomultiplier tube (PMT) can be considered.
 We will describe in detail this kind of photodetectors in the second part of this book. At this moment we can consider this sensor like an amplifier with extremely high internal gain. Since the amplification rate of PMT is really high, the noise of the following stage can be neglected.

The current of the PMT consists of 2 different components:

- a constant (DC) current I_D, due to the background light that reaches the photocathode plus the dark current
- a current signal $I_S(t)$ due to the optical signal detected

Due to the very short autocorrelation time, the PMT shot noise can be considered as white noise with 2 defined components:

- a stationary noise with constant intensity, i.e. with autocorrelation

$$R_{iD}(\tau) \approx S_D \, \delta(\tau)$$

- a non-stationary noise with intensity varying in time, i.e. with autocorrelation

$$R_{iS}(\alpha,\tau) \approx S_S(\alpha)\delta(\tau)$$

The stationary component has (bilateral) spectral density

$$S_D = Q \, I_D$$

The non-stationary component has (bilateral) spectral density

$$S_S(a) = Q \, I_S(a)$$

If we sum up these 2 components, we obtain:

$$S_B(a) = S_D + S_S(a) = Q * [\, I_D + I_S(a)]$$

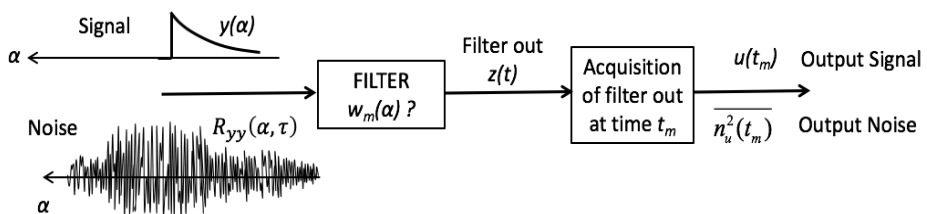Let's take a look at the weighting function in case of non stationary white nose.

Figure 11. 7 Acquisition chain for pulse signal and non-stationary noise

So, we have:
- Narrow autocorrelation, with width much smaller than the signal duration, therefore well approximated by a δ-function
- Intensity variable in time, i.e. area S_B of the δ-function variable with time

$$R_{yy}(\alpha,\tau) \approx S_B(\alpha)\delta(\tau)$$

The signal and noise acquired in the measurement are

$$u(t_m) = A \cdot \int_{-\infty}^{\infty} b(\alpha) w_m(\alpha) d\alpha$$

$$\overline{n_u^2(t_m)} = \int_{-\infty}^{\infty} S_B(\alpha) w_m^2(\alpha) d\alpha$$

Therefore

$$\left(\frac{S}{N}\right)^2 = \frac{u^2(t_m)}{\overline{n_u^2(t_m)}} = A^2 \cdot \frac{\left[\int_{-\infty}^{\infty} b(\alpha) w_m(\alpha) d\alpha\right]^2}{\int_{-\infty}^{\infty} S_B(\alpha) w_m^2(\alpha) d\alpha}$$

In this case the denominator is not $k_{ww}(0)$ because the integral includes also $S_B(\alpha)$. However, $S_B(\alpha)$ is among the data, i.e. it is known and not modifiable. We can thus define a weighting function $w_{m,a}(\alpha)$ *matched* to the shape of the signal *and adapted* to the variation of noise intensity, which has a known relation with the true weighting function $w_m(\alpha)$

$$w_{m,a}(\alpha) = \sqrt{S_B(\alpha)} \cdot w_m(\alpha)$$

and write

$$\overline{n_u^2(t_m)} = \int_{-\infty}^{\infty} w_{m,a}^2(\alpha) d\alpha = k_{ww,a}(0)$$

We can rewrite the numerator integral to introduce the adapted weighting function

$$u(t_m) = A \cdot \int_{-\infty}^{\infty} \frac{b(\alpha)}{\sqrt{S_B(\alpha)}} \cdot \sqrt{S_B(\alpha)} \cdot w_m(\alpha) d\alpha = A \cdot \int_{-\infty}^{\infty} \frac{b(\alpha)}{\sqrt{S_B(\alpha)}} \cdot w_{m,a}(\alpha) d\alpha$$

and define an adjusted signal shape $b_a(\alpha)$ adapted to the noise variation

$$b_a(\alpha) = \frac{b(\alpha)}{\sqrt{S_B(\alpha)}}$$

We can thus write the numerator as

$$u\left(t_m\right) = A \cdot \int_{-\infty}^{\infty} b_a\left(\alpha\right) w_{m,a}\left(\alpha\right) d\alpha = A \cdot k_{bw,a}\left(0\right)$$

and the S/N as

$$\left(\frac{S}{N}\right)^2 = A^2 \cdot \frac{k_{bw,a}^2\left(0\right)}{k_{ww,a}\left(0\right)}$$

and we can exploit again the property

$$\frac{k_{bw,a}^2\left(0\right)}{k_{ww,a}\left(0\right)} \le k_{bb,a}\left(0\right)$$

NB: $w_{m,a}$ and b_a are not the real weighting and signal, they are only a mathematical trick!
The maximum value is reached by taking

$$w_{m,a}\left(\alpha\right) = b_a\left(\alpha\right)$$

Given the definitions of $w_{m,a}(\alpha)$ and $b_a(\alpha)$, this leads to a true weighting function

$$w_m\left(\alpha\right) \propto \frac{b\left(\alpha\right)}{S_B\left(\alpha\right)}$$

This is an *adapted matched filter*, that is adapted to the non-stationary noise. The result is quite intuitive: since the noise power varies in time, it is convenient on one hand to give higher weight when the signal is higher, on the other hand to give lower weight where the noise is higher. The optimum S/N thus obtained is

$$\left(\frac{S}{N}\right)^2_{opt,ns} = A^2 \cdot \max\left[\frac{k_{bw,a}^2\left(0\right)}{k_{ww,a}\left(0\right)}\right] = A^2 \cdot k_{bb,a}\left(0\right) = A^2 \cdot \int_{-\infty}^{\infty} \frac{b^2\left(\alpha\right)}{S_B\left(\alpha\right)} d\alpha$$

By comparing to the previous result with stationary noise

$$\left(\frac{S}{N}\right)^2_{opt} = \frac{A^2}{S_B} \cdot \int_{-\infty}^{\infty} b^2\left(\alpha\right) d\alpha = \frac{signal\ energy}{noise\ power\ density\ (bilateral)}$$

we note that in the non-stationary case

$$\left(\frac{S}{N}\right)^2_{opt,ns} = A^2 \cdot \int_{-\infty}^{\infty} \frac{b^2(\alpha)}{S_B(\alpha)} d\alpha = \int_{-\infty}^{\infty} \frac{A^2 b^2(\alpha) d\alpha}{S_B(\alpha)}$$

- the integrand is the energy of the element of the signal in the interval $d\alpha$ at time α, divided by the noise power density $S_B(\alpha)$ at that time;
- that is, the integrand is the elementary S/N for the optimized measurement of a signal element at time α.
- The total optimized $(S/N)_{opt,ns}$ can be considered as a sum of these elementary contributions.

We can highlight that in real cases the input noise can be approximately white, with autocorrelation with finite width $2T_n$ (anyway much shorter than the signal duration) and finite noise power $\overline{n_y^2(\alpha)}$ that varies with time

We can thus write

$$\left(\frac{S}{N}\right)^2_{opt,ns} = A^2 \cdot \int_{-\infty}^{\infty} \frac{b^2(\alpha)}{S_B(\alpha)} d\alpha = \int_{-\infty}^{\infty} \frac{A^2 b^2(\alpha)}{n_y^2(\alpha)} \frac{d\alpha}{2T_n}$$

which shows that the $(S/N)_{opt,ns}$ is the sum of all the (S/N) of elementary measurements carried out by sampling the signal at all time instants.

As shown, with non-stationary white noise the optimal weighting depends not only on the signal waveform, but also on the variation in time of the noise intensity as follows

$$w_m(\alpha) \propto \frac{b(\alpha)}{S_B(\alpha)} \propto \frac{I_S(\alpha)}{S_D + S_S(\alpha)} \propto \frac{I_S(\alpha)}{I_D + I_S(\alpha)}$$

The shape of the weighting function thus changes as the relative size of the pulse current and background current change.

The optimal weighting in this case is

$$w_m(\alpha) \propto \frac{b(\alpha)}{S_B(\alpha)} \propto \frac{I_S(\alpha)}{S_D + S_S(\alpha)} \propto \frac{I_S(\alpha)}{I_D + I_S(\alpha)}$$

If the pulse current is much smaller than the background, i.e. with $I_S << I_D$ the noise is almost stationary and the weighting tracks the signal shape, it is the matched filter

$$w_m(\alpha) \approx \frac{I_S(\alpha)}{I_D}$$

If the pulse current is much greater than the background. i.e. with $I_S >> I_D$ the optimal weight is practically constant over the signal extent. The optimal processing is to *integrate over the interval occupied by the signal*, i.e. simply to collect all the pulse charge; it can be simply implemented by a GI

$$w_m(\alpha) \approx 1$$

If pulse and background are comparable, the optimal weighting function has an intermediate shape between the two cases above
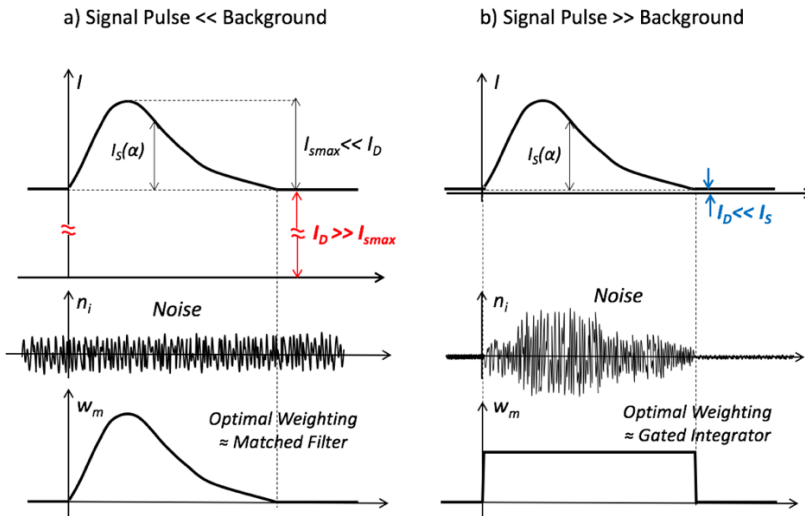
Figure 11. 8 Current pulse and related optimal weighting function in case of dominant background noise (left) or dominant pulse-related nois

Optimum filter 2

In this chapter we apply the optimum filter theory to signals coming from high-impedance sensors. We will study the whitening filter and the optimum filter. Finally, we will make some practical considerations about the approximation of this solutions in the real world.

12.1 High Impedance Sensor and Low-Noise Preamplifiers

Let us consider sensors that are seen by the circuits connected to their terminals as generators of current signals with high internal impedance (typically a small capacitance C_S with a high resistance R_S in parallel).

Typical examples are: p-i-n junction photodiodes and other photodetectors (CCDs, vacuum tube, photodiodes, etc.), piezoelectric Force Sensors in quartz or other piezoelectric ceramic materials.

These sensors typically have internal noise sources (e.g. shot current noise of a reverse-biased junction current) modeled by a current noise generator in parallel to the signal generator. The equivalent circuit model is sketched in figure 12.1.
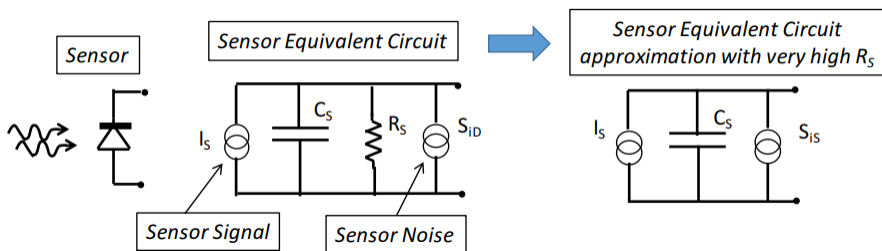
Figure 12.1 equivalent circuit model of an high impedance sensor

The amplifier can be modeled as shown in figure 12.2, where the R_{iA} is the true physical resistance between the input terminals (NOT the dynamic input resistance modified by the feedback in the amplifier; e.g. not the low dynamic resistance of the virtual ground of an operational amplifier).

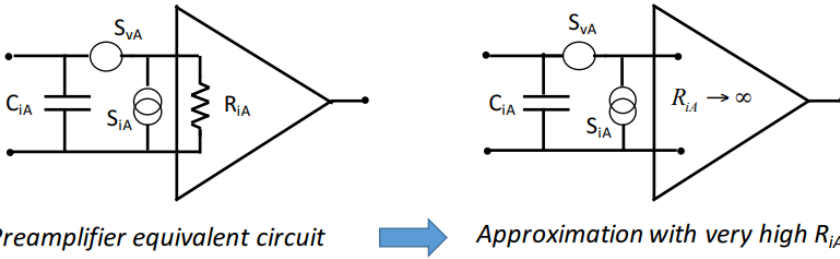Besides shot noise of bias currents, the S_{iA} includes Johnson resistor noise of R_{iA}

Figure 12.2 equivalent model of the amplifier connected to the sensor

$$S_{iR} = \frac{4kT}{R_{iA}}$$

The current noise directly faces the sensor current signal I_S. If R_{iA} is small the S_{iR} is overwhelming (e.g. with R_{iA}=50 Ω it is $\sqrt{S_{iR}} \approx 18pA/\sqrt{Hz}$) while other contributions to $\sqrt{S_{iA}}$ are much lower (about $1pA/\sqrt{Hz}$ or lower).
As a result, low-noise operation of high-impedance sensors requires a preamplifier with high input resistance R_{iA}.

So, the overall equivalent circuit of high impedance sensor and Preamplifier (approximation valid for very high sensor resistance Rs $\rightarrow\infty$) could be:
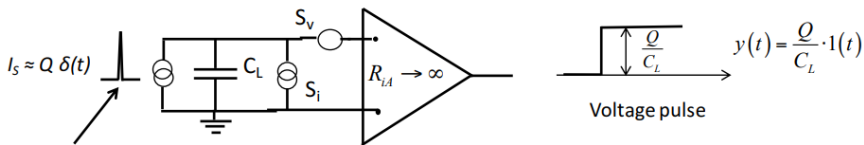
Figure 12.3 overall equivalent circuit of high impedance sensor and Preamplifier

The sensor generates a current pulse and we want to measure total charge Q; the amplitude of the signal step is Q/C_L while, from the noise point of view, we have:

$C_L = C_S + C_{iA}$ *total capacitance load*
$S_v = S_{vA}$ *voltage noise generator (wideband noise spectrum)*
$S_i = S_{iD} + S_{iA}$ *current noise generator (wideband noise spectrum)*

The voltage noise spectrum S_n has two components, it is NOT white at the preamplifier output

$$S_n(\omega) = S_v + \frac{S_i}{\omega^2 C_L^2}$$

The noise power spectrum sketched in figure 12.4 clearly shows the superposition of these two components: S_v is constant along all the frequency range, while the current component filtered by the load capacitance introduces a divergent trend towards low frequencies.
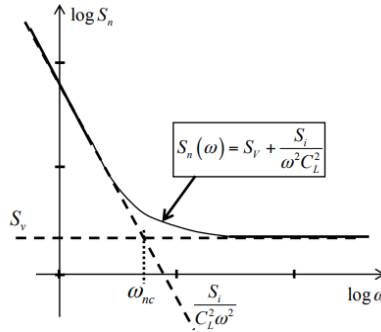

Figure 12.4 total noise power spectrum

The frequency where the two noise components have equal magnitude is called Noise corner angular frequency $\omega_{nc.}$, and it is derived by equating the two noise components.

$$S_v = \frac{S_i}{\omega^2 C_L^2} \qquad\qquad \omega_{nc} = \frac{\sqrt{S_i}}{C_L \sqrt{S_v}}$$

We define $T_{nc} = 1/\omega_{nc}$ as noise corner frequency time constant:

$$T_{nc} = \frac{1}{\omega_{nc}} = \frac{\sqrt{S_v}}{\sqrt{S_i}} C_L$$

T_{nc} and ω_{nc} are fundamental parameters of the optimum filter: we will see that T_{nc} rules the duration of the filter weighting and the ω_{nc} the filter bandwidth. We can define the Noise Corner resistance as:

$$R_{nc} = \sqrt{S_v}/\sqrt{S_i}$$

so that $T_{nc} = R_{nc} * C_L$

R_{nc} typically ranges from tens to hundreds of $k\Omega$, while T_{nc} ranges from a few nanoseconds to some hundreds of nanoseconds.
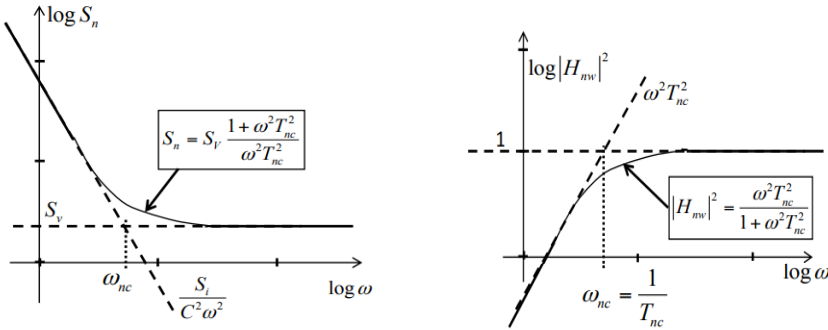
12.3 Noise whitening filter

All the considerations made in the previous chapter about optimum filter are valid for white noise only. When noise power spectrum is somehow colored, we saw that

need a whitening filter to make the output noise to be white. So, starting from the spectral density given by the sensor and the preamplifier:

$$S_n(\omega) = S_v\left(1 + \frac{S_i}{\omega^2 S_v C_L^2}\right) = S_v\left(1 + \frac{1}{\omega^2 T_{nc}^2}\right) = S_v\frac{1 + \omega^2 T_{nc}^2}{\omega^2 T_{nc}^2}$$

We see that the noise spectrum has a pole at $\omega_p = 0$ and a zero at $\omega_z = \omega_{nc} = 1/T_{nc}$

Figure 12.5 bode diagram of the total noise power spectrum (left) and whitening filter (right)

To make the output noise to be white, the noise whitening filter H_{nw} must
- Cancel the pole with a zero at $\omega = 0$
- Cancel the zero with a pole at $\omega_z = \omega_{nc} = 1/T_{nc}$

We can write the whitening filter transfer function (squared):

$$|H_{nw}(\omega)|^2 = \frac{\omega^2 T_{nc}^2}{1 + \omega^2 T_{nc}^2}$$

It is a simple high-pass filter.

A simple high-pass filter has a time constant $T_{hp} = T_{nc} = R_w * C_w$ that we can set to be equal to the noise corner frequency T_{nc}.

$$H_{nw}(\omega) = \frac{j\omega R_w C_w}{1 + j\omega R_w C_w} = \frac{sR_w C_w}{1 + sR_w C_w}$$

Appling the filter of figure 12.5, the output noise power spectrum is now equal to S_v, which is constant along the whole spectrum.
The signal pulse I_s is filtered by both C_L and $H_{nw}(\omega)$ so the overall effect is that of a low pass filtering with time-constant T_{nc}. The voltage step $y_s(t)$ is turned into a short exponential decaying pulse.
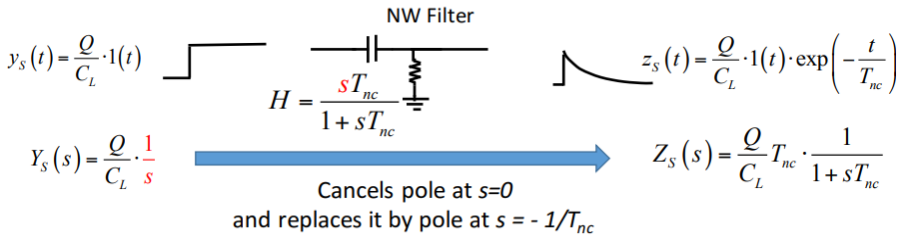
NW Filter

$$y_s(t) = \frac{Q}{C_L} \cdot 1(t)$$

$$H = \frac{sT_{nc}}{1 + sT_{nc}}$$

$$z_s(t) = \frac{Q}{C_L} \cdot 1(t) \cdot \exp\left(-\frac{t}{T_{nc}}\right)$$

$$Y_s(s) = \frac{Q}{C_L} \cdot \frac{1}{s}$$

$$Z_s(s) = \frac{Q}{C_L} T_{nc} \cdot \frac{1}{1 + sT_{nc}}$$

Cancels pole at s=0
and replaces it by pole at s = - 1/T_{nc}

Figure 12.8   Whitening filter effect on the signa

12.4 Matched Filter

The matched filter must have the same shape of the signal after the whitening process. The final optimum filter with colored noise will be the cascade of whitening and matched filters.
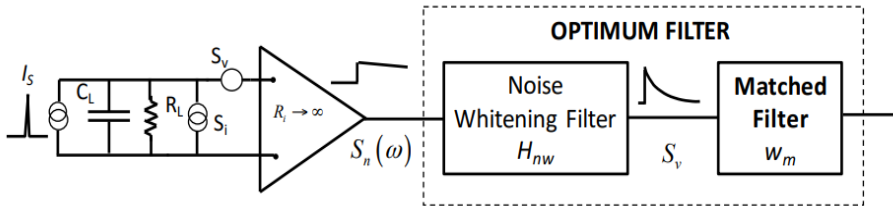
Figure 12.9   Explicit representation of the optimum filter split into the two stages

As we said before, the signal coming out from the noise whitening filter is:

$$z_s(t) = \frac{Q}{C_L} \cdot 1(t) \cdot \exp\left(-\frac{t}{T_{nc}}\right)$$

Therefore, the matched filter must have the same shape:

$$w_M = 1(t) \cdot \frac{1}{T_{nc}} \exp\left(-\frac{t}{T_{nc}}\right)$$

As a result:

$$\eta_o^2 = \left(\frac{S}{N}\right)_{opt}^2 = \frac{\left[\int_{-\infty}^{+\infty} z_s(\alpha) w_m(\alpha)\, d\alpha\right]^2}{S_v \cdot \int_{-\infty}^{\infty} w_m^2(\alpha)\, d\alpha} = \frac{Q^2}{C_L^2} \frac{T_{nc}^2}{S_v} \int_{-\infty}^{\infty} w_m^2(\alpha)\, d\alpha = \frac{Q^2}{C_L^2} \frac{1}{2} \frac{T_{nc}}{S_v}$$

At the output of the optimum filter (i.e. of the matched filter) we have:

Signal:

$$S_o = \int_0^\infty z_s(\alpha) w_m(\alpha) d\alpha = \frac{Q}{C_L} \frac{1}{T_{nc}} \int_0^\infty \exp\left(-\frac{2\alpha}{T_{nc}}\right) d\alpha = \frac{1}{2} \frac{Q}{C_L}$$

Noise:

$$\sqrt{n_o^2} = \sqrt{S_v} \cdot \sqrt{k_{ww}(0)} = \sqrt{S_v} \frac{1}{T_{nc}} \sqrt{\int_0^\infty \exp\left(-\frac{2\alpha}{T_{nc}}\right) d\alpha} = \sqrt{\frac{S_v}{2T_{nc}}}$$

So, the S/N is:

$$\eta_o = \frac{S_o}{\sqrt{n_o^2}} = \frac{1}{2} \frac{Q}{C_L} \sqrt{\frac{2T_{nc}}{S_v}}$$

As an example, we can apply these results to measure the signal charge Q, so we can consider the noise in terms of equivalent noise charge q^2_{no}

$$q_{no}^2 = \frac{2C_L^2 S_v}{T_{nc}}$$

We get another expression of S/N:

$$\eta_o^2 = Q^2 \frac{T_{nc}}{2C_L^2 S_v} = \frac{Q^2}{q_{no}^2}$$

Recalling that $T_{nc} = C_L * R_{nc} = C_L * (\sqrt{S_v}/\sqrt{S_i})$ we can express q^2_{no} in terms of the main parameters of the sensor and preamplifier:

$$q_{no}^2 = 2C_L \cdot \sqrt{S_v} \cdot \sqrt{S_i}$$

q^2_{no} is the charge pulse for which S/N = 1. It represents the minimum measurable pulse Q_{min} and the minimum difference between two distinguishable pulses. We can express q^2_{no} as the minimum number of electron charges (NB electron charge $e = 1,6 * 10^{-19}$ C) readable at the output of the system.

$$\sqrt{N_{eo}^2} = \frac{\sqrt{q_{no}^2}}{e} = \frac{\sqrt{2C_L \sqrt{S_v} \sqrt{S_i}}}{e} = 278 \cdot \sqrt{C_L[pF] \cdot S_v^{\frac{1}{2}}[nV/\sqrt{Hz}] \cdot S_i^{\frac{1}{2}}[pA/\sqrt{Hz}]}$$

It is instructive to evaluate and compare typical values of the optimized noise charge in typical cases of, for example, PIN photodiodes coupled to high-impedance preamplifiers.

a) PIN with wide sensitive area (1mm diameter) and discrete-component preamplifier with moderately low-noise. Fairly high capacitance of PIN & connections (C_L about 10pF) and not-so-low preamp noise (typical 20nV/$\sqrt{Hz}$) set a fairly high limit

$$\sqrt{N_{eo}^2} \approx 1800 \text{ electrons}$$

b) PIN with reduced sensitive area (about 100um diameter) and discrete-component preamplifier with low-noise. Reduced capacitance of PIN & connections (CL about 1pF) and reduced preamp noise (typical 10nV/$\sqrt{Hz}$) bring a strong reduction

$$\sqrt{N_{eo}^2} \approx 120 \text{ electrons}$$

c) PIN with a small sensitive area (less than 20um diameter) integrated with a very low-noise preamplifier. Capacitance minimization by monolithic integration (CL about 0,1 pF) and very low preamp noise (2nV/$\sqrt{Hz}$) bring further progress

$$\sqrt{N_{eo}^2} \approx 10 \text{ electrons}$$

12.5 Optimum Filtering with Finite Readout Time

The exponential weighting function w(t) of the matched filter reaches zero only at t → ∞, so that the measurement of the pulse in principle is available with infinite delay after the pulse onset. However, most of the weight of the matched filter is over a finite time interval of a few time constants T_{nc}. So we are interested in evaluating the result of a practical approximation of the matched filter.

We saw in the previous paragraphs how the whitening filter is simple and easy to implement. So, to compute the complete optimum filter, it is sufficient to find out how to approximate the matched filter.

The previous results, both in the frequency and time domain, pointed out that the overall effect of the acquisition chain (including the whitening filter) on the original signal is the one of a low pass filter; indeed, the original delta-like pulse is translated into an exponential pulse as reported in figure 12.8.

Therefore, a simple approximation of the matched filter could be a simple RC integrator. However, it is *extremely important* to point out that this is only an approximation, let's see it in detail.

With RC=T$_{nc}$ the delta response h(t) of the low pass filter seems identical to the weighting function W$_m$(t) of the matched filter. However, **the RC weighting function W$_f$(t) is reversed compared to the shape as Wm(t) of the matched filter.**
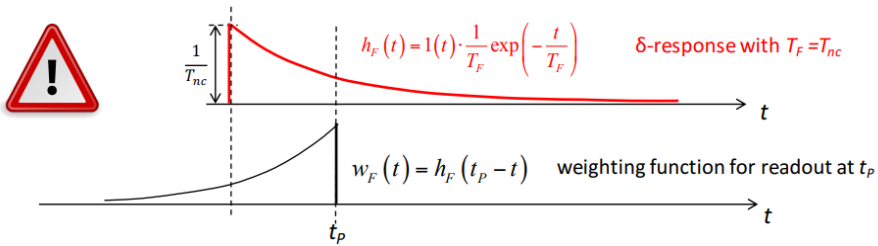

$$h_F(t) = 1(t) \cdot \frac{1}{T_F} \exp\left(-\frac{t}{T_F}\right)$$

δ-response with $T_F = T_{nc}$

$$w_F(t) = h_F(t_P - t)$$ weighting function for readout at t_P

Figure 12.10 δ-response and weighting function of RC single pole filter

From the noise point of view the filtering is the same of the matched filter, since it is independent from the weighting function inversion; the output is white noise with band limit set by a simple pole with time constant T$_f$.

Signal filtering instead is different from the matched filter, since it is affected by the time-inversion.

Sensor output current
$$I_s(t) = Q \cdot \delta(t)$$

Preamplifier output
$$y_S(t) = \frac{Q}{C_L} \cdot 1(t)$$

Whitening filter output
$$z_S(t) = \frac{Q}{C_L} \cdot 1(t) \cdot \exp\left(-\frac{t}{T_{nc}}\right)$$

δ-response of RC integrator
$$h_F(t) = 1(t) \cdot \frac{1}{T_F} \exp\left(-\frac{t}{T_F}\right)$$ with $T_F = T_{nc}$

Output of RC integrator
$$u_F(t) = z_S * h_F = \frac{Q}{C_L} \frac{t}{T_{nc}} \exp\left(-\frac{t}{T_{nc}}\right)$$

Weighting function of RC integrator
$$w_F(t) = h_F(t_P - t)$$

peaking time $t_P = T_{nc}$
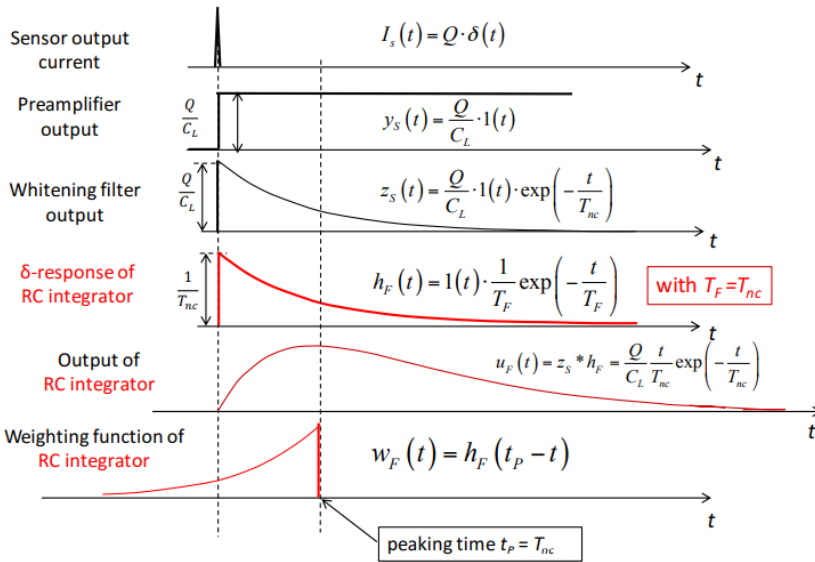
Figure 12.11 RC single pole filter features

The shape at the output of the matched filter can be seen as the overall weighting function of the optimum filter (whitening + matched). The δ response and w-function are shown in the next figure:

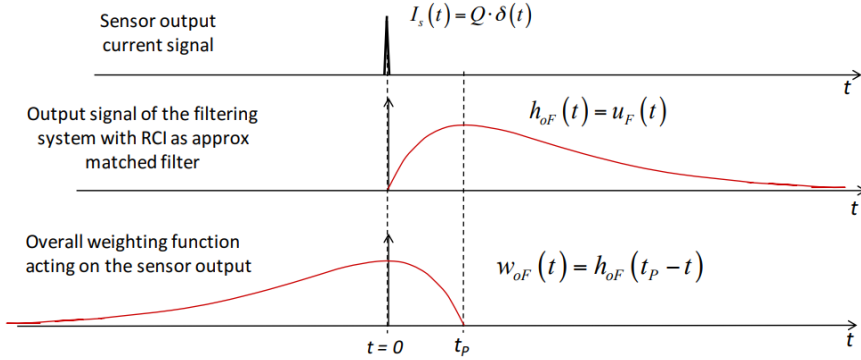**Figure 12.12** δ-response and weighting function of the overall optimum filter

The RC output signal waveform is:

$$u_F(t) = \frac{Q}{C_L} \frac{t}{T_{nc}} \exp\left(-\frac{t}{T_{nc}}\right)$$

Which peaks at time $t = T_{nc}$ with value:

$$s_F = u_F(T_{nc}) = \frac{1}{e} \frac{Q}{C_L}$$

The total noise power collected by the filter is:

$$\sqrt{\overline{n_F^2}} = \sqrt{S_v} \cdot \sqrt{k_{hh}(0)} = \sqrt{\frac{S_v}{2T_{nc}}}$$

Which gives a S/N:

$$\eta_F = \frac{s_F}{\sqrt{\overline{n_F^2}}} = \frac{1}{e} \frac{Q}{C_L} \sqrt{\frac{2T_{nc}}{S_v}}$$

Comparing the RC approximation with the ideal optimum filter system we see that maximum signal amplitude has been reduced by a factor:

$$s_F = \frac{2}{e} s_o \approx 0{,}736 \cdot s_o$$

As said before, the noise collected by the RC is the same of the optimum:

$$\sqrt{\overline{n_F^2}} = \sqrt{\overline{n_0^2}}$$

As a result, we have a S/N reduction of the factor $\frac{2}{e}$:

$$\eta_F = \frac{2}{e}\eta_o \approx 0{,}736 \cdot \eta_o$$

The performance of the filter system with RC approximation of matched filter is about 27% worse than the absolute optimum. The main causes of this reduction are related to the readout time t_r. The finite readout time $t_r = T_{nc}$ of the RC approximation of the matched filter, while the ideal optimum has infinite readout time.

We can compute a more objective assessment of the RC approximation by taking as reference the optimum filter with the constraint of finite readout time $t_r = T_{nc}$.

The optimum filter with the constraint of finite readout time t_r gives:

$$\eta_{Ro}(t_R) = \eta_o \cdot \sqrt{1 - \exp\left(-\frac{2t_R}{T_{nc}}\right)}$$

In the case with $t_r = T_{nc}$ it is

$$\eta_{Ro}(T_{nc}) = \eta_o \cdot 0{,}929$$

So, in conclusion:

$$\eta_F = \frac{2}{e}\eta_o \approx 0{,}736 \cdot \eta_o = 0{,}79 \cdot \eta_{Ro}(T_{nc})$$

The performance of the RC approximation is about 21% worse than the optimum filter with the constraint of finite readout time $t_r = T_{nc}$. This performance is remarkably lower than the optimum, but the RC is just a crude approximation: better results can be obtained with more sophisticated filter design.

12.6 Noise whitening filter: case with finite load resistance

Now we will analyze the behavior of a noise whitening filter without approximating as infinite the input load resistance. The main difference with the approximated case is the signal shape at the sensor output. In the $R_L \rightarrow \infty$ case we treated it as an infinite step, but now we are going to take in account a slowly exponential decaying pulse, whose time constant is $T_L = R_L C_L$. The equivalent circuit of a high-impedance Sensor and preamplifier configuration with finite load resistance R_L is shown in figure 12.13.
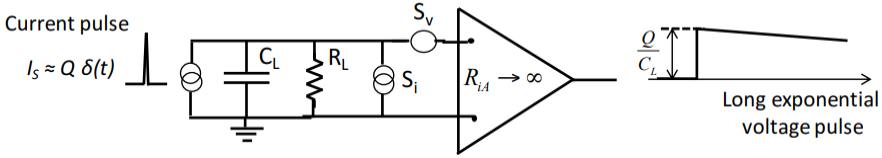
Figure 12.13 Equivalent circuit of a real high impedance Sensor and Preamp

Due to the introduction of the finite value R_L, we have an additional pole at low frequency. The noise power spectrum doesn't diverge anymore at low frequency since the load resistance sets a finite limit to the DC gain
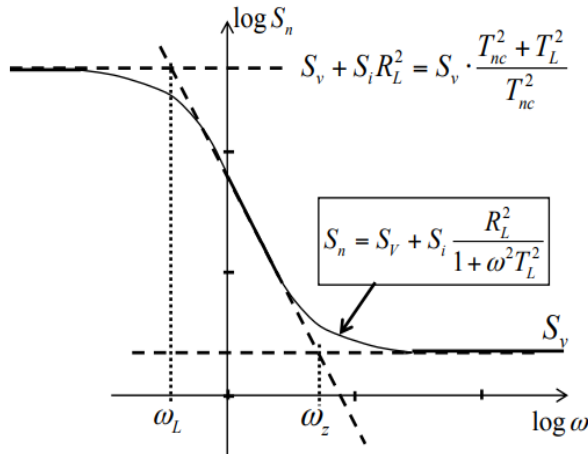

Figure 12.14 Sensor load network shaping of noise power spectral density

Let's analyze signal and noise at the preamplifier output.
The output signal is an exponential pulse with long time constant $T_L = R_L C_L$

$$y(t) = \frac{Q}{C_L} \cdot 1(t) \cdot \exp\left(-\frac{t}{T_L}\right)$$

The voltage noise spectrum S_n has two components and it is NOT white:

$$S_n(\omega) = S_v + S_i \frac{R_L^2}{1 + \omega^2 R_L^2 C_L^2} = S_v + S_i \frac{R_L^2}{1 + \omega^2 T_L^2}$$

$$S_n = S_v \left(1 + \frac{1}{T_{nc}^2} \frac{T_L^2}{1 + \omega^2 T_L^2}\right) = S_v \cdot \frac{T_{nc}^2 + T_L^2}{T_{nc}^2} \cdot \frac{1 + \omega^2 T_{nc}^2 \frac{T_L^2}{T_{nc}^2 + T_L^2}}{1 + \omega^2 T_L^2}$$

147

In which we can identify the pole frequency:

$$\omega_L^2 = 1/T_L^2$$

and the zero frequency:

$$\omega_z^2 = \frac{1}{T_z^2} = \frac{1}{T_{nc}^2}(1 + \frac{T_{nc}^2}{T_L^2})$$

With high load resistance R_L (namely with $R_L/R_{nc} = T_L/T_{nc} > 10$) the zero is about the same as with infinite load resistance:

$$\omega_z^2 \approx \frac{1}{T_{nc}^2} = \omega_{nc}^2$$

Now we can design a new whitening filter that takes into account the real shape of the noise. This filter must cancel out the pole (with a zero at the same angular frequency ω_L), and the zero (with a pole at ω_z). The Bode diagram of this whitening filter is that of a semi high-pass filter, with finite DC attenuation.

$$H_{nw}(\omega) = \frac{T_z}{T_L}\frac{1+j\omega T_L}{1+j\omega T_z} \approx \frac{T_{nc}}{T_L}\frac{1+j\omega T_L}{1+j\omega T_{nc}}$$
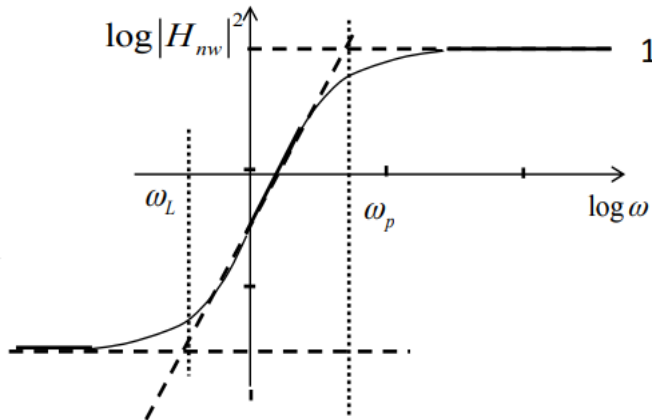
Figure 12.15 Transfer function and Bode diagram of a real case whitening filter

We can then study this implementation of a modified high-pass filter
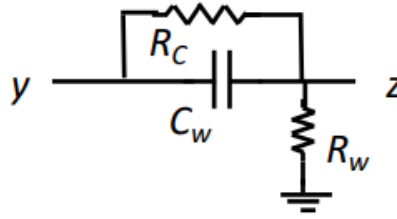
Figure 12.16 Schematic of the quasi high-pass filter

Whose transfer function is:

$$H_{nw}(s) = \frac{T_z}{T_L} \frac{1 + sT_L}{1 + sT_z} \qquad \text{with} \qquad \frac{R_C R_w}{R_C + R_w} C_w = T_z \approx T_{nc}$$

As a conclusion, the effect of this whitening filter is to turn the noise power spectrum white and to shape the signal into a short exponential pulse with time-constant T_{nc}. Figure 12.17 shows the overall pulse response of the system stage by stage.
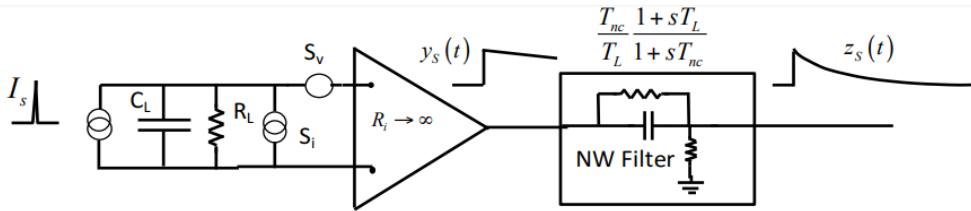
Input (current)	Preamp Output (voltage)	NW Filter Output (voltage)
δ − pulse	Long exponential pulse	Short exponential pulse
$I_s(t) = Q \cdot \delta(t)$	$y_s(t) = \dfrac{Q}{C_L} \cdot 1(t) \cdot \exp\left(-\dfrac{t}{T_L}\right)$	$z_s(t) = \dfrac{Q}{C_L} \cdot 1(t) \cdot \exp\left(-\dfrac{t}{T_{nc}}\right)$
$I_s(s) = Q$	$Y_s(s) = \dfrac{QT_L}{C_L} \dfrac{1}{1 + sT_L}$	$Z_s(s) = \dfrac{QT_{nc}}{C_L} \dfrac{1}{1 + sT_{nc}}$

Figure 12.17 Summary: high impedance sensor and preamp with whitening filter.

CHAPTER 12

Filtering: 1/f Noise and High-Pass Filter 1

Understanding 1/f noise is crucial in a wide variety of applications dealing with low amplitude and low frequency signal (for example in bioelectronics). This chapter aims at providing an insight on 1/f noise and some techniques to reduce its contribution to the total output noise are presented.

13.1 Introduction

1/f Noise was firstly reported in 1925 as *flicker noise* in electronic vacuum tubes, and it was later observed in all electronic devices. It is characterized by random fluctuations with power spectral density $S(f) \propto \frac{1}{|f|}$ and with different intensity in different devices. It has also been observed in many cases outside electronics: for example, in biology (cell membrane potential), in oceanography (ocean current flux), in astronomy (gravitational-wave), etc. Random fluctuations that characterize 1/f noise arise from physical processes that generate random superposition of elementary pulses with random pulse duration ranging from very short to very long.
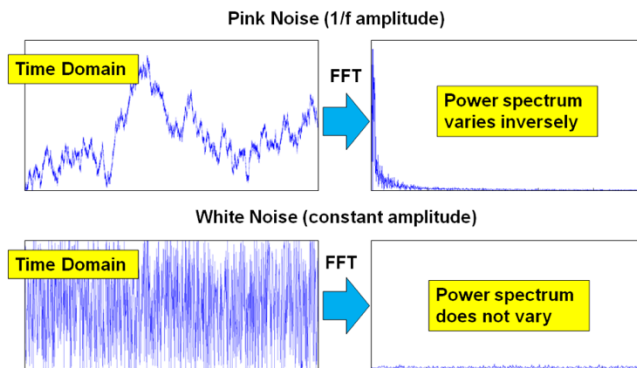
Figure 13.1 difference between 1/f and white noise in time and frequency domain.

The basic distinction between 1/f and white noise, for our purpose, is *the time span of interdependence* between the samples; for the 1/f noise, the samples are *strongly correlated even* at *long* time distance, while white noise samples are *uncorrelated even* at *short* time distance.

13.2 1/f Noise

The real observed power density at low frequency is often not exactly

$$S(f) \propto \frac{1}{|f|}$$

but rather

$$S(f) \propto \frac{1}{|f|^k}$$

with k close to unity 0.8<k<1.2.

Nevertheless, 1/f spectrum is a good approximation and we will use it in this text. 1/f noise is typically specified in relative terms referred to the white noise S_b of the same system; in particular, the *corner frequency* f_c at which $S_f = S_b$ is typically provided, and since the spectral density is

$$S_f(f) = \frac{P}{f}$$

we find out that

$$P = S_B \cdot f_c$$

As we can see from this formula, increasing the frequency f_c will determine a stronger role of the 1/f noise in the total noise computation, and for a given S_B the higher the corner frequency, the higher the intensity P.
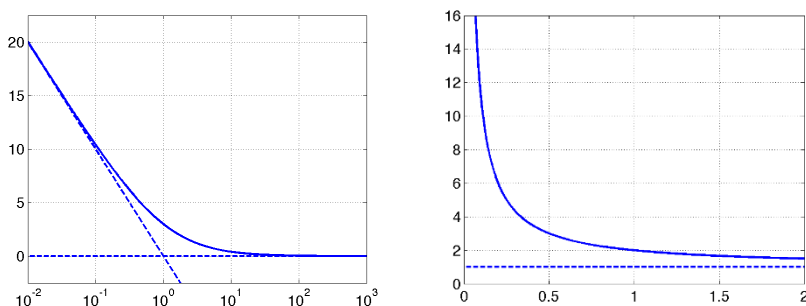
Figure 13.2 1/f and white noise spectrum: bode diagram on the left and linear diagram on the right.

Typical values for low-noise voltage amplifiers are:

$$S_B \approx 10^{-18} V^2 / Hz$$

$$10Hz < f_c < 10kHz$$

That is

$$3.2\, nV < \sqrt{P} < 100nV$$

The ideal 1/f noise spectrum runs from f=0 to $f \to \infty$ and so ideal 1/f noise has a divergent power $\overline{n_f^2} \to \infty$ both considering the low and high frequencies.

$$\overline{n_f^2} = \int_0^\infty \frac{P}{f} df \to \infty$$

However, we will see that real 1/f noise spectrum has span limited at both ends and is not divergent due to physical limitations.

If there is *wide spacing* between these high-frequency and low-frequency limitations, the formula of the noise power can be approximated by sharp cutoff at low frequency f_i and high frequency f_s and the noise power can be evaluated as

$$\overline{n_f^2} \approx \int_{f_i}^{f_s} \frac{P}{f} df = P ln\left(\frac{f_s}{f_i}\right) = S_b f_c ln\left(\frac{f_s}{f_i}\right)$$

This is correct only if $f_s \gg f_i$, otherwise the approximation will lead to an incorrect value of the noise power. (the values of f_i and f_s are the ones given by the filters).

From the above equation we can derive valuable information such as:

- $\overline{n_f^2}$ depends on the ratio $\frac{f_s}{f_i}$ and not on the absolute value f_i and f_s.
- $\overline{n_f^2}$ is divergent for $f_s \to \infty$ (like white noise). To reduce the noise power it is necessary to have a limit at high frequency. Such a limit always exists in real cases.
- $\overline{n_f^2}$ is divergent for $f_i \to 0$. To reduce the noise power, it is necessary to have a limit at low frequency. We will see that such a limit always exists in real cases.

These considerations can also be inferred from both the linear and the bode diagram of Figure 13.2 (remembering that for the bode plot f=0 is at -∞)

Another important consideration that we have to keep in mind is that the noise power $\overline{n_f^2}$ is *slowly* divergent for $f_i \to 0$ or $f_s \to \infty$ because of the logarithmic dependence. This will be a useful information to remember trying to optimize the overall noise of a circuit (considering both white and 1/f noise). The two noise

contributions, white and 1/f, have different dependency on the frequency limitations imposed by the filters.

Let's consider an example to better evaluate this dependence for the 1/f. We can compare two different situations starting from 1/f noise with $\sqrt{P} = \sqrt{S_B f_c} = 100nV$

    I.   Filtered with $f_i = 1KHz$ and $f_s = 10KHz$ ($\frac{f_s}{f_i}$=10)

$$\sqrt{\overline{n_{f,I}^2}} = \sqrt{2.3}\sqrt{S_B f_c} = 151nV$$

    II.   Filtered with $f_i = 1Hz$ and $f_s = 10MHz$ ($\frac{f_s}{f_i} = 10^7$ i.e. $x10^6$higher)

$$\sqrt{\overline{n_{f,II}^2}} = \sqrt{7*2.3}\sqrt{S_B f_c} = 401nV \text{ (just x2.7 higher)}$$

This example denotes the features of a logarithmic dependence and it shows that it is not necessary to know very precisely f_i and f_s to make a good assessment of $\overline{n_f^2}$. For example, considering a high frequency pole f_s set by a low pass filter RC we have:

    I.   Filtered with $f_i = 1KHz$ and $f_s = 10KHz$ ($\frac{f_s}{f_i}$=10)

$$\sqrt{\overline{n_{f,I}^2}} = \sqrt{2.3}\sqrt{S_B f_c} = 151nV$$

    II.   With bandlimit f_s corrected to $f_{sn} = \frac{\pi}{2} f_s = 15.7KHz$ (50% higher)

$$\sqrt{\overline{n_{f,I}^2}} = \sqrt{2.3}\sqrt{S_B f_c} = 166nV \text{ (just 10\% higher)}$$

13.3 1/f Noise Filtering

As discussed earlier, 1/f noise power depends on the ratio between the upper band limit and the lower band limit of the noise spectrum, that is $\frac{f_s}{f_i}$, and not on the individual values f_s and f_i. To better understand filtering 1/f noise we can change the integration variable from f to ln(f)

$$\overline{n_f^2} = S_b f_c \int_0^\infty |W(f)|^2 \frac{1}{f} df = S_b f_c \int_{-\infty}^\infty |W(lnf)|^2 \, d(lnf)$$

From the formula we can see that the filtered power $\overline{n_f^2}$ is proportional to the *area* of $|W|^2$ plotted in logarithmic frequency scale *(LIN-LOG plot, not a bode plot!)*, whereas the filtered power $\overline{n_B^2}$ (white noise) is proportional to the area of $|W|^2$ plotted in linear frequency scale (LIN-LIN).

In both cases, the noise power depends on the frequency span covered by $|W|^2$. However, the frequency span is measured differently:

- For white noise, by the difference of the bounds $(f_s - f_i)$
- For 1/f noise, by the *logarithmic* difference, i.e. by the ratio of the bounds $(\frac{f_s}{f_i})$

The noise effect of a filter for white noise is well visualized in the linear-linear diagram of the weighting function $|W(f)|^2$ while the noise effect of a filter for 1/f noise is well visualized in the linear-log diagram of the weighting function $|W(lnf)|^2$.

Let's clarify this concept with a couple of examples:

Example 1:
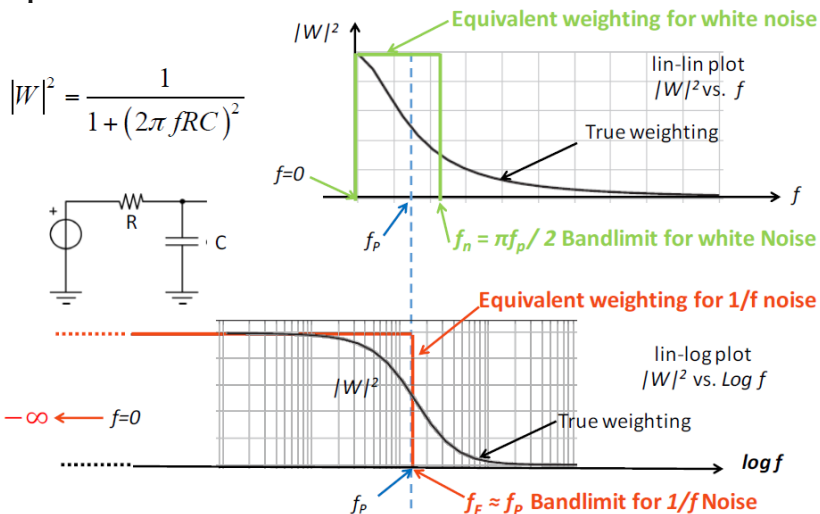
Figure 13.3 Noise bandlimit of an RC integrator

In this example we can see that:

- white noise spans from f=0 to f_n and it gives a finite noise power
- 1/f noise spans from f=0 to f_p and it gives a divergent noise power. The area of the equivalent weighting is infinite because in the lin-log plot the f=0 is at -∞. That's why we also need to filter at low frequency.
- the upper band limit for the 1/f noise is approximated, nevertheless, the approximation gives a negligible error as shown before.

Example 2:
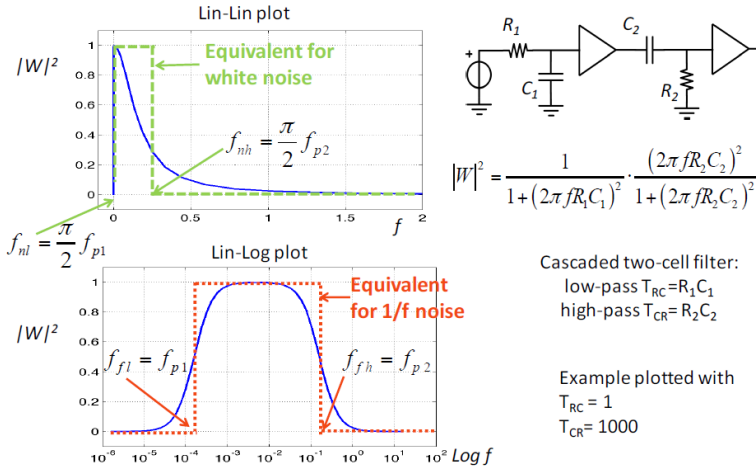
Figure 13.4 Noise bandlimits: RC lowpass plus CR high-pass

In this example we have both an integrator (RC) and a differentiator (CR). We can see that:

- white noise spans from f_{nL} to f_{nh} and it gives a finite noise power. The frequency span is $f_{nh} - f_{nL}$. In most of the cases $f_{nh} \gg f_{nL}$ and the span is well approximated by f_{nH}
- 1/f noise spans from f_{fL} to f_{fh} and it gives a finite noise power.

In both the cases the areas of the green (white noise) rectangle and the red (1/ noise) rectangle give us the value of the total noise power in our system. This graphical representation will be extremely useful to compare different filter from the intuitive point of view.

13.4 Intrinsic High-Pass Filtering by Correlated Double Sampling (CDS)

In all real cases, even with DC coupled electronics, the weighting function is inherently not extended down to zero frequency because as intrinsic high-pass filtering is always present. This intrinsic high-pass filtering action arises because the system operation is started at *some time before* the acquisition of the measure and because the operation is *started from zero value*.

Let's clarify this point with an example: if we want to measure the amplitude of the output signal of a DC amplifier, we must perform a zero-setting; the baseline voltage is preliminary adjusted to zero, or it is measured, recorded and then subtracted from the measured signal. This operation may be done a long time before the signal

measurements or repeated before each measurement. This is exactly the same procedure we perform when we turn on our weight scale, or when we make its tare. We will see now how zero-setting produces a high-pass filtering action, in fact the baseline sample is subtracted from the signal sample (both acquired with instant (delta) sampling). In the time domain, the weighting is

$$w_B(t) = \delta(t) - \delta(t + T)$$
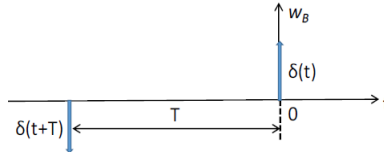
Figure 13.5 weighting function in time

This is called Correlated Double Sampling (CDS).
Applying the Fourier Transform, we obtain the weighting for the frequency domain

$$W_B(\omega) = F[w_B] = 1 - e^{i\omega t} = 1 - \cos(\omega T) - i \sin(\omega T)$$

For the noise computation we need $|W(\omega)_B|^2$.

$$|W(\omega)_B|^2 = [1 - \cos(\omega T)]^2 + [\sin(\omega T)]^2 = 2[1 - \cos(\omega T)] = 4[\sin(\frac{\omega T}{2})]^2$$

At $\omega T \ll 1$ a low frequency cutoff is produced

$$|W(\omega)_B|^2 \approx (\omega T)^2$$

(for x<<1 it is sinx ≈ x)

In figure 13.6 we can see a comparison of the action of a baseline subtraction to the action of a high-pass CR filter (differentiator) in a bode diagram.
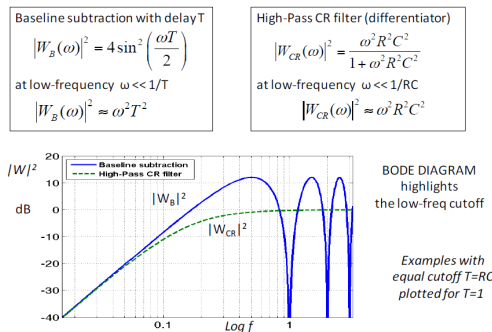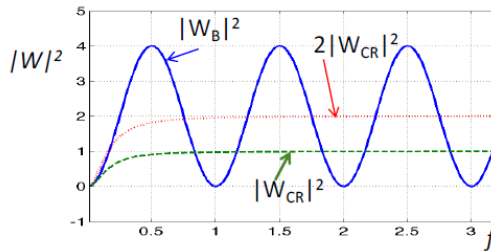

Figure 13.6 Bode diagram of $|W|^2$ for CDS and CR filter

157

We have just seen that subtracting the baseline of a signal corresponds to performing a high-pass filtering action.

Let's now analyse and compare in detail the action of a CDS and of a CR with respect to the white noise and 1/f noise.

White noise analysis:

Figure 13.7 LIN-LIN plot of $|W|^2$. Remember that the white noise power is $\propto$ area of $|W|^2$. NB: the 3 filters shown have equal cutoff T=RC=1

As we have seen, from the graphical point of view, it is easier to study the white noise in a lin-lin plot. From the figure 13.7 we can infer that the CDS $|W_B|^2$ oscillates around 2; it's area is exactly the same as for a constant $|W_B|^2 = 2$; wheareas the CR differentiator has a cutoff at low frequency $f < f_i = \frac{1}{4RC}$ while at higher frequency $|W_{CR}|^2 \approx 1$.

Therefore, the white noise output power of the CDS is *double* of the unfiltered noise, and approximately double of the filtered output of the CR (actually even more than double). This can be intuitively understood also in the time domain since the white noise is acquired twice, in the baseline and in the signal sampling, and the two noise samples are uncorrelated, hence their power sum up quadratically.

$$\overline{n_B^2} = S_B \int_0^{f_s} |W(f)|^2 \, d$$

With Baseline sampling & subtraction	With CR high-pass filter
$\overline{n_B^2} = S_B \int_0^{f_s} 2 \cdot [1 - \cos \omega T] \, df$ that is $\overline{n_B^2} = 2\, S_B f_S$	$\overline{n_B^2} = S_B (f_S - f_i)$ and since $f_s \gg f_i$ $\overline{n_B^2} \approx S_B f_S$

Figure 13.8 comparison between CDS and CR. Note the doubling factor for the CDS white noise
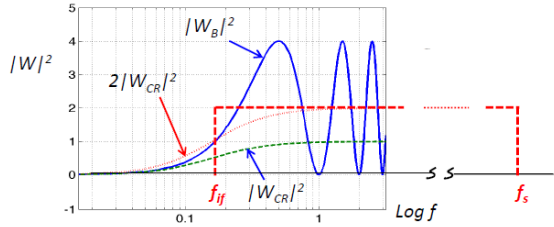
1/f noise analysis:

Figure 13.9 LIN-LOG plot. Examples with equal cutoff T=RC plotted for T=1

We can now perform the same analysis for the 1/f noise. We saw that, in this case, from the graphical point of view it is clearer to use a lin-log plot. Also in this case we limit 1/f noise power $\overline{n_f^2}$ with a low pass f_s, but with $f_s \gg \frac{1}{T}$

$$\overline{n_f^2} = \int_0^{f_s} |W(f)|^2 \frac{S_B f_c}{f} df = S_B f_c \int_0^{f_s} |W(f)|^2 \, d(lnf)$$

At low frequency $f \ll \frac{1}{T}$ the $|W_B|^2$ and $|W_{CR}|^2$ have the same cut off (T=RC). At higher frequency $|W_{CR}|^2 \approx 1$ whereas the $|W_B|^2$ oscillates around a mean value 2, so that:

$$\int_0^{f_s} |W_B(f)|^2 d(\ln f) \approx 2 \int_0^{f_s} |W_{CR}(f)|^2 d(\ln f)$$

Therefore

$$\overline{n_{f,B}^2} \approx 2\overline{n_{f,CR}^2}$$

Also 1/f noise power output of CDS is approximately double with respect to a CR high-pass with equal cut off.

13.4.1 Time domain analysis of CDS and CR for white noise performance

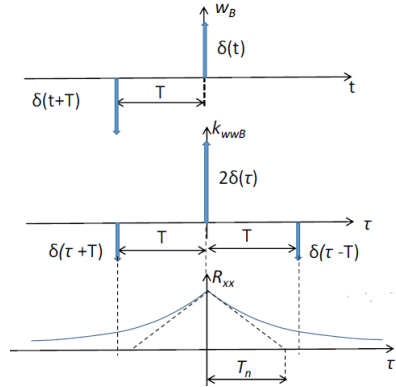We can understand how thw white noise is filtered by a CDS analysing it also in the time-domain.
White noise power at the output of the filter can be computed as

$$\overline{n_B^2} = S_B \int_{-\infty}^{+\infty} R_{xx}(\tau) \, k_{wwB}(\tau) d\tau$$

159

$R_{xx}(\tau)$ is the autocorrelation of band-limited white noise and $k_{wwB}(\tau)$ is the autocorrelation of CDS weighting function.

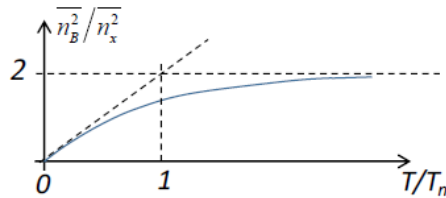$$R_{xx}(\tau) = \overline{n_x^2} e^{\frac{-|\tau|}{T_n}}$$

T_n is the noise autocorrelation time.

Figure 13.10 CDS weighting function, autocorrelation of CDS weighting function, autocorrelation of band-limited white noise

Substituting in the above equation, we find:

$$\overline{n_B^2} = S_B \int_{-\infty}^{+\infty} R_{xx}(\tau)\, k_{wwB}(\tau) d\tau = 2\overline{n_x^2} - R_{xx}(T) - R_{xx}(-T) =$$

$$= 2\overline{n_x^2} \cdot \left(1 - e^{\frac{-T}{T_n}}\right)$$

- Noise with very short correlation time is doubled:
  If $T_n \ll T$ we have $\overline{n_B^2} \approx 2\overline{n_x^2}$

- Noise with long correlation time is strongly attenuated:
  If $T_n \gg T$ we have $\overline{n_B^2} \approx \overline{n_x^2} \cdot 2\frac{T}{T_n} \ll \overline{n_x^2}$

Now we can do the same analysis for a CR filter. Let's start with writing and plotting the weighting function of a CR filter.

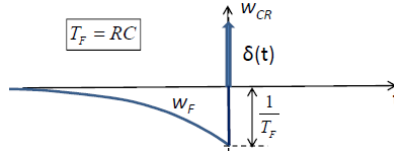$$w_{CR}(t) = \delta(t) - w_F(t) = \delta(t) - \frac{1}{T_f}e^{\frac{-|t|}{T_f}}$$

Figure 13.11 weighting function w_{CR}

Now we calculate the autocorrelation of the CR weighting function:

$$k_{wwCR}(\tau) = \delta(t) - \frac{1}{2}w_F(|\tau|) = \delta(t) - \frac{1}{2T_f}e^{\frac{-|\tau|}{T_f}}$$
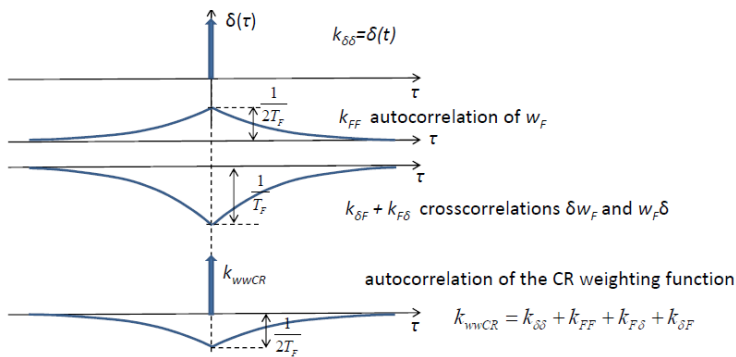

Figure 13.12 steps for the calculation of the autocorrelation of w_{CR}

We can calculate the white noise output power, using the same formula we used before

$$\overline{n_B^2} = S_B \int_{-\infty}^{+\infty} R_{xx}(\tau)\,k_{wwCR}(\tau)d\tau = R_{xx}(0) - 2\int_0^{+\infty} R_{xx}(\tau)\frac{1}{2}w_F(|\tau|)d\tau$$

$$= \overline{n_x^2} - \overline{n_x^2}\cdot\int_0^{\infty}\frac{1}{T_F}\exp\left[-\tau(\frac{1}{T_n}+\frac{1}{T_F})\right]d\tau$$

$$\overline{n_B^2} = \overline{n_x^2}\left(1 - \frac{T_n}{T_n + T_F}\right) = \overline{n_x^2}\frac{T_F}{T_n + T_F}$$

With CR we can state that:

- Noise with very short correlation time is practically passed as it is, not doubled as for CDS:

  If $T_n \ll T_F$ we have $\overline{n_B^2} \approx \overline{n_x^2}$

- Noise with long correlation time is strongly attenuated at half level of CDS:

  If $T_n \gg T_F$ we have $\overline{n_B^2} \approx \overline{n_x^2} \cdot \frac{T_F}{T_n} \ll \overline{n_x^2}$

As a conclusion, zero-setting by correlated double sampling (CDS) produces a high-pass filtering action that limits the power of 1/f noise.

The interval T between zero setting and measure in most real cases is quite long (from a few seconds to several minutes) so the high-pass band-limit $f_{i,f}$ is quite low. This is a main drawback because the filtering is not very effective since 1/f noise is limited just to a moderately low level, which may be higher than that of white noise.

Comparing CDS and CR shows a further drawback: CDS output noise power is approximately double than CR one even with equal band-limit $f_{i,f}$. This occurs because in the baseline sampling all the frequency components are acquired, but in the subtraction only those with $f \ll \frac{1}{T}$ concur to reduce 1/f noise.

13.5 Correlated Double Sampling with Filtered Baseline (CDS-FB)

As we have seen, baseline sampling is intended to acquire the contributions of the low frequency components that we want to subtract from the measurement. However instant sampling(delta-like) acquires all the frequency components at low and high frequency and by subtracting them we double the noise passed above the CDS cutoff. An intuitive remedy is to modify the baseline sampling in such a way to *acquire only the low-frequency components*.

We can perform this task sampling the baseline with a low-pass weighting function $w_F(t)$ with band-limit f_{Fn} which includes only the frequencies to be subtracted. An example can be acquiring the baseline by means of a gated integrator with narrower filtering band $f_{fn} \ll f_s$, where f_s is the upper band-limit of the noise.
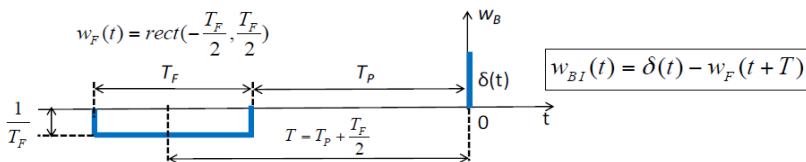

Figure 13.13 weighting function with filtered baseline.

NB: we still consider cases with long interval $T_P \gg T_F$, from zero-setting to measurement.

Let's compute the weighting function in the frequency domain and analyze its behavior at low and high frequency as we made for the CDS:

$$W_{BI}(\omega) = F[w_{BI}(t)] = F[\delta(t) - w_F(t+T)] = 1 - e^{i\omega T} W_F(\omega)$$

Since $W_F(\omega) = sinc(\frac{\omega T_F}{2})$ is real at any ω, we have

$$W_{BI}(\omega) = 1 - W_F(\omega)\cos(\omega T) - iW_F(\omega)\sin(\omega T)$$

$$|W_{BI}(\omega)|^2 = 1 + W_F{}^2(\omega) - 2W_F(\omega)\cos(\omega T)$$

- At low frequency $f \ll 1/T$ it is $W_F(f) \approx 1$ and W_{BI} has a high-pass cut off equivalent to a CR differentiator with RC=T

$$|W_{BI}(\omega)|^2 \approx (\omega T)^2 = \omega^2 (T_p + \frac{T_F}{2})^2$$
$$f_{if} \approx \frac{1}{2\pi T} = \frac{1}{2\pi(T_P + \frac{T_F}{2})} \quad \text{cut off frequency}$$

- At high frequency above the GI low-pass cut off $f_n \gg 1/2T_F$ it is $|W_F(f)| \approx 0$ so that $|W_{BI}(f)|^2 \approx 1$

- In the intermediate range $\frac{1}{T} \ll f \ll 1/2T_F$ it is roughly $W_F \approx 1$ so that roughly it is $|W_{BI}(f)|^2 \approx 2(1 - cos2\pi fT)$. In this range the average value is about $|W_{BI}(f)|^2 \approx 2$, hence we can denote it as double-noise range
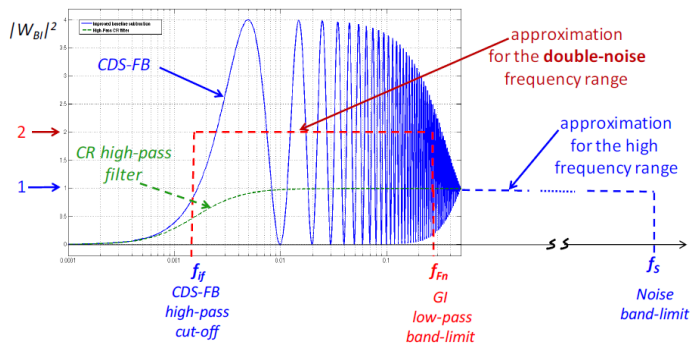
Figure 13.14 example of CDS-FB with $T_P = 101$ and $T_F = 2$. For comparison a CR filter with equal cutoff $RC=T=T_P + T_F$ is reported.

$$\overline{n^2} = \int_0^{f_s} S(f)\,|W_{BI}(\omega)|^2 df = \int_0^{f_s} S(f)\,[1 + W_F{}^2 - 2W_F\cos(2\pi fT)]df$$

By approximating W_{BI} as outlined, the 1/f noise power and white noise power can be approximately evaluated as:

$$\overline{n_{f,BI}^2} \approx S_B f_C \ln\left(\frac{f_s}{f_{if}}\right) + S_B f_C \ln\left(\frac{f_{Fn}}{f_{if}}\right)$$

$$\overline{n_{B,BI}^2} \approx S_B(f_s - f_i) + S_B(f_{Fn} - f_i) \approx S_B f_s + S_B f_{Fn}$$

In CDS-FB the noise-doubling effect is strongly reduced with respect to the simple CDS as it occurs only in the range from the low-frequency cutoff to the GI filtering band-limit.

In cases where the GI band-limit is much smaller than the noise band-limit ($f_s \gg f_{Fn}$) the effect of noise doubling is practically negligible

$$\overline{n_{f,BI}^2} \approx S_B f_C \ln\left(\tfrac{f_s}{f_{if}}\right) \qquad \text{and} \qquad \overline{n_{B,BI}^2} \approx S_B f_s$$

13.6 Correlated Double Filtering (CDF)

Both in the CDS and in the CDS-FB we acquired the signal using a delta in order to compare the filters with the simple sample&hold. We understood in the previous chapters that sample&hold it is not efficient for the noise reduction. So, in this paragraph we will study the use a gated integrator for the signal filtering in a CDS configuration. The gated integrator is quite efficient for the white noise component, but not for the 1/f component. An improvement is obtained by subtracting from the GI acquisition of the pulse another GI acquisition over an equal time interval where the pulse is not present (before it or after it). This approach has the same conceptual foundation as CDS, but here the two samples are both filtered by the GI: it is therefore called "Correlated Double Filtering" (CDF).

This approach can be extended to cases where any constant-parameter low-pass filter LPF is employed for filtering the white noise component and 1/f component is also present. In such cases, the measure can be obtained as a difference of two samples of the LPF output: a sample taken at the pulse peak and a sample taken where the pulse is not present.
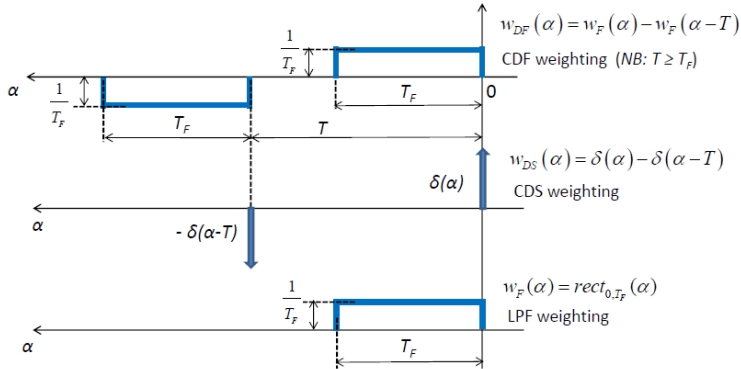
Figure 13.15 weighting function with filtered baseline and filtered signal. The LPF is a gated integrator only for graphical reasons.

The weighting of the CDF is equal to the convolution of CDS' weighting with LPF weighting

$$w_{DF}(\alpha) = w_{DS}(\alpha) * w_F(\alpha)$$

This is translated in the frequency domain to

$$W_{DF}(\omega) = W_{DS}(\omega) \cdot W_F(\omega)$$

For the noise computation

$$|W_{DF}(\omega)|^2 = |W_{DS}(\omega)|^2 \cdot |W_F(\omega)|^2$$

And since

$$|W_{DS}(\omega)|^2 = 2(1 - \cos \omega T) = 4\sin^2(\frac{\omega T}{2})$$

We have

$$|W_{DF}(\omega)|^2 = 2(1 - \cos \omega T)|W_F(\omega)|^2 = 4\sin^2(\frac{\omega T}{2})|W_F(\omega)|^2$$

The main features of CDF reflect the fact that it is a combination of CDS and LPF:
- The LPF cuts the noise at high frequencies with its LPF band-limit f_F
- The CDS cuts the noise at low frequencies with its HPF band-limit $f_{iD} \approx \frac{1}{2\pi T}$
- The CDS enhances the noise in the passband between the band-limits (with enhancement factor roughly 2)

We can now compute the weighting function in the frequency domain when a GI is used as LPF

$$w_F(\alpha) = rect_{0,T_F}(\alpha) = rect_{-\frac{T_F}{2},+\frac{T_F}{2}}\left(\alpha - \frac{T_F}{2}\right) \leftrightarrow W_F(\omega) = sinc(\frac{\omega T_F}{2})e^{\frac{-j\omega T_F}{2}}$$
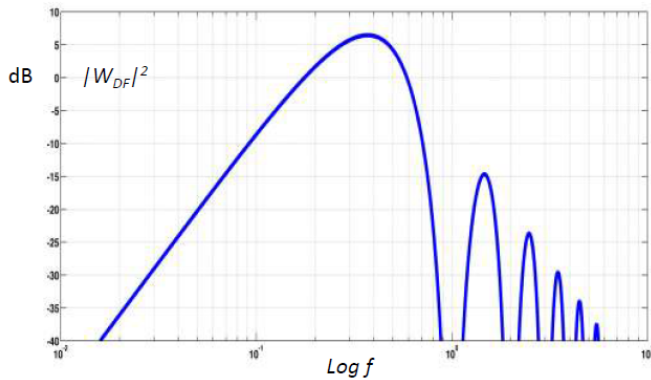
But the module does not depend on the phase factor (i.e. on the time shift)

$$|W_F(\omega)| = \left|sinc\left(\frac{\omega T_F}{2}\right)\right| = \left|sinc\left(\frac{\omega T_F}{2}\right)/\frac{\omega T_F}{2}\right|$$

Therefore

$$|W_{DF}|^2 = 2(1 - cos\omega T) \cdot |W_F|^2 = 2(1 - cos\omega T) \cdot sinc^2\left(\frac{\omega T_F}{2}\right)$$

$$= 4sinc^2\left(\frac{\omega T_F}{2}\right) \cdot sin^2\left(\frac{\omega T}{2}\right)$$

That is

$$|W_{DF}|^2 = 2(1 - cos\omega T) \cdot \frac{sin^2\left(\frac{\omega T}{2}\right)}{\left(\frac{\omega T}{2}\right)^2} = 4 \cdot \frac{sin^4\left(\frac{\omega T}{2}\right)}{\left(\frac{\omega T}{2}\right)^2}$$

Figure 13.19 Plot of $|W_{DF}|^2$ computed for the case of time shift $T=T_f$ and integration time=1

Filtering: 1/f Noise and High-Pass Filter 2

After studying 1/f noise in the previous chapter, we will here analyze some solutions to reduce its effect using both constant-parameters and switched-parameters filters. The proposed solutions will be described and discussed both in time and frequency domain.

14.1 Pulse signals in presence of 1/f noise

A typical situation in signal recovery is the need to measure the amplitude of a signal in presence of significant noise. Usually, noise consists of both 1/f and wideband contribution.

In presence of colored noise, the classic approach to evaluate the optimum filtering option requires a noise-whitening filter first, and then the design of a matched filter; unfortunately, in presence of 1/f noise the task is arduous since:

- 1/f noise sets a remarkably difficult mathematical problem
- it makes the whitening filter difficult to design: it can't be implemented with lumped circuit components, but only with distributed-parameters solutions (distributed RC delay lines, etc.)

We can try to simplify this task resorting to some properties of 1/f noise. In the previous chapter we observed that 1/f noise power mainly depends on the span of the filter band-pass measured in terms of bandlimit ratio; therefore, 1/f noise is markedly sensitive to the lower bandlimit level while it weakly depends on the shape of the filter weighting function.

On the other side, wideband noise depends on the span of the band-pass measured by the bandlimit difference, hence it is weakly sensitive to the lower bandlimit level and it markedly depends on the shape of the weighting function.

We can exploit these differences between 1/f and wideband noise to devise an alternative approach leading to quasi-optimum filtering.

We will proceed in two steps: first of all, we will design a main filter for signal and wideband noise only (that is, neglecting the presence of 1/f noise); then, we will take into account the 1/f component and we will evaluate the additional noise power that it adds to the main filter output. In the (lucky) cases where this 1/f noise power is

smaller than the wide-band noise (or at least comparable), the main filter may be considered sufficient without any additional filtering stage; otherwise, we will design an additional filter to limit 1/f noise contribution trying to avoid a significant loss of signal. It is quite obvious that filtering 1/f noise will require some sort of high-pass filtering action.

We will now consider a well-known specific case to understand some issues related to the two-step design approach introduced to deal with both 1/f and wideband noise. A pulse signal coming with both 1/f and wideband noise is generated by a high-impedance sensor; this case has been already discussed in chapter 12 considering only wideband noise. Since the two-step approach here presented suggests neglecting the 1/f noise contribution at first, we will use the same approach of chapter 12 in this phase and we will design a whitening filter. At the output, the signal will feature an exponential waveform, while the noise is wideband with low-frequency spectral density S_B. We have seen that the optimum filter in this case is not practically feasible; therefore, we will here use an approximation, that is a constant-parameter RC Integrator. We have seen that a good approximation of the matched filter is obtained with $RC = T_{nc}$, giving an upper bandwidth limitation at frequency f_S. As a result, the signal amplitude is roughly $V_S \approx A f_S$, where A is the DC amplitude of the pulse transform, while the contribution of white noise at the output of the RC Integrator is:

$$\overline{n_B^2} \approx S_B f_S$$

Let's now take into account also the 1/f noise component, which brings a significant 1/f spectral density $S_B f_C / f$ at the whitening filter output.

At high frequency, the 1/f component is limited by the upper bandlimit f_S of the matched filter.

At low frequency, the 1/f component can be limited by a lower band-limit f_i set by an additional constant-parameter filter. With $f_i \ll f_S$, the output power of the 1/f noise can be evaluated as

$$\overline{n_{fn}^2} \approx S_B f_C \ln\left(\frac{f_S}{f_i}\right)$$

However, the constant-parameter high-pass filter operates also on the signal: it attenuates the low frequency components and thus causes a loss in pulse amplitude, hence a loss in S/N. The reduced amplitude is roughly evaluated as

$$V_S \approx A(f_S - f_i) = A f_S \left(1 - f_i / f_S\right)$$

In order to limit the signal loss, f_i / f_S must be low; e.g. to keep the loss below 5% it must be

$$f_i / f_S \leq 0,05$$

that is

$$\ln\left(f_S/f_i\right) \geq 3$$

To limit 1/f noise below the white noise level it is necessary that:

$$S_B f_C \, \ln\left(f_S/f_i\right) \leq S_B f_S$$

Which means that

$$f_C \leq \frac{f_S}{\ln\left(f_S/f_i\right)}$$

and since for keeping the signal loss <5% it must be

$$\ln\left(f_S/f_i\right) \geq 3$$

we need to have

$$f_C < f_S/3$$

This means that the goal can be achieved only if the 1/f noise component is low or moderate. Note that f_C and f_S are *data* of the problem, they cannot be changed. In cases where f_C exceeds the above limit, a constant-parameter high-pass filter is NOT a suitable solution to limit the 1/f noise power.

In conclusion, constant-parameter high-pass filters can be useful to limit 1/f noise only in specific cases, i.e. when 1/f noise intensity is moderate. In all other cases, the introduction of a high-pass filter can either have a detrimental effect on the signal pulse amplitude or no significative effect on 1/f noise contribution, depending on the filter parameters.

14.2 Basic constant-parameter High-Pass Filter

In the previous paragraph we tried to use a constant parameter high-pass filter to reduce the effect of the 1/f noise. Before proceeding with the study of more complicated filters, let's take a step back and refresh the knowledge on the CR filter. First of all, the delta response of the filter is:
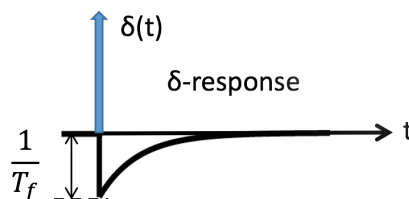
$$h(t) = \delta(t) - 1(t)\frac{1}{T_f}e^{-\frac{t}{T_f}}$$

Figure 14.1 delta response of constant-parameter high-pass filter

And its step response is

$$u(t) = 1(t)\, e^{-\frac{t}{T_f}}$$

And finally, its transfer function is

$$H(f) = \frac{j\, 2\pi f T_f}{1 + j\, 2\pi f T_f}$$

While the step response is typically intuitive for the student, the same does not hold for the delta response. In the following, we will try to give a point of view that can make the calculation of the delta easier.

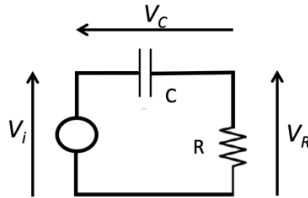Let's start from this easy and intuitive consideration:

High-Pass Filter = All-Pass - Low-Pass Filter

We can then use this consideration to move in the frequency domain

$$H(f) = 1 - \frac{1}{1 + j\, 2\pi f T_f} = \frac{j\, 2\pi f T_f}{1 + j\, 2\pi f T_f}$$

And then to calculate the delta response coming back to the time domain.
The same approach can be followed also from the circuit point of view. Let's take

a very simple network:
Where

V_i = input voltage
V_C = low-pass filtered V_i
V_R = high-pass filtered V_i

The Kirchoff's voltage law give us

$$V_i = V_C + V_R$$

And therefore

$$V_R = V_i - V_C$$

That bring us to the same results

High-pass filtered V_i = resistor voltage =
= input voltage V_i – capacitor voltage =
= input voltage V_i – Low-pass filtered V_i

Let's now apply the constant-parameter high-pass filter to white noise. With only a high-pass CR filter the white noise power $\overline{n_B^2}$ is divergent, therefore we consider here also a low-pass filter with band-limit $f_s \gg 1/RC$.

The high-pass band-limit f_i of the CR filter with weighting function $W(f)$ is defined by

$$\overline{n_B^2} = S_B \int_0^{f_s} |W(f)|^2 \, df = S_B \int_0^{f_s} \frac{\left(f/f_p\right)^2}{1+\left(f/f_p\right)^2} \, df = S_B (f_s - f_i)$$

The computation of the integral can be avoided by recalling that

$$\text{CR high pass filter} = \text{all-pass} - \text{RC low-pass filter}$$

and therefore, the high-pass band-limit f_i of the CR filter is equal to the low-pass band-limit f_h of the RC filter

$$f_{i\,CR} = f_{h\,RC} = \frac{1}{4RC}$$

We can now apply the same filter to the 1/f noise. Also in this case, with only a high-pass CR filter the noise power $\overline{n_f^2}$ is divergent, therefore we consider again also a low-pass filter with band-limit $f_s \gg 1/RC$.

The high-pass band-limit f_{if} of the CR filter is defined by

$$\overline{n_f^2} = S_B f_c \int_0^{f_s} \frac{\left(f/f_p\right)^2}{1+\left(f/f_p\right)^2} \, \frac{df}{f} = S_B f_c \int_{f_{if}}^{f_s} \frac{df}{f} = S_B f_c \ln(\frac{f_s}{f_{if}})$$

In this case the first integral is fairly easily computed, and it shows that

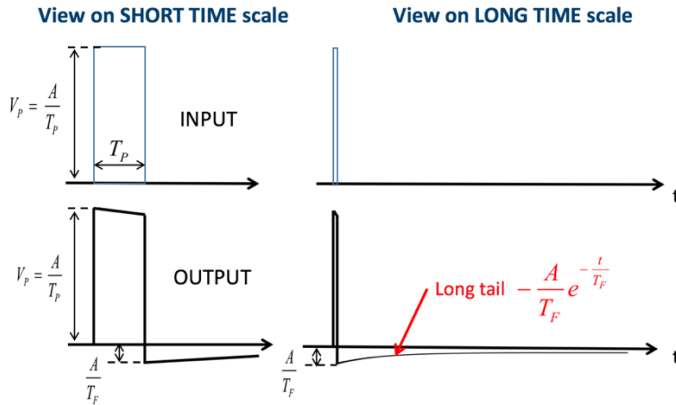$$f_{if} = \frac{f_p}{\sqrt{1+\left(f_p/f_s\right)^2}}$$

that is, for $f_s \gg f_p$

$$f_{if} \approx f_p = \frac{1}{2\pi RC}$$

14.3 Constant-Parameter High-Pass Filters in measurements of pulses in sequence
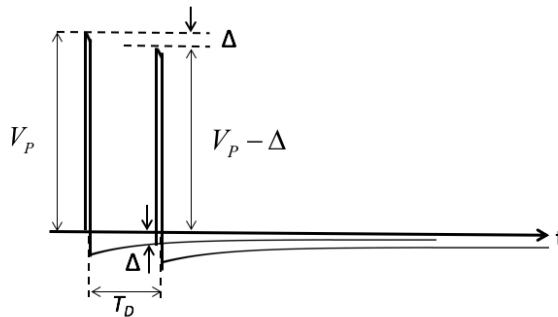
Using a constant-parameter high-pass filter, we have to consider some aspects when we play with a sequence of pulses.

Let's analyze the effect of a high-pass filter ($RC = T_F$) on a pulse rectangular signal
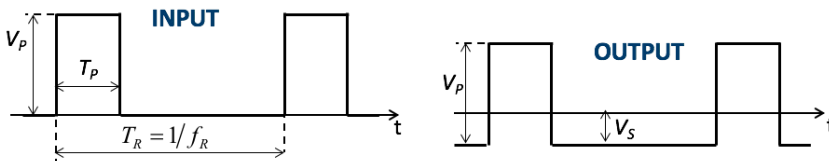
Figure 14.2 constant-parameter high-pass filter response to a rect pulse

A pulse that follows a previous one within a fairly short time interval ($T_D < 5\ T_F$) sums up to the slow tail of the first pulse. Therefore, it starts from a down-shifted baseline, so that its measured amplitude is smaller than the real one.

Figure 14.3 constant-parameter high-pass filter response to a couple of rect pulse

For *periodic* pulses with fairly short repetition period $T_R \ll T_F$, the superposition of slow pulse-tails shifts down the baseline by a V_S that makes zero the net area of the output signal

Figure 14.4 constant-parameter high-pass filter response to a periodic pulse

We can so calculate the repetition-rate-dependent baseline-shift as

$$V_S = V_P \frac{T_P}{T_R} = A\, f_R$$

We can therefore draw certain final considerations on the high-pass filtering (differentiator action) of the CR filter.

- The effect on noise is *advantageous*: by cutting off the low frequencies it markedly decreases the 1/f noise power (and mildly reduces the white noise contribution)
- The effect on the signal is *disadvantageous*:
  - it decreases the signal amplitude by cutting off the low frequencies of the signal, hence f_i must be kept low ($f_i \ll f_S$ of the pulse) in order to limit the signal loss. However, this limits also the reduction of 1/f noise
  - it generates slow tails after the pulses, which down shifts the baseline and thus it can cause an error in the amplitude measurement of a following pulse
  - With a *periodic* sequence of equal pulses, all pulses find the same baseline shift. The amplitude error is constant, systematically dependent on the repetition rate and can be calculated.
  - With *random-repetition* pulses (e.g. pulses from ionizing radiation detectors, for example) the pulses occur randomly in time. Hence the random superposition of tails produces a randomly fluctuating baseline shift. The resulting amplitude error is random: in this case the effect is equivalent to that of an additional noise source.

In conclusion, a differentiator action is desirable on noise, but NOT on the signal.

14.4 Switched-Parameter High-Pass Filter: The Baseline Restorer

In the previous paragraph we understood that the main problem in the use of a constant-parameter high-pass filter to limit the effect of the 1/f noise is that we have to find a trade-off between attenuation of the noise and effect on the signal.
To solve this trade off we introduce now a switched-parameter filter. The idea is to work in a different way on noise and signal using the possibility to switch the parameters. The circuit we will use is shown in figure 14.5
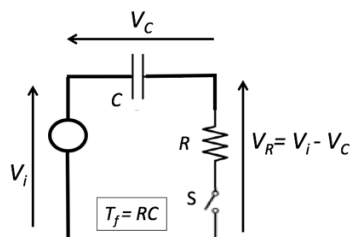

Figure 14.5 switched-parameter high-pass filter

The circuit is quite the same of the standard constant-parameter filter, but we introduced a switch to stop the differentiator effect when the signal is present.
We can compare the weighting function of both constant and switched parameter filter on the same signal
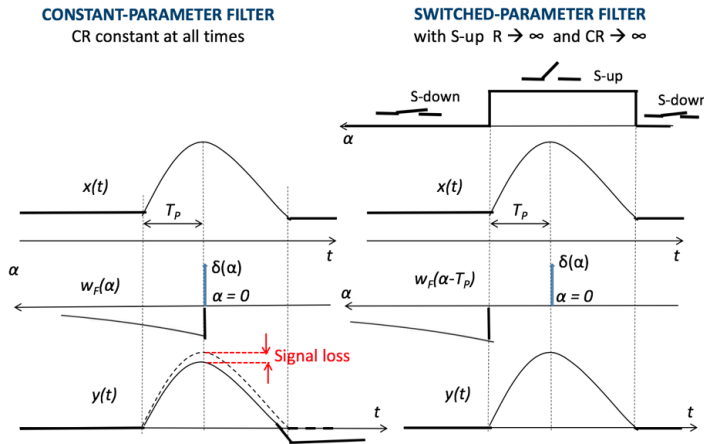

Figure 14.6 constant-parameter vs switched-parameter high-pass filtering

The weighting function of the switched-parameter filter is then composed of two different parts: a delta, corresponding to the sampling we make when the switch is open and the signal passes without modifications; a low-pass filtering action which is carried out when the switch is closed and it is used to read the value of the low frequency components of the input waveform

$$w_B(\alpha) = \delta(\alpha) - w_F(\alpha - T_P)$$

with

$$w_F(\alpha - T_P) = 1(\alpha - T_P)\frac{1}{T_f}e^{-\frac{\alpha - T_P}{T_f}}$$

In fact, as soon as S is open (at the pulse onset, i.e. at $\alpha=T_P$), charging of C stops and voltage V_C stays constant at the stored value.
The name of this filter is *Baseline Restorer* (BLR).
In the case of a sequence of pulses, we can open the switch every time the signal is present; in this case, we can easily see that our filter can be written as:

BLR weighting = (All-Pass – Low-pass Boxcar) weighting
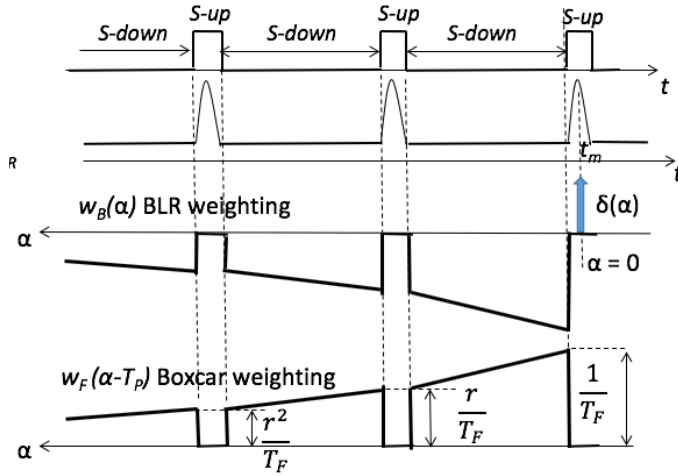
$$w_B(\alpha) = \delta(\alpha) - w_F(\alpha - T_P)$$

Figure 14.7 switched-parameter filter for sequence of pulses

BLR principle is alike filtered zero-setting, but with a basic advantage: the high-frequency band limit f_{if} (high-pass) can be very high providing a very short T_P, which can be achieved with a fast electronically-controlled switch.
Let's move in the frequency domain.

BLR weighting = All Pass – Low-pass

$$w_B(\alpha) = \delta(\alpha) - w_F(\alpha - T_P)$$

The low-pass weighting function in the frequency domain is:

$$W_F(\omega) = F[w_F(\alpha)] = R_F(\omega) + i\, I_F(\omega)$$

so, we can find the BLR weighting in the frequency domain as follows:

$$W_B(\omega) = 1 - e^{j\omega T_P} W_F(\omega) = 1 - [\cos \omega T_P - j \sin \omega T_P] \cdot [R_F + jI_F] =$$
$$= [1 - R_F \cos \omega T_P - I_F \sin \omega T_P] - j[I_F \cos \omega T_P - R_F \sin \omega T_P]$$

Now, we can analyze the BLR weighting function for noise

$$|W_B(\omega)|^2 = [1 - R_F \cos \omega T_P - I_F \sin \omega T_P]^2 + [I_F \cos \omega T_P - R_F \sin \omega T_P]^2 =$$
$$= 1 + R_F^2 + I_F^2 - 2R_F \cos \omega T_P - 2I_F \sin \omega T_P =$$
$$= 1 + |W_F|^2 - 2R_F \cos \omega T_P - 2I_F \sin \omega T_P$$

The study of this function can be very complicated and strongly depends on the number of pulses that are used. To try to better understand how it works using

relatively simple computation let's consider only cases where the interval between pulses is much longer than T_F so that

$$w_F(\alpha) = 1(\alpha) \frac{1}{T_f} e^{-\frac{\alpha}{T_f}}$$

and

$$W_F(\omega) = \frac{1}{1 + j\omega T_F}$$

therefore

$$|W_B(\omega)|^2 = 1 + \frac{1}{1 + \omega^2 T_F^2} - 2\frac{1}{1 + \omega^2 T_F^2}\cos \omega T_P + 2\omega T_F \cdot \frac{1}{1 + \omega^2 T_F^2}\sin \omega T_P$$

In the low-frequency region $\omega \ll \dfrac{1}{T_P}$ with the approximations

$$\sin \omega T_P \approx \omega T_P$$

$$\cos \omega T_P = 1 - \frac{\omega^2 T_P^2}{2}$$

we get

$$|W_B(\omega)|^2 \approx 1 + \frac{1}{1 + \omega^2 T_F^2} - \frac{2}{1 + \omega^2 T_F^2} + \frac{\omega^2 T_P^2}{1 + \omega^2 T_F^2} + 2\frac{\omega^2 T_P T_F}{1 + \omega^2 T_F^2} =$$

$$= \frac{\omega^2 (T_P + T_F)^2}{1 + \omega^2 T_F^2} = \frac{\omega^2 T_F^2}{1 + \omega^2 T_F^2}\left(1 + \frac{T_P}{T_F}\right)^2$$

and in the lower region $\omega \ll \dfrac{1}{T_F} \ll \dfrac{1}{T_P}$
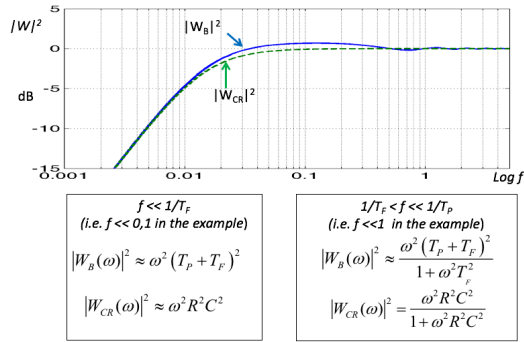
$$|W_B(\omega)|^2 \approx \omega^2 (T_P + T_F)^2$$

As a result, the BLR has a cutoff equivalent to a CR high-pass with RC= $T_P + T_F$

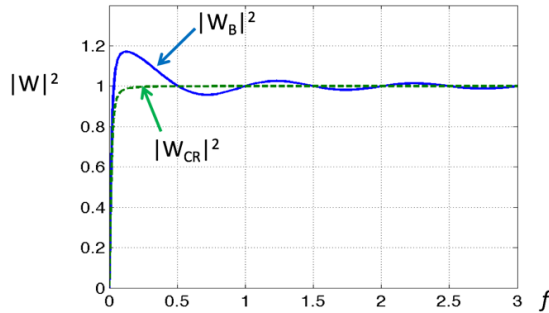It is possible to derive that the high-pass band-limit for the white noise is

$$f_{ni} \approx \frac{\pi}{2} f_p = \frac{1}{4RC} = \frac{1}{4(T_F + T_P)} \approx \frac{1}{4T_F}$$

- f_{ni} is equal to that of the equivalent CR High-pass filter
- f_{ni} is equal to bandlimit of the low-pass section in the BLR circuit

Figure 14.8 The Bode diagram highlights the low frequency cut-off, in this example: BLR with TP = 1 and TF = 10, CR filter with RC = TP + TF

Figure 14.9 Lin-Lin diagram highlights the low frequency cut-off, in this example: BLR with TP = 1 and TF = 10, CR filter with RC = TP + TF

We can look at the effect of the BLR on 1/f noise; to this aim, we plot the weighting function in a Lin-Log scale in order to evaluate the area of the $|W|^2$
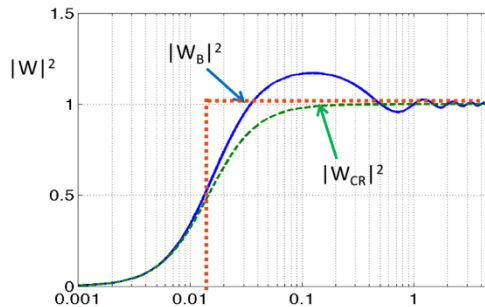

Figure 14.10 Lin-Log diagram highlights the low frequency cut-off, in this example: BLR with $T_P = 1$ and $T_F = 10$, CR filter with RC = $T_P + T_F$

It is possible to derive that the high-pass band-limit for 1/f noise is

$$f_{fi} \approx f_p = \frac{1}{2\pi RC} = \frac{1}{2\pi(T_F + T_P)} \approx \frac{1}{2\pi T_F}$$

- f_{fi} is equal to that of the equivalent CR High-pass filter
- f_{fi} is equal to bandlimit of the low-pass section in the BLR circuit

Let's go in more detail about how to fit it in the real application. The BLR filtering is ruled by:

1. T_P = **time delay** from switch opening to pulse-amplitude measurement. There is **no choice:** T_P is equal to the rise time from pulse onset to peak. In fact, T_P can't be shorter than the rise of the pulse signal and it should be as short as possible to effectively filter 1/f noise.
2. $T_F = RC$ **differentiation time constant:** it has **to be selected** to optimize the overall filtering of noise. The question is: how T_F should be selected to
   a) provide a good reduction of the 1/f noise power and
   b) avoid a significant enhancement of the white noise power

Since the BLR cut-off is set by $1/(T_P + T_F)$, a very short T_F might look advisable, but it is not: a BLR with $T_F \ll T_P$ operates like a CDS, hence it doubles the white noise and remarkably enhances also 1/f noise above the cut-off frequency.

In the following discussion about the selection of T_F, we will consider a wideband noise spectrum coming with the signal. Let's start with a *time-domain analysis of BLR filtering.* To this aim, we need to study the autocorrelation function of the BLR weighting function. In figure 14.11, a graphical method to calculate this function and the corresponding result are sketched.

Starting from the weighting function, that is

$$w_B = \delta(\alpha) - w_F(\alpha) = \delta(\alpha) - 1(\alpha - T_P) \cdot \frac{1}{T_F} \exp\left(-\frac{\alpha - T_P}{T_F}\right)$$

we can calculate the $k_{\delta\delta}$ autocorrelation of $\delta(\alpha)$, that is still a delta, the k_{FF} autocorrelation of w_F (that has been studied in the previous chapters) and finally the cross-correlation function of w_F and δ.

Considering a white noise band-limited by a single pole, we recall that its autocorrelation function is:

$$R_{xx}(\tau) = \overline{n_x^2} e^{\frac{-|\tau|}{T_n}}$$

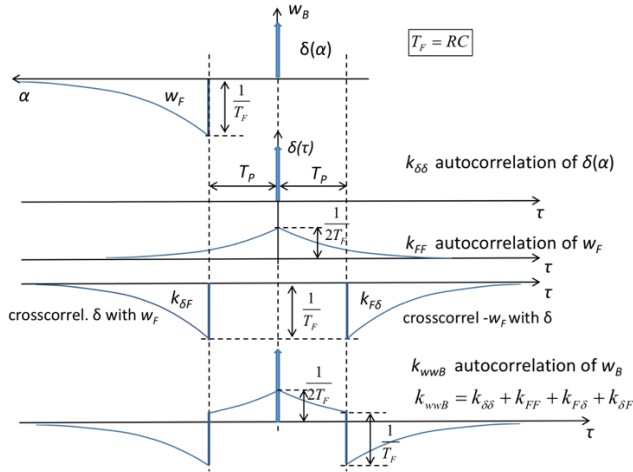Now, we need to multiply the two autocorrelation functions (noise and BLR) and compute the integral of the product.

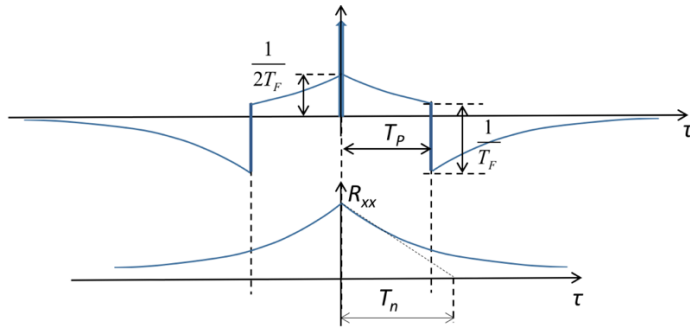**Figure 14.11** Graphical construction of the BLR autocorrelation

Figure 14.12 BLR autocorrelation and noise autocorrelation

The total noise is then

$$\overline{n_B^2} = \int_{-\infty}^{\infty} R_{xx}(\tau) k_{wwB}(\tau) d\tau = R_{xx}(0) + 2\int_0^{\infty} R_{xx}(\tau)\frac{1}{2} w_F(\tau) d\tau - 2\int_{T_P}^{\infty} R_{xx}(\tau) w_F(\tau - T_P) d\tau =$$

$$= R_{xx}(0) + \int_{-\infty}^{\infty} R_{xx}(\tau)\frac{1}{2} w_F(\tau) d\tau - 2\int_0^{\infty} R_{xx}(\beta + T_P) w_F(\beta) d\beta$$

Denoting

$$r_{xx}(\tau) = \frac{R_{xx}(\tau)}{R_{xx}(0)} = \frac{R_{xx}(\tau)}{n_x^2}$$

we have

$$\overline{n_B^2} = \overline{n_x^2}\left\{1 + \int_{-\infty}^{\infty} r_{xx}(\tau)\frac{1}{T_F} e^{-\frac{\tau}{T_F}} d\tau - 2\int_0^{\infty} r_{xx}(\beta + T_P)\frac{1}{T_F} e^{-\frac{\beta}{T_F}} d\beta\right\}$$

and finally

$$n_B^2 = n_x^2 \left[1 + \frac{T_n}{T_n + T_F} \left(1 - 2e^{-\frac{T_P}{T_n}} \right) \right]$$

With *fast differentiation*, i.e. with $T_F \ll T_n$, it is quantitatively confirmed that the BLR acts like a CDS with $T = T_P$

$$n_B^2 \approx 2n_x^2 \cdot \left(1 - e^{-\frac{T_P}{T_n}} \right)$$

With $T_F \ll T_n$ the effect of BLR on band-limited white noise depends on how long the correlation time T_n is with respect to the delay T_P

- with **short correlation time** (wide band) the noise is *doubled*:
  with $T_n < T_P/5$ it is

$$\overline{n_B^2} \approx 2\overline{n_x^2}$$

- with **moderate correlation time** (moderately wide band) the noise is enhanced:
  with $T_n \approx T_P/2$ it is

$$\overline{n_B^2} \approx 1,73\,\overline{n_x^2}$$

- only with **long correlation time** (low-frequency band) the noise is *attenuated* (*note that anyway the level is double of that given by a simple CR filter with equal cut-off, that is with $T_F = RC = T_P$*).
  with $T_n > 10T_P$ it is

$$\overline{n_B^2} \approx \overline{n_x^2} \cdot 2\frac{T_P}{T_n} < 0,2\,\overline{n_x^2}$$

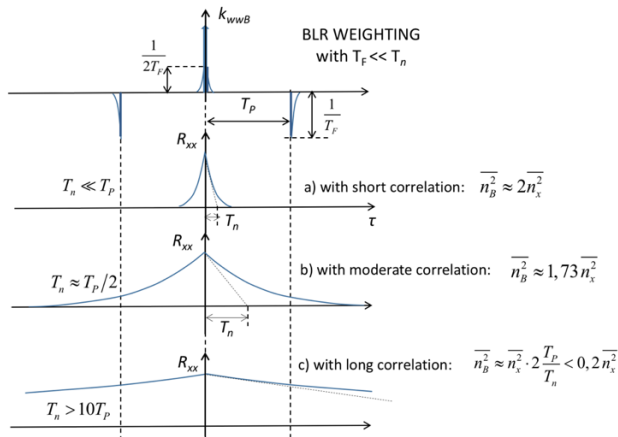A graphical summary is shown in figure 14.13.

Figure 14.13 BLR filtering with fast differentiation

Going back to the expression of noise

$$\overline{n_B^2} = \overline{n_x^2}\left[1+\frac{T_n}{T_n+T_F}\left(1-2e^{-\frac{T_P}{T_n}}\right)\right]$$

We can now consider the case where T_F is NOT negligible with respect to T_n. In this case, the effect on white noise depends also on the size of T_F compared to T_n and T_P: a long T_F can limit the white noise enhancement so let's evaluate how long T_F must be in the various cases of noise correlation.

With **short correlation time,** i.e. $T_n \approx T_P /10$, it is

$$\overline{n_B^2} \approx \overline{n_x^2}\left(1+\frac{T_n}{T_n+T_F}\right)$$

to keep $\overline{n_B^2} < 1.05\,\overline{n_x^2}$ we need $T_F > 20\ T_n \approx 2\ T_P$

With **moderate correlation time,** i.e. $T_n \approx T_P /2$, it is

$$\overline{n_B^2} \approx \overline{n_x^2}\left[1+\frac{T_n}{T_n+T_F}\left(1-\frac{2}{e^2}\right)\right] = \overline{n_x^2}\left[1+0,73\frac{T_n}{T_n+T_F}\right]$$

To keep $\overline{n_B^2} < 1.05\,\overline{n_x^2}$ in this case we need $T_F > 7T_n = 3,5T_P$

With **long correlation time,** i.e. $T_n > 10\ T_P$, it is

$$\overline{n_B^2} \approx \overline{n_x^2}\left[1-\frac{T_n}{T_n+T_F}\right] = \overline{n_x^2}\,\frac{T_F}{T_n+T_F}$$

With such a low-frequency noise, the BLR acts as a CR constant-parameter filter (with equal time constant $T_F = RC$).
The most interesting case for us is noise with moderate T_n. In fact, when the BLR works on the output of an optimum (or approximate-optimum) filter for wideband noise, the correlation time T_n and delay T_p are comparable, since they are both closely related to the band-limit of the signal pulse.
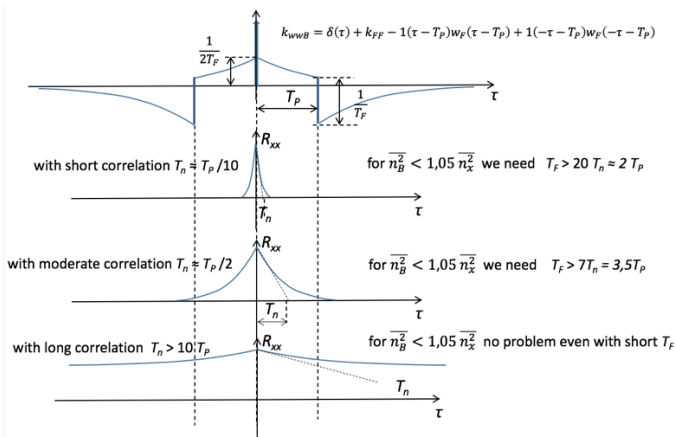
Figure 14.14 BLR filtering with slow differentiation

We conclude that avoiding any enhancement of the white noise requires a fairly slow BLR differentiation, i.e. a quite long T_F

$$T_F \geq 5\,T_P$$

This approach is suitable also for 1/f noise filtering, notwithstanding that making T_F longer than T_P shifts down the BLR cut-off frequency, hence reduces the attenuation of 1/f noise. This is counterbalanced by the fact that the enhancement of 1/f noise at frequencies above the cutoff is limited by the low-pass filtering in the baseline subtraction, whereas with short T_F it is remarkable.

14.5 BLR summary

- The BLR is a high-pass filter that acts on noise and disturbances without affecting the pulse signal
- The BLR is a switched-parameter filter: the low-pass section within the high-pass filter structure is a boxcar integrator that acquires the baseline only in the intervals free from pulses
- The BLR can thus establish a high-pass band-limit at a high value (suitable for reducing efficiently the 1/f noise output power) without causing the signal loss suffered with a constant-parameter high-pass filter having the same band-limit
- The high-pass band-limit enforced by the BLR is given (with good approximation) by the low-pass bandlimit of the low-pass section in the BLR circuit structure
- The combination of: (1) optimum filter designed for the case of pulse signal in presence of wideband noise only (i.e. without 1/f noise) and (2) BLR specifically designed (for reducing the actual 1/f noise without worsening the wide-band noise) provides in most cases a quasi-optimum filtering solution

Band-Pass Filters 1

Bandpass filters are very useful to recover narrow-band signals. The aim of this chapter is to introduce the student to the bandpass filter, discussing why and how this type of filter is used, which are its figures of merit, how it can be implemented using simple circuits and its applications. The theory of amplitude and frequency modulation will also be presented in this chapter.

15.1 Introduction

In electronics, information is often carried by signals that do not have a DC component and all their energy (and so information) is concentrated in a narrow portion of the spectrum. In these cases, bandpass filters are useful since they eliminate both the low-frequency and high-frequency components of the system that are undesired. Some types of sensors (e.g. MEMS gyroscopes) naturally encode the information using narrow-band high-frequency signals, but we will see that bandpass filters can be very useful even when working with sensors that output quasi-dc signals (e.g. load cells).

In fact, we know that many physical systems, including semiconductor devices, are subject to statistical fluctuations with power spectral density proportional to the inverse of frequency, commonly called 1/f noise or pink noise. 1/f noise is very strong at low frequency (theoretically infinite at DC) and sensors that normally have quasi-dc output are often driven in such a way that their output is shifted at high frequency without losing information, this practice is called modulation and will be explained in paragraph 15.4. After modulation, if a bandpass filter is used, the pink noise can be filtered out obtaining a higher S/N.

Narrow signals (and so bandpass filters) are also very common in RF telecommunication for two main reasons: low-frequency signals cannot be transmitted easily using electromagnetic waves since they require very long antennas (antenna length must be of the same order of magnitude as the wavelength). Also, the use of narrow-band signals allows multiple streams to be transmitted at the same time on the same medium: the disposable bandwidth can be divided in contiguous non-overlapping portions called channels. At the receiver end, a single stream can be then reconstructed using a bandpass filter that filters out all the channels except the one of interest.

Narrow-band signals concentrate their power in a small part of the spectrum, and this allows filtering out most of the noise reaching a higher S/N with respect to a broadband signal with the same power. Narrow-band signals are therefore preferred

in applications that require a low power consumption, a long transmitting range and a high reliability.

As rule of thumb, a signal can be considered narrow-band if its energy spectrum is centered around f_s and its spectrum width Δf_s is much smaller than f_s.

15.2 Characteristics

In figure 15.1 the generic modulus and phase of a bandpass-like transfer function is shown. From the modulus diagram, we can see that the transfer function has a zero in the origin and then two poles. As a result, the filter attenuates the frequencies before the first pole f_L and those after the second pole f_H: these regions are called *Stop Band*. Between the two poles, we have the range of frequencies that we do not want to attenuate: this is called *Pass Band* and it also represents the bandwidth of the filter. To make the frequencies in the pass band to pass unaltered through the filtering stage, we ideally should have a gain equal to 1 (0dB) and a phase shift equal to 0°. Actually, if we could implement an ideal band-pass filter, this would be able to completely eliminate the unwanted frequencies; however, the ideal filter would require vertical transitions at the edges of the pass band, but this is not feasible. Therefore, we will always speak of *attenuation* in the stop band.
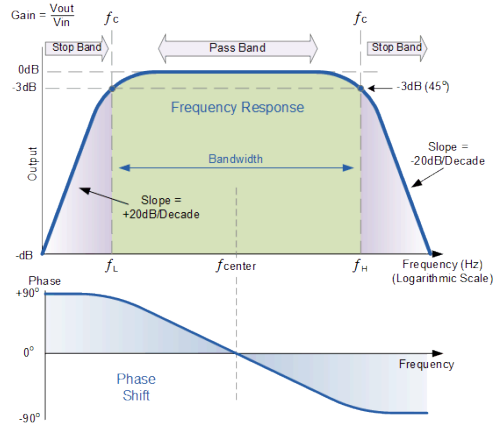
Figure 15.1: Modulus and Phase frequency Response

Here we can introduce the concept of *roll-off*. The roll-off is the steepness of the transition between stop band and pass band; it is usually measured in dB/decade.

The roll-off provides information about the selectivity of the filter we are considering: a larger roll-off means a steeper transfer function. This parameter gives us an idea on how the filter is able to cut the unwanted frequencies and so it is a very important information when we have to choose a filter for a specific application. Another important parameter of a bandpass filter is the fractional bandwidth s, which

is defined as the ratio between the filter bandwidth and the central frequency of the pass-band range:

$$s = \frac{\Delta f}{f_0}$$

In resonant filters, we will see in the next chapter that the quality factor Q of the oscillator is equal to inverse of the fractional bandwidth:

$$Q = \frac{f_0}{\Delta f}$$

15.3 Real Case Examples

Let see how narrow signals can be measured in presence of white noise and also in presence of 1/f noise. Let us consider a signal centered at f_S and with bandwidth

$$\Delta f_s = 1\ Hz$$

and a small amplitude

$$Vs \leq 100\ nV.$$

In order to make it suitable for processing circuits, the signal is amplified by a DC-coupled wide-band preamplifier with upper band-limit of 1 MHz. This preamplifier introduces also some noise. In particular, we consider the case of a white noise with spectral density $\sqrt{S_v} = 5\ \frac{nV}{\sqrt{Hz}}$ and a $\frac{1}{f}$ component with a corner frequency $f_c = 2\ kHz$.

Depending on the value of f_s, the situation is different, and the results show opposite conclusions.

For instance, if we have $f_s = 100\ kHz$, that is a frequency much higher than the noise corner frequency, we could apply a simple high-pass filter with a band limit of $10\ kHz$ in order to eliminate the 1/f component obtaining a rms noise

$$\sqrt{v_n{}^2} = \sqrt{S_b} \cdot \sqrt{f_h - f_i} \approx \sqrt{S_b} \cdot \sqrt{f_h} = 5\ \mu V$$

Which corresponds a very low Signal-to-Noise Ratio:

$$\frac{S}{N} = \frac{100\ nV}{5\ \mu V} = 0,02 \ll 1$$

In particular, a Signal-to-Noise Ratio smaller than 1 means that the signal is completely buried into noise.

This result can be observed experimentally on an oscilloscope display, as reported in figure 15.2: it is easy to notice that our signal is practically invisible on the display since it is completely covered by the noise.
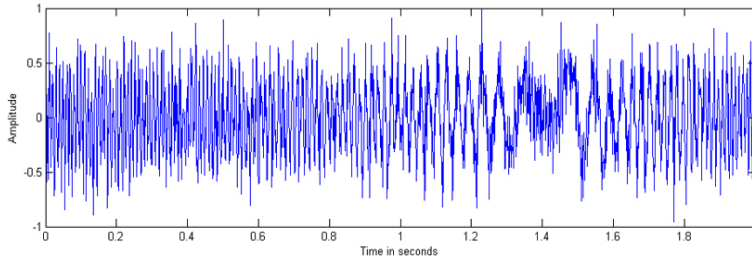
Figure 15.2: Small signal buried in white noise

Nevertheless, if we analyze the power spectrum we can point out the effective power density of the signal and of the noise.
The power

$$P_s = \frac{V_s^{\,2}}{2} = 50 \cdot 10^{-6}\ V^2$$

is within a bandwidth

$$\Delta f_s = 1\ Hz$$

So, the effective power density of the signal is

$$\sqrt{S_s} \approx \sqrt{\frac{P_s}{\Delta f_s}} = 70\ \frac{nV}{\sqrt{Hz}}.$$

It is thus evident that the narrower is our signal the higher is our power density.
On the other hand, the effective power density of our white noise is

$$\sqrt{S_b} = 5\ \frac{nV}{\sqrt{Hz}}.$$

If we then consider the ratio between the effective power densities, we can notice that:

$$\frac{\sqrt{S_s}}{\sqrt{S_b}} = 14 \gg 1$$

So, we can conclude that, in this case, with a bandpass filter having bandwidth Δf_b matched to the signal band Δf_s, we can obtain a good Signal-to-Noise ratio:

$$\frac{S}{N} = \sqrt{\frac{P_s}{S_b \Delta f_b}} = \sqrt{\frac{S_s \, \Delta f_s}{S_b \, \Delta f_b}} \sim \sqrt{\frac{S_s}{S_b}} = 14 \gg 1$$

If f_s is comparable to f_c, we have to consider also the contribution of 1/f noise in the calculation of the total spectral density:

$$\sqrt{S_n(f_s)} = \sqrt{S_b + S_b \frac{f_c}{f_s}}$$

In this case we could replicate the same calculations finding a value of the spectral density at f_s equal to $8.7 \frac{nV}{\sqrt{Hz}}$ and so a signal-to-noise ratio of 8 that is lower than 14 but greater than 1.

Finally, if f_s is lower than f_c, the situation is different: in this case we have to use, for example, $f_s = 10 \, mHz$, which is well below the corner frequency. We will again use a high pass filter with a band limit of $f_i = 1 \, Hz$ in order not to filter out the signal. Then the noise referred to the input is

$$\sqrt{v_n{}^2} \approx \sqrt{S_b \cdot f_h + S_b \cdot f_c \cdot ln\left(\frac{f_h}{f_i}\right)} \approx \sqrt{S_b \cdot f_h} \approx 5 \, \mu V$$

So, this situation in practically equal to the previous cases. In fact:

$$\frac{S}{N} = \frac{V_s}{\sqrt{v_n^2}} \leq 0{,}02 \ll 1$$

Again, the effective power density of the signal is

$$\sqrt{S_s} \approx \sqrt{\frac{P_s}{\Delta f_s}} = 70 \, \frac{nV}{\sqrt{Hz}}.$$

but if we focus on the noise at the signal frequency, now we have a significant contribution given by the $\frac{1}{f}$ noise:

$$\sqrt{S_n(f_s)} = \sqrt{S_b + S_b \frac{f_c}{f_s}} = 14.2 \sqrt{S_b} \approx 71 \, \frac{nV}{\sqrt{Hz}}$$

In this case we get a signal peak that is not so easy to distinguish from the noise. In particular, the ratio between the two power densities is about one so the signal to-noise ratio is very poor even if we employ a bandpass filter.

In this case we have to find a way to transfer the signal to higher frequency in order to be able to eliminate $\frac{1}{f}$ noise. In particular, we have to move the signal to higher frequencies before it mixes with $\frac{1}{f}$ noise introduced, for example, by the preamplifier.

15.4 Sinusoidal Signal Modulation

Until here we have seen that a narrow-band signal with some superimposed white noise can be recovered in a very efficient way using a matched Band-Pass Filter even if the signal is buried into noise.

If there's also some 1/f noise the situation is more complex: the lower the signal frequency, the higher is 1/f noise and recovering the signal might be impossible even using a matched bandpass filter (i.e. the SNR is very poor). In this scenario, a possible solution is the following: we shall try to "move" the signal away from the dominant noise. In other words, we can try to transfer the information at higher frequency in order to avoid the dominant influence of 1/f noise.

The problem is how to operate this transfer without losing information. The solution is the *modulation*. It is worth noting that we have to pay attention about when we operate this modulation: if we modulate the signal after it mixed with 1/f noise, modulation would be useless because we would treat the noise and the signal in the same way, and we would not obtain any SNR improvement.

Modulation consists in multiplying in time our input signal $x(t)$ with a reference signal called *carrier* $m(t)$.

A typical carrier is a sinusoidal signal

$$m(t) = B \cos (2\pi f_m t)$$

that performs the so-called sinusoidal modulation. In fact, in this case we have that the spectrum of the carrier is equal to

$$F[m(t)] = \frac{B}{2}\delta(f - f_m) + \frac{B}{2}\delta(f + f_m)$$

so multiplying the signal in time with $m(t)$ means shifting the signal spectrum $X(f)$ around $\pm f_m$, as we wanted.

We recall that the Fourier Transform of $B \cos (2\pi f_m t)$ is equal to two deltas centered in $\pm f_m$ with amplitude $\frac{B}{2}$, $F[Bcos(2\pi f_m t)] = \frac{B}{2}\delta(f - f_m) + \frac{B}{2}\delta(f + f_m)$; convolving a signal with deltas centered in f_m means to shift the signal itself at f_m.

15.4.1 Modulating ideal DC signals

An ideal DC signal concentrate all its information at f=0. In time we have $x(t) = A$, while in frequency we have $X(f) = A\delta(f)$.

If we modulate a DC signal with a cosinusoidal signal, at output we obtain:

$$y(t) = A \cdot B \cos (2\pi f_m t + \varphi m)$$

where φ_m has to be constant, and thus we can consider it now equal to zero. In frequency we have:

$$Y(f) = X(f) * M(f) = \int_{-\infty}^{+\infty} X(\tau) \cdot M(f - \tau)d\tau$$

That is

$$Y(f) = A \cdot \frac{B}{2} [\delta(f - f_m) + \delta(f + f_m)]$$

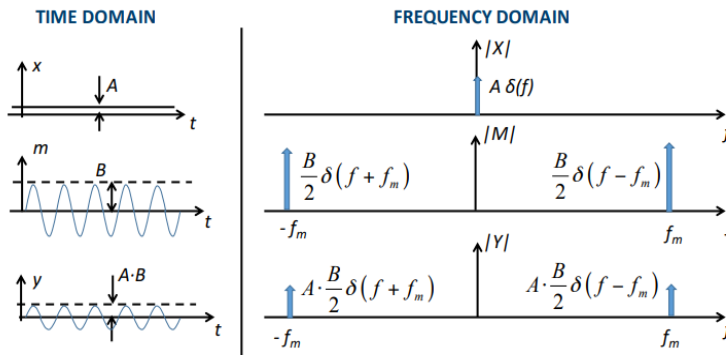As we wanted, now the information is carried by the amplitude of a (high-frequency) cosine.

Figure 15.3: Ideal DC Signal Modulation

15.4.2 Modulating Narrow-Band signal

Now we can try to analyze the case of Narrow-Band signals. As we saw in the introduction this type of signals is an approximation of a low frequency signal with a sufficiently small bandwidth.

The main issues of analyzing Narrow Band signals arises from the fact that the convolution in Frequency Domain is more complicated than in Time Domain because:

- The spectrum (differently from time domain) runs from $-\infty$ to $+\infty$ (instead of 0 to $+\infty$)
- The spectrum $X(f)$ is a *complex function* and so we have to deal with complex values.
- In general, modulus of convolution is not convolution of modulus

$$|Y(f)| \neq |X(f)| * |M(f)|$$

The issue presented above is simplified and it can be neglected if

- $X(f)$ is confined in a Narrow Bandwidth Δf_s
- $M(f)$ has a f_m much greater than the bandwidth of the signal: $f_m \gg \Delta f_s$

If these two hypotheses are satisfied, there is no sum of complex numbers because at each frequency there is only one dominant term and the others can be neglected. So, the convolution between the two signals can be computed in this way:

- Shift in frequency every component of $X(f)$ by $+f_m$ and $-f_m$
- Shift in phase every component of $X(f)$ by $+\varphi_m$ and $-\varphi_m$

we can conclude that *in this particular case* it is true that

$$|Y(f)| \approx |X(f)| * |M(f)|$$

and

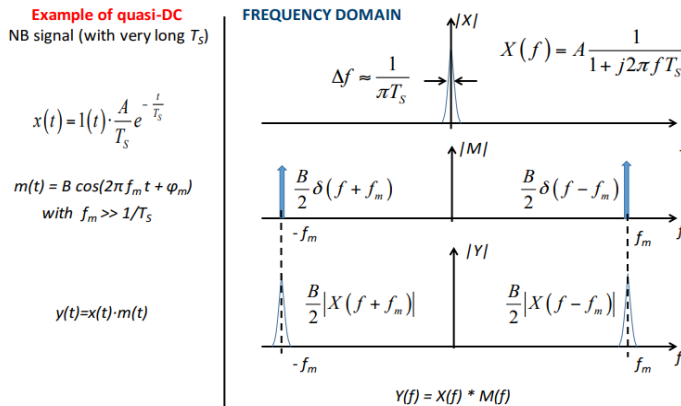$$|Y(f)|^2 \approx |X(f)|^2 * |M(f)|^2$$

Figure 15.4: Narrow-Band Signal Modulation

15.4.3 Modulating Sinusoidal signal

In case we have to deal with a Sinusoidal signal $x(t) = A\cos(2\pi f_s t)$, the situation is only a bit different. Recall that the Fourier transform of $x(t)$, $X(f)$, is the sum of two symmetric deltas centered in $\pm f_s$ with amplitude $\frac{A}{2}$. The same holds for the carrier $m(t) = B\cos(2\pi f_m t + \varphi)$.

Under the hypotesis $f_c \gg f_s$, the final result is graphically shown in figure 15.5
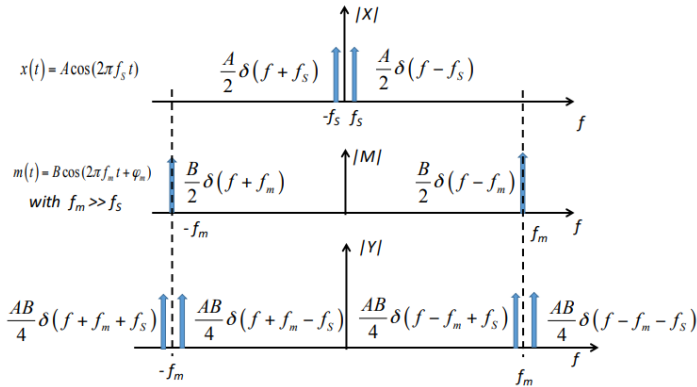

Figure 15.5: Sinusoidal Signal Modulation

It can also be verified analytically by recalling the trigonometric equivalence

$$cos(\alpha)cos(\beta) = \frac{1}{2}cos(\alpha - \beta) + \frac{1}{2}cos(\alpha + \beta)$$

we obtain

$$y(t) = x(t) \cdot m(t) = \frac{AB}{2}cos[2\pi(f_s - f_m)t - \varphi] + \frac{AB}{2}cos[2\pi(f_s + f_m)t + \varphi]$$

15.5 Square Wave Signal Modulation

The implementation of sinusoidal modulation requires analog multipliers that are slightly complex and noisy. What if we use a square wave instead of a sinusoidal signal as a carrier?

Modulation with a square wave reference *m(t)* can be implemented with circuits based simply on switches and amplifiers, without analog multipliers.

Let's consider a square wave $m(t)$ with period $\frac{1}{f_m}$, amplitude B ($2B$ peak to peak) and DC component B_0

If we look at its Fourier transform, we have delta shaped frequency components at odd multiples of square wave frequency, $(2h - 1)f_m$,with amplitude $\frac{2B}{\pi(2h-1)}$

$$M(f) = B_0\delta(f) + \sum_{h=1}^{\infty}(-1)^{-h+1}\frac{2B}{\pi(2h - 1)}[\delta(f - (2h - 1)f_m) + \delta(f + (2h - 1)f_m)]$$
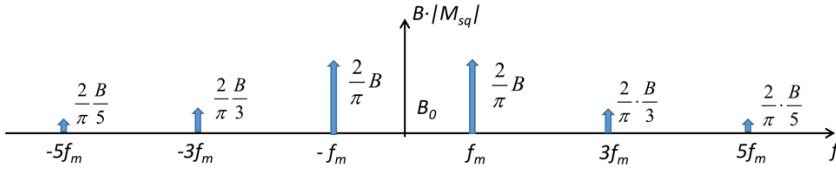
Figure 15.6: Spectral components of a generic square wave

As in sinusoidal modulation, also square wave modulation shifts the input signal spectrum in frequency, and if $\Delta f_s \ll f_m$:
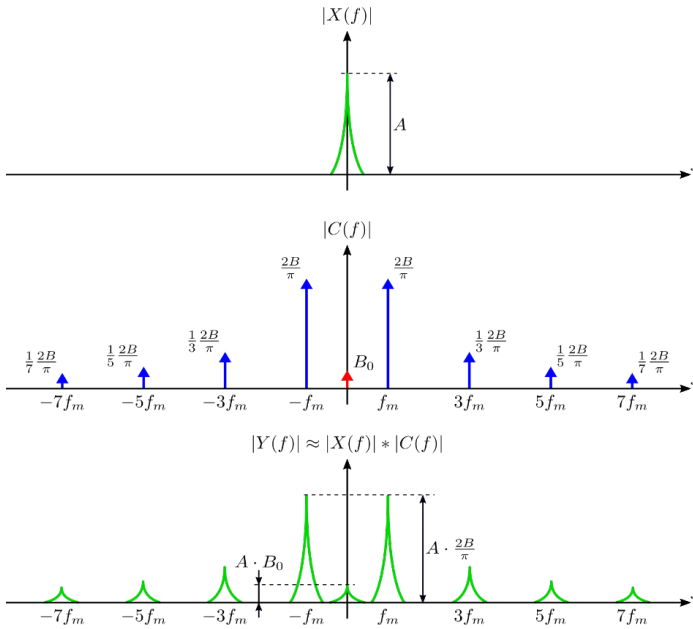
We have

$$|Y(f)| \approx |X(f)| * |M(f)|$$

Figure 15.7: Square wave modulation

If the square wave is not perfectly symmetrical (e.g. it has asymmetrical amplitude and/or duration of positive and negative parts) there is also a finite DC component with amplitude B_0 (possibly very small). The carrier DC component does NOT transfer the signal X in frequency, it just «amplifies» it by B_0 so it must be avoided in order to reduce the 1/f noise contribution.

After the analysis of the square-wave modulation effect on signals, we will now discuss how to implement a circuit to perform such modulation.

A simple implementation employing two single-ended amplifiers can be the one shown in figure 15.8
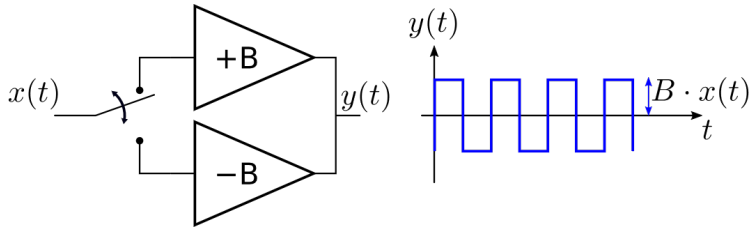

Figure 15.8: Simple circuit for square wave modulation

In principle, the sketched circuit performs the modulation using a square wave with zero DC value. However, to have zero-DC component the gains of the two amplifiers must be exactly the same (in module), but this is practically not feasible leading to an unwanted DC component.

To solve this problem, we can use either a differential amplifier, as shown in figure 15.9
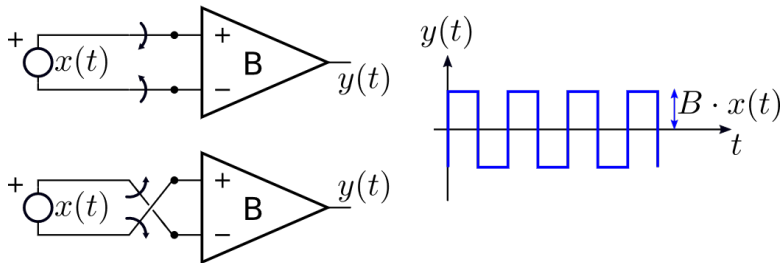

Figure 15.9: Single amplifier square wave modulator

or a single single-ended amplifier used as chopper, as shown in figure 15.10.
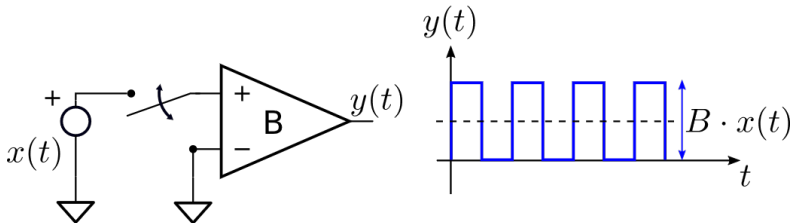

Figure 15.10: Chopper as modulator

The last configuration is the simplest, but it generates a DC component which amplitude is dependent on the signal value ($B_0 = \frac{B}{2}$).

In any case, square-wave modulators are much simpler than analog multipliers that are required to perform a sinusoidal modulation; this approach requires a lower number of transistors (lower cost) and it also improves the noise performance.

The main source of noise in square-wave modulator topologies are the switches that can be implemented either using MOSFETs (presence of $\frac{1}{f}$ noise, fast switching frequencies) or using mechanical metal switches, which are less noisy but have switching frequency limited to about 100Hz. In any case, the total added noise is typically much lower than that introduced by similar analog multipliers.

15.6 Summary

- As intuitive, narrow-band filtering is very effective to recover narrow-band signals immersed in wide-band noise
- Besides wide-band noise, however, other components with power density increasing as the inverse frequency (1/f noise) are ubiquitous in electronic circuitry (amplifiers etc.). In the low-frequency range they are indeed dominant.
- At low frequencies the 1/f noise added by the circuitry is overwhelming, so that the solution of narrow-band filtering becomes progressively less effective and finally insufficient for recovering signals with progressively lower frequency
- An effective approach to recover a low-frequency signal is to move it to higher frequency before the addition of 1/f noise. That is, to modulate the signal before the circuitry that contains the 1/f noise sources
- Narrow-band filtering can then be employed to recover the modulated signal
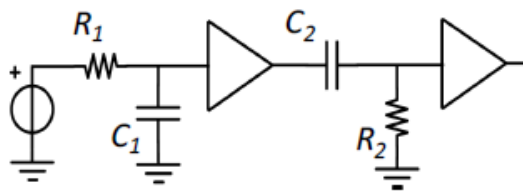
Band-Pass Filters 2

In the previous chapter we discussed the benefits of exploiting a band-pass filter when the signal is limited to a relatively small range of frequencies, allowing the attenuation or rejection of any component outside the range of interest. In this chapter, two solutions to implement a band-pass filter with analogue passive circuits are discussed, i.e. CR – RC and LRC resonant filters.

16.1 Introduction

In previous chapters we studied in details low-pass filters and high-pass ones. Now we focus on band-pass filters that can be used to filter out a certain range of frequencies and to attenuate frequencies outside this range. In these terms, band-pass filters can be seen as a combination of a low-pass filter (RC) and a high-pass one (CR): this is the case of CR – RC filters. Another example of an analogue passive band-pass filter is an LRC circuit, that is a combination (it can be used either in the series or in the parallel configuration) of a resistor, an inductor and a capacitor. In this chapter, both these types of circuits are discussed.

16.2 Band-pass filtering with a combination of High-pass and Low-Pass filters (CR - RC)

One of the band-pass filtering methods is the CR– RC. This filter basically consists of an integrator (low-pass filter) followed in cascade by a differentiator (high-pass filter) as sketched in figure 16.1. This is a second-order filter, because it has two reactive components (the two capacitors).

Figure 16.1: RC-CR band-pass filter circuit.

The transfer function can be computed considering the property of the Laplace transfer functions stating that the equivalent transfer function of the cascade of

multiple independent stages is the product of the transfer functions of its single blocks:

$$H(s) = \frac{V_{out}}{V_{in}} = \frac{1}{1 + sR_1C_1} \cdot \frac{sR_2C_2}{1 + sR_2C_2}$$

To figure out the noise filtering performance of the circuit, the transfer function in the frequency domain is required. The function can be obtained by replacing $s = j\omega = j2\pi f$:

$$H(f) = \frac{1}{1 + jf/f_{p1}} \cdot \frac{jf/f_{p2}}{1 + jf/f_{p2}},$$

where $f_{p1} = \frac{1}{2\pi R_1 C_1}$ is the pole of the low-pass filter and $f_{p2} = \frac{1}{2\pi R_2 C_2}$ is the pole of the high-pass filter.

Considering the case with equal poles $R_1C_1 = R_2C_2$, so $f_{p1} = f_{p2} = f_p$, the transfer function becomes:

$$H(f) = \frac{jf/f_p}{\left(1 + jf/f_p\right)^2}$$

and its square magnitude is:

$$|H(f)| = \frac{f/f_p}{1 + \left(f/f_p\right)^2}$$

At the center of the band, $f = f_p$, the peak value and the phase are:

$$|H(f_p)| = \frac{1}{2} \quad \arg\left(H(f_p)\right) = 0$$

In figure 16.2 the time-domain response of a CR-RC filter to a step input is shown and compared with the output of a CR and a RC filter. The CR filter, given a sufficiently-small time constant, outputs a voltage that is almost proportional to the time-derivative of the input waveform (CR differentiator); as a result, it is able to preserve the fast variations of the input while its output goes to zero in steady state. On the other hand, the RC filter, given a sufficiently-large time constant, outputs a signal that is almost proportional to the integral of the input signal (RC integrator); as a result, it reaches the same final value of the step input signal with a slow transition. The CR-RC filter shows a step-response that is the combination of the main features of its two building blocks, i.e. a smooth response upon fast input variations and a null output in steady state.
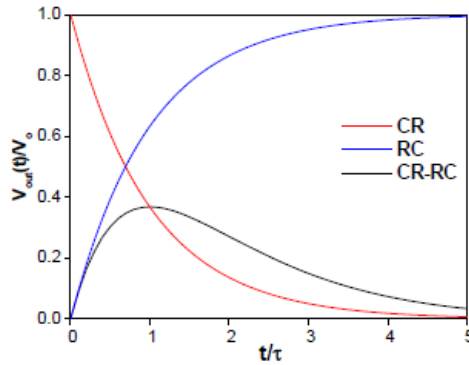
Figure 16.2: CR, RC and CR-RC time domain responses for a step function input.
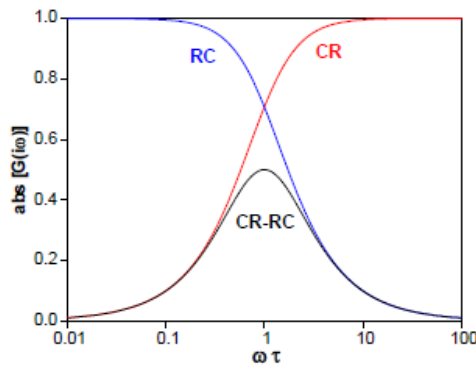

Figure 16.3: Absolute value of the frequency domain transfer function for CR, RC and CR-RC.

In figure 16.3, a comparison in the frequency domain is reported; in particular, the modulus of the CR-RC transfer function is shown along with the modulus of its building blocks. The modulus of both CR and RC filter is equal to $1/\sqrt{2}$ at the frequency $f = f_p$. The CR filter attenuates low frequency signals ($f < f_p$) while passes the high frequency signals. The RC filter shows the opposite trend. As a result of the cascade of these two stages, the CR-RC filter has a gain equal to 1/2 at $f = f_p$ and it is able to attenuate both low and high frequency components.

As a rule of thumb, we can consider the input components to be attenuated by the CR-RC filter if the modulus of the transfer function is lower than $1/\sqrt{2}$, that is 3dB lower than its peak amplitude; the corresponding boundary frequencies are called *cut-off* frequencies. For a band-pass filter, we can define the signal bandwidth as the frequency range that exists within two cut-off frequencies. Hence, the bandwidth can be simply defined as the difference between the lower cut-off frequency (f_{pl}) and the higher cut-off frequency (f_{ph}) points.

$$|H(f_{pl})| = |H(f_{ph})| = \frac{|H(f_p)|}{\sqrt{2}} = \frac{1}{2\sqrt{2}}$$

With $x = f/f_p$, the equation can be written as:

$$\frac{x}{1+x^2} = \frac{1}{2\sqrt{2}}$$

that gives as solution $x_{l,h} = \sqrt{2} \pm 1$, so in terms of frequency $f_{pl,ph} = (\sqrt{2} \pm 1)f_p$. Therefore, the signal bandwidth is

$$\Delta f_p = f_{ph} - f_{pl} = 2f_p.$$

We can observe that the bandwidth obtained with this topology is not narrow, since it's the double of the peaking frequency. In particular, this might be a problem if the input signal is centered at high frequencies, i.e. MHz. In this case, the bandwidth of the bandpass filter could be of the order of MHz. This means that the input signal may not be filtered as needed and a great amount of noise is not filtered at all.
To confirm this theory, the equivalent white-noise bandwidth will be computed in the following.
From the definition of white-noise bandwidth, we can write:

$$\overline{n_B{}^2} = S_B |H(f_p)|^2 \Delta f_n = S_B \, {}^1/_4 \, \Delta f_n$$

And by comparison with the computed output power:

$$\overline{n_B{}^2} = S_B \int_0^\infty |H(f_p)|^2 \, df_n = S_B \int_0^\infty \frac{(f/f_p)^2}{\left[1 + (f/f_p)^2\right]^2} \, df_n$$

$$= S_B f_p \int_0^\infty \frac{x^2}{[1+x^2]^2} \, dx = S_B f_p \int_0^\infty \frac{2x}{[1+x^2]^2} \frac{x}{2} dx$$

$$= S_B f_p \left\{ \left| -\frac{x}{2} \frac{1}{1+x^2} \right|_0^\infty + \frac{1}{2} \int_0^\infty \frac{1}{1+x^2} \, dx \right\}$$

$$= S_B f_p \frac{1}{2} |arctg(x)|_0^\infty = S_B f_p \frac{\pi}{4}$$

the equivalent white noise bandwidth may be calculated as follows:

$$\Delta f_n = \pi f_p = \frac{1}{2RC} = \frac{\pi}{2} \Delta f_p$$

As discussed for the signal, this is a quite large value that will prevent the use of this kind of filter in many applications.

For the sake of completeness, the δ-response of the CR-RC will be now computed:

$$H(s) = \frac{sRC}{(1 + sRC)^2} = \frac{1}{1 + sRC} - \frac{1}{(1 + sRC)^2}$$

By applying the inverse Laplace transform we obtain

$$h(t) = \frac{1}{RC} e^{-\frac{t}{RC}} - \frac{t}{(RC)^2} e^{-\frac{t}{RC}}$$

Starting from the $h(t)$, we can compute the autocorrelation function, and then we can calculate its value in zero. By definition of autocorrelation function in $t = 0$:

$$k_{hh}(0) = \int_0^\infty h^2(\alpha) \, d\alpha = \frac{1}{4RC}$$

By handling the Fourier transform properties:

$$k_{hh}(0) = \int_{-\infty}^\infty |H(f)|^2 df = |H(f_p)|^2 \, 2\Delta f_n = \frac{1}{2} \, \Delta f_n$$

which confirms that: $\Delta f_n = \frac{1}{2RC} = \frac{\pi}{2} \Delta f_p$.

16.3 LRC Resonant Filter

In this section, the LRC resonant filter is discussed. This kind of filter can be implemented with different topologies. In particular, we will here discuss the parallel resonant filter.

The circuit that implements the parallel resonant filter is composed by a resistor, an inductor and a capacitor all in parallel to each other as shown in figure 16.4. This is typically used with current input signals.
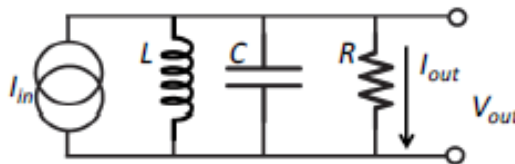

Figure 16.4: LRC parallel circuit configuration.

A parallel resonant circuit stores the circuit energy in the magnetic field of the inductor and in the electric field of the capacitor.

The circuit analysis is based on the calculation of the transfer function between the voltage output signal and the current input one, which is an equivalent impedance.

$$Z = \frac{V_{out}}{I_{in}} = \frac{1}{j\omega C + \frac{1}{j\omega L} + \frac{1}{R}} = R \frac{j\omega \frac{L}{R}}{(1 - \omega^2 LC) + j\omega \frac{L}{R}}$$

The total equivalent impedance of the circuit is given by:

$$Z_{tot} = R \,||\, Z_L \,||\, Z_C = \frac{R}{1 - jR\left(\frac{1}{X_C} + \frac{1}{X_L}\right)} = \frac{R}{1 + jR\left(\omega C - \frac{1}{\omega L}\right)}$$

The impedance of the circuit has its maximum magnitude at the resonant frequency f_0, that is when the inductive impedance is the opposite of the capacitive impedance, i.e. $X_L = - X_C$. Hence, the resonant frequency is given by the relationship:

$$X_L + X_C = 0 \qquad \omega_o = \sqrt{\frac{1}{LC}}$$

where $\omega_o = 2\pi f_o$ is the resonant angular frequency.

As a result, the total impedance of the parallel circuit at resonance assumes its maximum value and it is equal to just the value of the resistance:

$$Z_{tot}(\omega_o) = R$$

Let's study in more detail the impedance of this filter. Denoting by

$$H(\omega) = \frac{I_{out}}{I_{in}} = \frac{1}{R}\frac{V_{out}}{I_{in}} = \frac{j\omega \frac{1}{RC}}{\left(\frac{1}{LC} - \omega^2\right) + jw\frac{1}{RC}}$$

we can write
$$Z = R\, H(\omega).$$

At the resonance frequency, the reactive impedances cancel each other so that the impedance is purely resistive. As a result

$$H(\omega_o) = 1 \quad and \quad \arg H(\omega_o) = 0$$

Another basic parameter is the characteristic resistance R_o, which for the oscillation at $\omega = \omega_o$ represents the ratio

(amplitude of voltage on C) / (amplitude of current in L)

and can be written as

$$R_0 = \sqrt{\frac{L}{C}}$$

The poles of H(ω) are defined as:

$$s_{1,2} = -\frac{1}{2RC} \pm \frac{1}{2RC}\sqrt{1 - 4\frac{RC}{\frac{L}{R}}} = -\frac{1}{2RC} \pm \frac{1}{2RC}\sqrt{1 - \frac{R^2}{\left(\frac{R_0}{2}\right)^2}}$$

$$= -\frac{1}{2RC} \pm \sqrt{\left(\frac{1}{2RC}\right)^2 - \frac{1}{LC}}$$

defining

$$\alpha_0 = \frac{1}{2RC}$$

we can rewrite the poles of the transfer function as

$$s_{1,2} = -\alpha_0 \pm \sqrt{\alpha_0^2 - \omega_o^2}$$

and we can study the behavior of the δ-response as a function of these parameters. The δ-response h(t) is:

- damped (real poles)   if $\omega_o^2 < \alpha_0^2$ which implies $R^2 < \left(\frac{R_0}{2}\right)^2$
- critically damped (real poles)   if $\omega_o^2 = \alpha_0^2$ that is if $R^2 = \left(\frac{R_0}{2}\right)^2$
- oscillatory (complex poles)   if $\omega_o^2 > \alpha_0^2$ that is with $R^2 > \left(\frac{R_0}{2}\right)^2$

The higher is R with respect to R_o, the lower is the dissipation and the slower is the damping of the oscillation.

Starting from these observations we can define the *quality factor* that is an important parameter of many systems as, for example, a resonator and oscillator. Physically speaking, Q is defined as the ratio of the total stored energy divided by the energy dissipated over one radian of the oscillation. In LRC resonant filters, it describes how under-damped the resonator is; the higher the Q factor the less damping there

is. In an electrical resonator circuit, damping is caused by the loss of energy in resistive components.

The Q factor can therefore be described as:

$$Q(\omega) = \omega \, \frac{\text{Maximum Energy Stored}}{Power\ Loss}$$

The Q factor depends on frequency, but it is most often represented as function of resonance ω_o.

The maximum energy stored in a resonant circuit at the resonant frequency can be easily calculated from either the maximum energy stored in the capacitor or the maximum energy stored in the inductor (in general, the total energy is equal to the sum but here the computation is simplified since when one component is holding its maximum energy, the other one has zero energy in it).

$$\text{Maximum Energy Stored} = L\, I_{rms}{}^2 = C V_{rms}{}^2$$

Where $I_{rms}{}^2$ and $V_{rms}{}^2$ are the root mean square (rms) current through the inductor and the rms voltage across the capacitor respectively.

The resistor is the only component in the circuit that dissipates power:

$$Power\ Loss = V_{rms} I_{rms} = \frac{V_{rms}{}^2}{R} = I_{rms}{}^2 R$$

In a parallel LRC resonant circuit:

$$Q = \frac{\omega_o}{2\alpha_0} = \frac{R}{R_0}$$

The higher is R ($R \gg Ro$) the lower is the dissipation (Q $\to \infty$ for R $\to\infty$).

We can express the transfer function of the circuit as a function of the quality factor and w_o

$$H(w) = \frac{j\omega\frac{\omega_o}{Q}}{(\omega_o{}^2 - \omega^2) + j\omega\frac{\omega_o}{Q}}$$

It is interesting to analyze the behavior of this function varying the quality factor Q; in figure 16.5 it is possible to see the δ-response of the filter for different values of the Q.
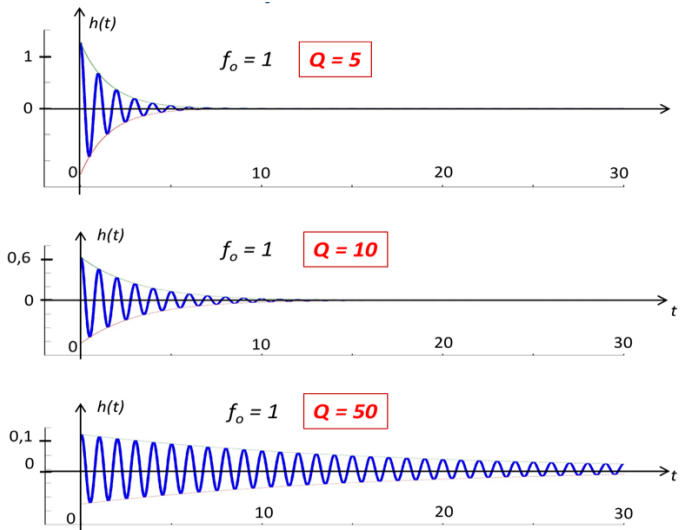
Figure 16.5: Plot of the δ-response as a function of the quality factor.

An extremely important parameter of the band pass filter for our purposes is the phase and, even more important is the variation of the phase as a function of the frequency. We will use this property in the next chapter.
For the LRC circuit, the phase can be written as:

$$\varphi = \arg H\left(\omega\right) = arctg\left[\frac{Q}{\omega\omega_o}(\omega_o + \omega)(\omega_o - \omega)\right]$$

so   for $\omega \to +\infty$ $|H| \to 0$ $\phi = \arg H(\omega_o) \to -\pi/2$ (- 90°)
          for $\omega \to -\infty$ $|H| \to 0$ $\phi = \arg H(\omega_o) \to +\pi/2$ (90°)
while   for $\omega = \omega_o$ $H(\omega_o) = 1$ $\phi = \arg H(\omega_o) = 0$

The phase-shift given by the filter is exactly zero at exactly the band center, but it rapidly increases at different ω. We can calculate the derivative of the phase shift as a function of ω at ω= ω0:

$$\left(\frac{d\varphi}{d\omega}\right)_{\omega=\omega_o} = -2\frac{Q}{\omega_o}$$

The higher is Q the steeper is the increase, as it possible to see in the figure 16.6.
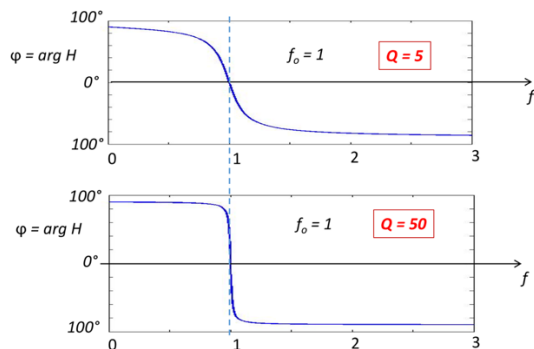
Figure 16.6: Phase of the LCR as a function of the frequency for two different quality factor.

16.4 Approximations of the LRC transfer function

We have seen that the transfer function of LRC filter is a quite complicated expression so working with it is not trivial. To simplify the task, we can do some approximations in order to have a simpler expression of the module of the function. First of all, let's rewrite the square modulus of the transfer function as follows:

$$|H(\omega)|^2 = \frac{\omega^2 \frac{\omega_o^2}{Q^2}}{(\omega_o^2 - \omega^2)^2 + \omega^2 \frac{\omega_o^2}{Q^2}} = \frac{1}{1 + \frac{Q^2(\omega_o^2 - \omega^2)^2}{\omega_o^2 \omega^2}}$$

$$= \frac{1}{1 + \frac{Q^2(\omega_o + \omega)^2(\omega_o - \omega)^2}{\omega_o^2 \omega^2}}$$

Now, we can have three possible cases:

1. $\omega \gg \omega_o$, **"Higher wing"** approximation:

$$|H(\omega)| = \sqrt{\frac{1}{1 + \frac{Q^2(\omega_o + \omega)^2(\omega_o - \omega)^2}{\omega_o^2 \omega^2}}} \approx \sqrt{\frac{1}{1 + \frac{Q^2(\omega)^2(\omega)^2}{\omega_o^2 \omega^2}}} \approx \frac{\omega_o}{\omega} \frac{1}{Q}$$

2. $\omega \ll \omega_o$, **"Lower wing"** approximation:

$$|H(\omega)| = \sqrt{\frac{1}{1 + \frac{Q^2(\omega_o + \omega)^2(\omega_o - \omega)^2}{\omega_o^2 \omega^2}}} \approx \sqrt{\frac{1}{1 + \frac{Q^2(\omega_o)^2(\omega_o)^2}{\omega_o^2 \omega^2}}} \approx \frac{\omega}{\omega_o} \frac{1}{Q}$$

3. $|\omega - \omega_o| \ll \omega_o$, **"Central lobe"** approximation:

$$|H(\omega)| = \sqrt{\cfrac{1}{1 + \cfrac{Q^2(\omega_o + \omega)^2(\omega_o - \omega)^2}{\omega_o^2\omega^2}}} \approx \sqrt{\cfrac{1}{1 + \cfrac{Q^2(2\omega_o)^2(\omega_o - \omega)^2}{\omega_o^2\omega^2}}}$$

$$\approx \cfrac{1}{1 + 4Q^2\left(\cfrac{\omega_o - \omega}{\omega_o}\right)^2}$$

Let's see with a simulation the validity of these approximations.
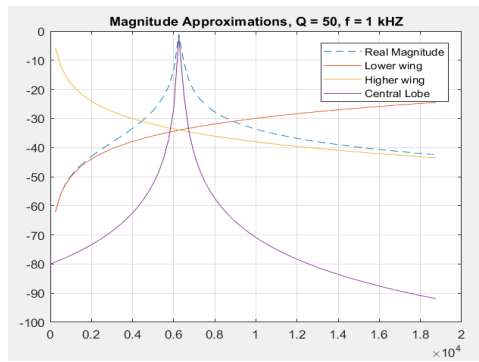

Figure 16.7: Magnitude approximations with Q=50 and f=1kHz.

Figure 16.7 shows that the lower-wing, central-lobe and right-wing approximations well fit the actual transfer function at frequencies much lower, close to or much higher than the resonance frequency, respectively.

In the first part of this chapter, we have introduced the concept of noise bandwidth for a band-pass filter. We can do the same for the signal: as frequency boundaries of the pass band interval we can take the two -3dB points ω_{dL} and ω_{dH}, where $|H(\omega_{dL})|^2 = |H(\omega_{dH})|^2 = 1/2$. The selectivity of the circuit is a measure of its ability to reject any frequencies outside this interval. A more selective circuit will have a narrower bandwidth, whereas a less selective circuit will have a wider bandwidth.

$$\Delta\omega_s = \omega_{dH} - \omega_{dL}$$

If we use the central lobe approximation for cases in which Q>>1 we get:

$$|H_C(\omega_d)|^2 = \cfrac{1}{1 + 4Q^2\left(\cfrac{\omega - \omega_o}{\omega_o}\right)^2} = \cfrac{1}{2}$$

that is

$$4Q^2\left(\cfrac{\omega_d - \omega_o}{\omega_r}\right)^2 = 1$$

and we get

$$\omega_{dH} - \omega_o = \omega_o - \omega_{dL} = \frac{\omega_o}{2Q}.$$

The signal bandwidth thus is

$$\Delta\omega_s = \frac{\omega_o}{Q}$$

Two basic advantages with respect to the CR-RC bandpass filter are quite evident:
• No signal attenuation at the center frequency;
• Narrow filtering bandwidth even with moderately high Q values.

The bandwidth for white noise is defined by

$$\Delta f_n = \int_0^\infty |H(f)|^2 df$$

In cases with $Q \gg 1$ we can use for H(f) the central lobe approximation, taking into account that with good approximation $|H_c(f)|^2$ is symmetrical with respect to the band center f_r, thus obtaining:

$$\int_0^\infty |H(f)|^2 df \approx 2 \int_{f_0}^\infty |H_c(f)|^2 df$$

Using central lobe approximation, we get:

$$\Delta f_n = 2 \int_{f_0}^\infty \frac{1}{1 + 4Q^2 \left(\frac{f - f_r}{f_r}\right)^2} df$$

if we substitute $2Q \left(\frac{f - f_r}{f_r}\right) = x$, we obtain

$$\Delta f_n = \frac{f_r}{Q} \int_0^\infty \frac{1}{1 + x^2} dx = \frac{f_r}{Q} \, artg(x)_0^\infty = \frac{f_r}{Q} \frac{\pi}{2}$$

since $\frac{\Delta f_s}{f_r} = \frac{1}{Q}$, we can rewrite bandwidth for white noise as follows:

$$\Delta f_n = \frac{f_r}{Q} \frac{\pi}{2} = \frac{\pi}{2} \Delta f_s$$

	SERIES	PARALLEL
CIRCUIT		
TRANSFER FUNCTION	$H(\omega) = \dfrac{V_{out}}{V_{in}}$ $= \dfrac{j\omega \frac{R}{L}}{\left(\frac{1}{LC} - \omega^2\right) + j\omega \frac{R}{L}}$	$H(\omega) = \dfrac{I_{out}}{I_{in}}$ $= \dfrac{j\omega \frac{1}{RC}}{\left(\frac{1}{LC} - \omega^2\right) + j\omega \frac{1}{RC}}$
RESONANT FREQUENCY	$\omega_o = \sqrt{\dfrac{1}{LC}}$	$\omega_o = \sqrt{\dfrac{1}{LC}}$
½ POWER FREQUENCY	$\omega_{dL} = -\dfrac{R}{2L} + \sqrt{(\frac{R}{2L})^2 + \frac{1}{LC}}$ $\omega_{dH} = +\dfrac{R}{2L} + \sqrt{(\frac{R}{2L})^2 + \frac{1}{LC}}$	$\omega_{dL} = -\dfrac{1}{2RC} + \sqrt{(\frac{1}{2RC})^2 + \frac{1}{LC}}$ $\omega_{dH} = +\dfrac{1}{2RC} + \sqrt{(\frac{1}{2RC})^2 + \frac{1}{LC}}$
BANDWIDTH	$\Delta\omega_s = \dfrac{R}{L}$	$\Delta\omega = \dfrac{1}{RC}$
Q FACTOR	$Q = \dfrac{1}{\omega_o RC}$	$Q = \dfrac{R}{\omega_o L}$

Figure 16.7: Summary of the properties of RLC resonant circuits

16.5 PRO'S AND CON'S OF REAL TUNED FILTERS

Real capacitors and inductors are not pure C and L. Their equivalent circuits include also finite resistances that model the internal sources of energy dissipation that inherently limit the Q of resonant circuits.

- In general, the dissipation is higher in components with higher value of L or C. Good quality capacitors with low dissipation are available from pF to about 1 µF. Inductors are more problematic than capacitors. Good quality components are available from nH to a few 100nH. Even components with fairly small L (typically a few 10 nH) have non-negligible internal resistance.
- Stray reactance must not be overlooked. In discrete circuitry stray capacitances are in the order of pF and stray inductances are in the order of nH. In integrated

circuits the values are much smaller, thanks to the very small physical size of the components.

- Since the resonance is at $f_o = 1/2\pi\sqrt{LC}$, for obtaining a low frequency f_o high values of both L and C are required: in fact, with C=1 μF and L= 100 nH one gets f_o =1,26 MHz. Therefore, the Q values really obtained in the tuned filters progressively decrease as the desired resonant frequency decreases.

For high frequencies f_o > 100MHz values of Q > 10 are currently obtained, up to almost Q ≈ 100 with clever design and high-quality components.

- For intermediate frequencies 1MHz < f_o < 100MHz values up to Q ≈ 10 are obtained with careful design and implementation
- For f_o < 1MHz it becomes progressively more difficult to obtain high Q values as the frequency decreases. Anyway, even with moderate Q the performance of the tuned filters is remarkable and in many practical cases filters with Q ≈ 5 are really satisfactory.
- Constant-parameter tuned filters are a simple and economical solution, widely employed in prefiltering stages and other simple situations, but their use in high-performance filtering is hindered by some intrinsic drawbacks.

The accuracy and relative stability of f_o directly depends on that of the C and L values. Drift of f_o due to aging and temperature must be kept smaller than the filter bandwidth, in order to avoid uncontrolled variations of the output signal amplitude and phase. In particular, strong phase variations are caused by even small variations of f_o because of the strong $d\phi/df$ at band-center of filters with high Q

- Cascading simple filter stages for improving the cutoff characteristics is not practical for narrow-band filters, because they should have very accurately equal and stable f_o.
- The value of C influences both the center frequency f_o and the bandwidth Δf_s so that it is not easy to design a filter with specified f_o and specified Δf_s .
- It is even more difficult to design a filter with adjustable f_o and constant bandwidth Δf_s.
- In cases where the frequency of a narrow-band signal is not very stable, a filter with very narrow bandwidth can be employed only if its center frequency can be adjusted to track that of the signal. As above outlined, this is not easy to obtain.

Band-Pass Filters 3

In this chapter we will introduce the basic function and features of the lock-in amplifier. We will describe the principle of synchronous and asynchronous measurements and we will show the advantages of synchronous measurements compared to the others. Then, we will describe the principle of lock-in amplifier and analyze stage-by-stage the behavior of both signal and noise.

17.1 Introduction

In this chapter, we will discuss the measurement of sinusoidal signals as an example of the wider class of narrow-band signals. The measurement of these signals is particularly challenging because, due to its narrow band, it is not so easy to acquire just the signal without collecting too much surrounding noise. We will study both synchronous and asynchronous measurements.

17.2 Asynchronous measurement of sinusoidal signals

The asynchronous measurement is so called because it does not exploit a reference signal, which is a signal that provides information on the frequency and phase of the modulated input, but it just works on the input signal. On one hand, these solutions are easy to implement and useful to acquire the signal; on the other hand, they present a wide set of limitations concerning noise filtering performance and nominal working conditions.

We'll analyze three reasonable configurations:

- the Mean-square Detector;
- the Half-Wave Rectifier;
- the Full-Wave Rectifier;

Before going into details, it's worth mentioning that these circuit blocks only work properly when the input is an almost perfect sinusoidal wave without any tangible DC component. To guarantee that this condition is verified, a high-pass (or band-pass) filter might be needed before them.
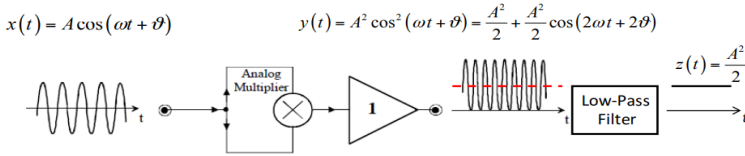
17.2.1 The Mean-Square Detector

Figure 17.1 Mean-square detector scheme

The idea is to measure the total input mean power. To do that, we can just multiply the signal by itself and make the average. Let's see in more detail:

$$y(t) = A^2 \cos^2(\omega t + \vartheta) = \frac{A^2}{2} + \frac{A^2}{2} \cos(2\omega t + 2\vartheta)$$

in which the following relationship has been used

$$cos^2 x = \frac{1 + \cos(2x)}{2}$$

We can then remove the components at the double of the original signal frequency obtaining the amplitude information. It is important to notice that the low-pass filter has no effect of noise reduction: in fact, it does not average the input, it averages the square of the input. For improving the S/N it is necessary to insert a filter before the Mean-Square detector.
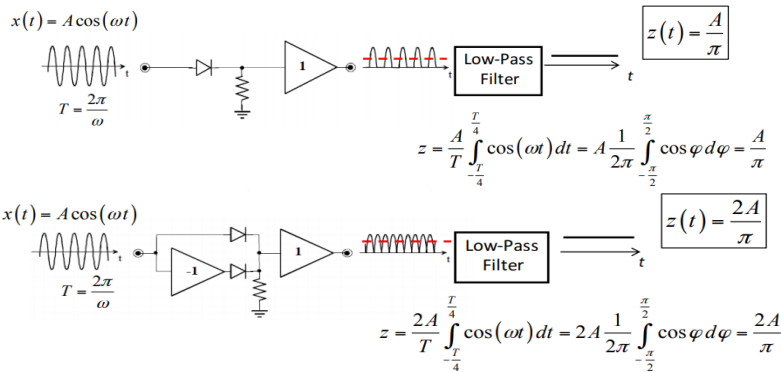
17.2.2 The Half-Wave and Full-Wave Rectifier

Figure 17.2 Half-wave and full-wave rectifier schemes

In both cases, we are implementing the classical rectifier in order to extract the amplitude of our signal.
The measurement with a rectifier is not really asynchronous, it is self-synchronized.

The sinusoidal signal itself decides when it has to be passed with positive polarity and when it passes with negative polarity (in the full-wave rectifier) or not passed at all (in the half-wave rectifier). In such operation, the LPF reduces the contribution of the wide-band noise, thus improving the output S/N. However, this is true only if the input signal is already remarkably higher than the noise, i.e. if the input S/N is high.

As the input signal is reduced, the noise gains increasing influence on the switching time of the rectifier, which progressively loses synchronism with the signal and tends to be synchronized with the zero-crossings of the noise.

The loss of synchronization progressively degrades the noise reduction by the LPF. With moderate S/N the improvement due to LPF is modest; with low S/N it is very weak. With S/N < 1 there is no improvement, there is not even a measure of the signal: the output is a measure of the noise mean absolute value.

In conclusion, meters based on rectifiers can just improve an already good S/N. They can't help to improve a modest S/N and it is out of the question to use them when S/N <1. For improving S/N it is necessary to employ filters before the meter.

17.3 Synchronous measurement of sinusoidal signals

Synchronous detection is a useful technique for extracting low-level signals buried below the noise floor. Some examples include measuring light absorption in the presence of strong background light or making strain measurements in the presence of high noise levels.

Many electrical systems have increasing noise as the frequency approaches dc (e.g. 1/f noise in amplifiers, background noise, etc.). The aim is then to move the measurement away from these low frequency noise sources, so that you can achieve a much better signal-to-noise ratio and detect much weaker signals.

For example, connecting a resistance measurement bridge, like in the figure, to an AC signal source V_A will make it possible to measure signals that would otherwise be buried in the noise at lower frequencies. This type of signal conversion is known as *modulation* and it makes use of an auxiliary signal, the *carrier*. The *reference* is instead a signal with equal frequency and phase of the carrier.

In our example, the AC voltage supply (usually a sine wave) acts as the carrier. The result of the modulation is a shift of the signal frequency band to a position around the carrier frequency, as explained in the previous chapter.

In our example we consider a bridge with only one resistance (R_T) that is sensitive to the the measurand ϑ (i.e. a temperature or a strain). Assuming equal values of the three other resistances, V_3 is just half the carrier and V_T is half the carrier modulated by R_T.

The bridge output $V^+ - V^-$ is therefore an amplitude modulated signal with no DC component.

We want to measure the amplitude A of our output signal as it is proportional to the measurand ϑ (R_T is a strain sensor and its resistance varies following a mechanical strain ϑ):

- in cases with **constant** strain ϑ: A is constant and x(t) is a pure sinusoidal signal;
- in cases with **slowly variable** strain $\vartheta = \vartheta(t)$: A = A(t) is variable and x(t) is a modulated sinusoidal signal;

We will discuss different possible setups (worst to best) to perform the measurement:

- Sample & Hold (*single peak sampling*)
- Synchronous measurement with averaging over many samples N >>1 of the peak
- Synchronous measurement with DC suppression by summing positive peak and subtracting negative peak samples
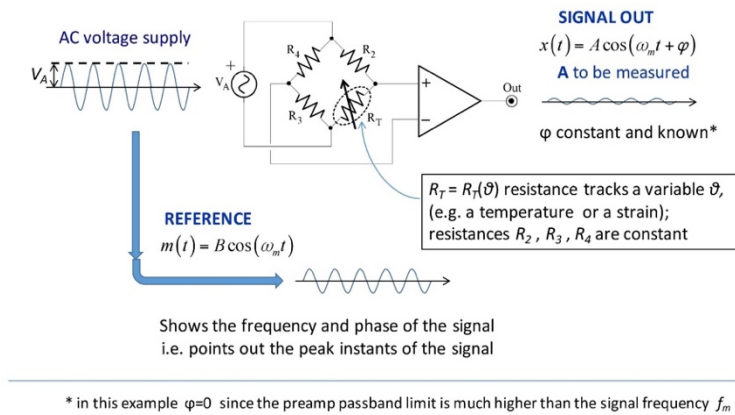- Continuous Sinusoidal Weighting

Figure 17.3: Simple modulation scheme

17.3.1 Elementary Synchronous Measurement: Peak Sampling

A first elementary approach is the *peak sampling*. We are interested in the signal's peak as its amplitude is equal to the quantity A that we want to measure.
This setup relies on the reference signal, which is required to meet the frequency precision constraints. The reference signal is fed to a sampling driver which detects the peak and outputs a $\partial(t)$ to enable the S&H.
The peak is then sampled by the S&H which provides a DC output signal of amplitude A.

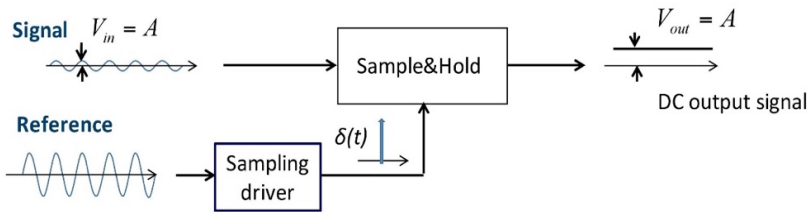**Figure 17.4:** Synchronous measurement with peak sampling
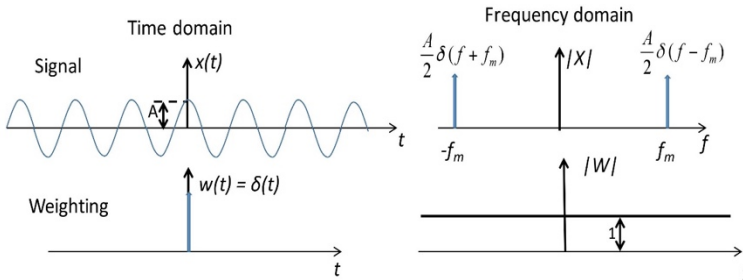

Figure 17.5: time domain and frequency domain of peak sampling approach

With this approach, although easy to implement, there is one big drawback in the use of this approach: it provides *no filtering action*.

As we can see in the fig. above, the frequency response is constant among all frequencies and therefore the output noise power is equal to the full input noise power.

17.3.2 Synchronous measurement with averaging over many samples

This second approach is an attempt to improve the filtering action with respect to the previous method by exploiting the repetitive occurrence of the peaks.

The modulated signal

$$x(t) = A \cos(2\pi f_m t)$$

is sampled at its peak N times. Taking N samples is equivalent to gate (or window) a free-running sampler and therefore the resulting weighting function is

$$w(t) = m(t) \cdot r_T(t)$$

where $r_T(t)$ is a rect of duration 2T. In the frequency domain, since we have a multiplication in the time domain, we will have

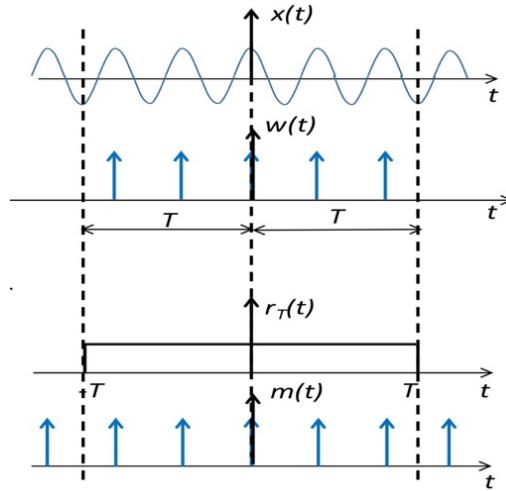$$\left| W(f) \right| \approx \left| M \right| * \left| R_T \right|$$

Figure 17.6: Synchronous measurement with averaging over many samples

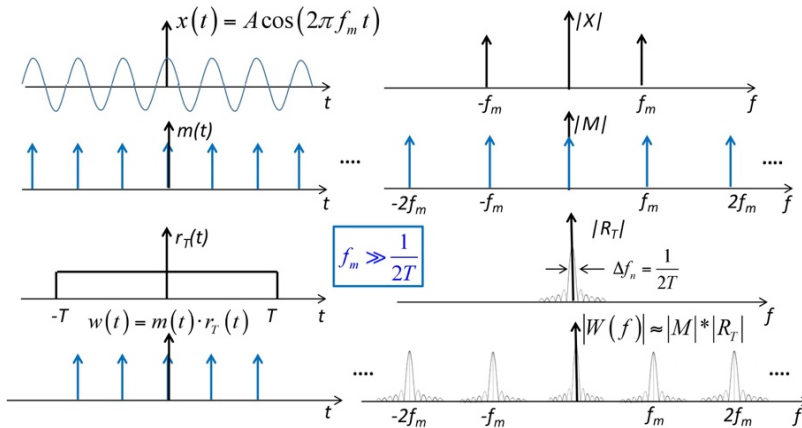where $|R_T|$ is a sinc function centered at DC and with $\Delta f_n = \frac{1}{2T}$.

Figure 17.7: Time domain and frequency domain behavior of the synchronous measurement with averaging over many samples

Comparing the frequency spectrum of $|W|$ here obtained (figure 17.7) with the one obtained with a Single Sampling (figure 17.5) we can see how the filtering action was improved. Instead of collecting white noise at all the frequency, now the collection is limited to some portions of the spectrum.
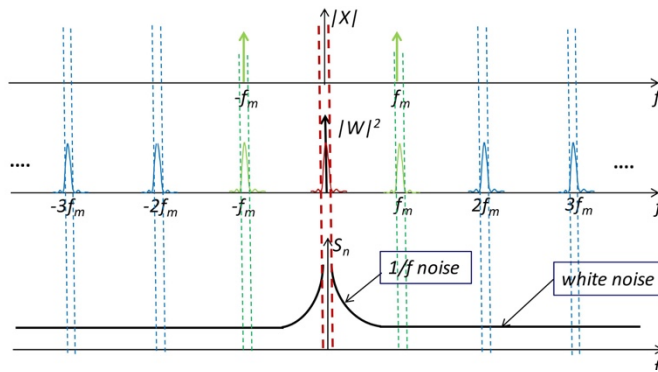
Figure 17.8: noise collecting bands of the synchronous measurement with averaging over many samples

However, the Sample-Averaging's noise filtering is still poor. Let's see why:

- at f_m: band is useful, since we are collecting the signal (and some inevitable white noise around it);
- at DC: band is *very* harmful, because we are collecting $\frac{1}{f}$ noise without acquiring any signal;
- at $2f_m, 3f_m, ..$: band is harmful, as we are collecting white noise without acquiring any signal;

17.3.3 Synchronous measurement with DC suppression by summing positive peak and subtracting negative peak samples

As we have seen, the biggest drawback of the Sample-Averaging is constituted by the narrow band at 0 frequency which collects $\frac{1}{f}$ noise and greatly worsens the Signal-to-Noise ratio. This new setup offers a solution to this issue.

Once again, to take 2N samples is equivalent to gate a free-running sampler:

$$w(t) = m(t) \cdot r_T(t)$$

Note that the number of samples is now 2N (instead of N, as in the previous case) to underline the fact that it must be even. The number of *positive* and *negative* samples has, in fact, to be equal so that the net area is *zero*. This is necessary as net area zero ↔ null DC component, which is precisely what we aimed for.

Looking at figure 17.10, it is possible to see how the narrow band at 0 frequency is suppressed.
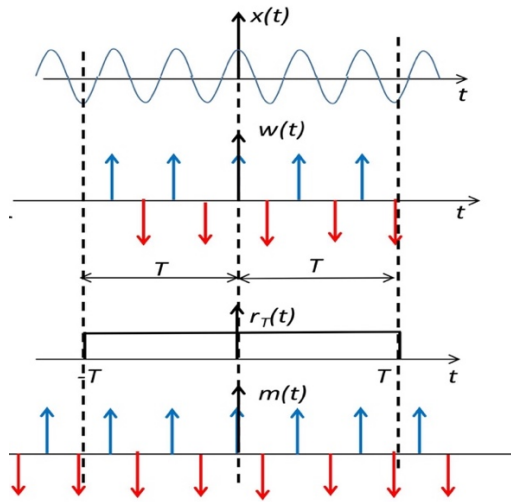
Figure 17.9 Synchronous measurement with DC suppression by summing positive peak and subtracting negative peak samples
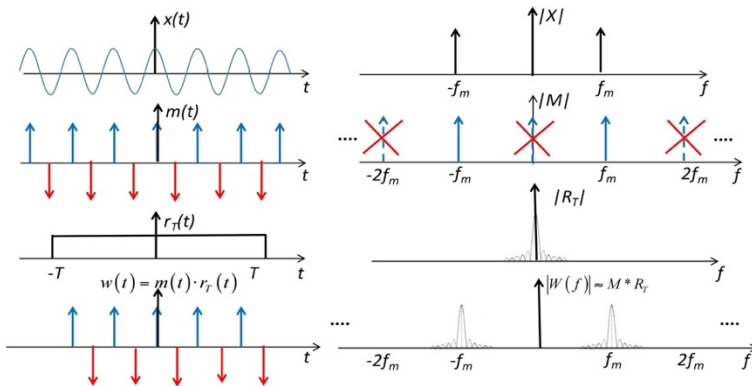
Figure 17.10: Time domain and frequency domain behavior of synchronous measurement with DC suppression by summing positive peak and subtracting negative peak samples

Comparing the filtering with the previous case:

- at f_m: band is useful, since we are collecting the signal (and some inevitable white noise around it)
- at DC: there is no band, so we are not collecting $\frac{1}{f}$ noise anymore
- at $3f_m, 5f_{m,...}$: residual harmful bands, as we are collecting white noise without acquiring any signal;

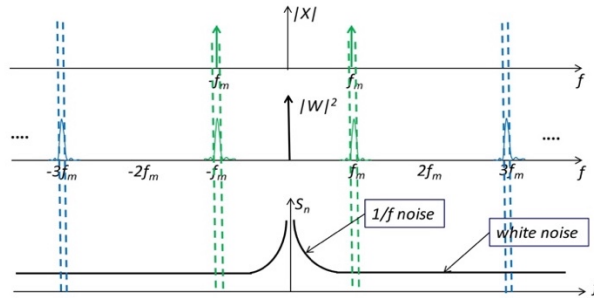Is it possible to further improve this kind of filter?

Figure 17.11: Noise collecting bands of synchronous measurement with DC suppression by summing positive peak and subtracting negative peak samples

17.3.4 Continuous Sinusoidal Weighting

With the Sample-Averaging (alternated peak samples) filtering scheme we have achieved the DC suppression; hence the next step is to get rid of the residual bands at odd multiples of f_m. The idea is to adopt continuous sinusoidal weighting instead of peak sampling. The approach is similar to the one we used for the optimum filter. This is a truly efficient filtering: there is just one narrow band at f_m
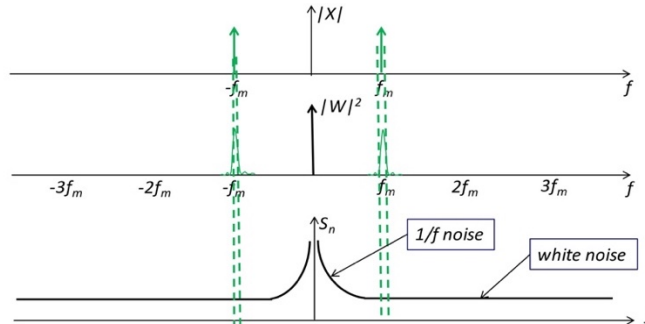
Figure 17.12: Noise collecting bands of Continuous Sinusoidal Weighting

Let's check if the filtering is optimized:

- at f_m band is useful, since we are collecting the signal (and some inevitable white noise around it)
- at DC: there is no band, so we are not collecting $\frac{1}{f}$ noise
- at $3f_m, 5f_{m,...}$: there are *no more* residual harmful bands, thus we are not collecting white noise where no signal is present;

In the following paragraph, the implementation of this optimized synchronous measurement is addressed.

17.4 Basic set-up for synchronous measurement with optimized noise filtering
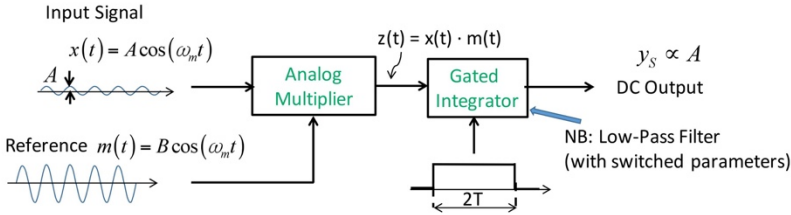

Figure 17.13: Basic Set-up for synchronous measurement with optimized noise filtering

This setup consists of two main blocks:

- an Analog Multiplier, which multiplies the input signal x(t) by the reference signal m(t) shifting the input signal down by $-f_m$ and up by $+f_m$
- a Gated Integrator, which gates the output of the analog multiplier;

Recalling that the weighting function in the time domain is $w(t) = m(t) \cdot r_t(t)$, it is straightforward that it results from the combined action of the multiplier and the GI.
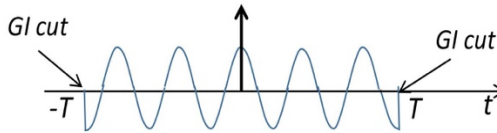
Figure 17.14: Weighting function of the optimized synchronous measurement in the time domain

In the frequency domain the weighting function's module is:

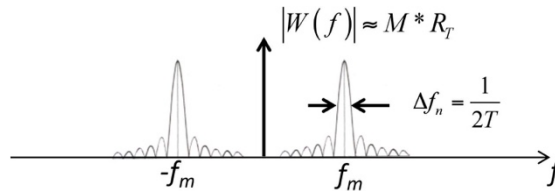**Figure 17.15:** Weighting function of the optimized synchronous measurement in the frequency domain

There are two remarks worth making:

- the reference input of the multiplier is a *standard* waveform, that does not depend on the input signal and is thus the same for every input signal
- the setup is a *linear* time-variant parameter filter

This linear time-variant filter composed by analog multiplier and gated integrator (low-pass filter with switched-parameter) has a weighting function similar to that of a tuned filter with constant-parameter, but has basic advantages over it:

- Center frequency f_m and width Δf_n are independently set
- The center frequency is set by the reference $m(t)$ and locked at the frequency f_m
- In cases where f_m has not a very stable value the filter band-center tracks it: the signal is thus kept in the admission band even if the width Δf_n is very narrow.
- The width $\Delta f_n = 1/2T$ is set by the GI, it is the (bilateral) passband of the GI
- Narrow Δf_n and high-quality factor Q can thus be easily obtained at any f_m

17.5 LOCK-IN amplifier and weighting function

With averaging performed by a gated integrator, like in the previous paragraph, the amplitude A can be measured only at discrete times (spaced by at least the averaging time 2T). However, by employing a constant-parameter low-pass filter instead of the GI, continuous monitoring of the slowly varying amplitude A(t) is obtained. Hence the setup becomes:
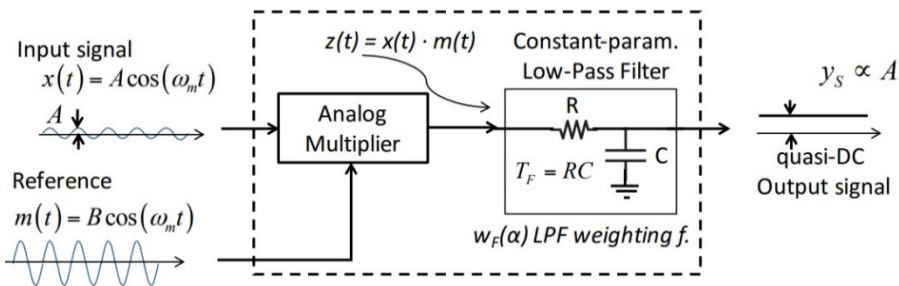
Figure 17.16: Lock-in amplifier scheme

The constant parameter LPF performs a running average of the output $z(t)$ of the output of the multiplicator. The output is continuously updated and tracks the slowly varying amplitude $A(t)$. Writing the output of our system we have

$$y(t) = \int_0^\infty z(\alpha)\, w_F(\alpha)d\alpha = \int_0^\infty x(\alpha)\, m(\alpha) w_F(\alpha)d\alpha$$

By comparison with the definition of the LIA weighting function $w_L(\alpha)$

$$y(t) = \int_0^\infty x(\alpha)\, w_L(\alpha)d\alpha$$

we can extract the weighting function $w_L(\alpha)$ of the LIA both in time and frequency domain

$$w_L(\alpha) = m(\alpha) \cdot w_F(\alpha) \quad \rightleftarrows \quad |W_L(f)| \cong |M| * |W_F|$$

The choice of the LPF's time constant on one side has to preserve the useful **signal bandwidth**, while on the other side it has to strongly attenuate the $2f_m$ component.
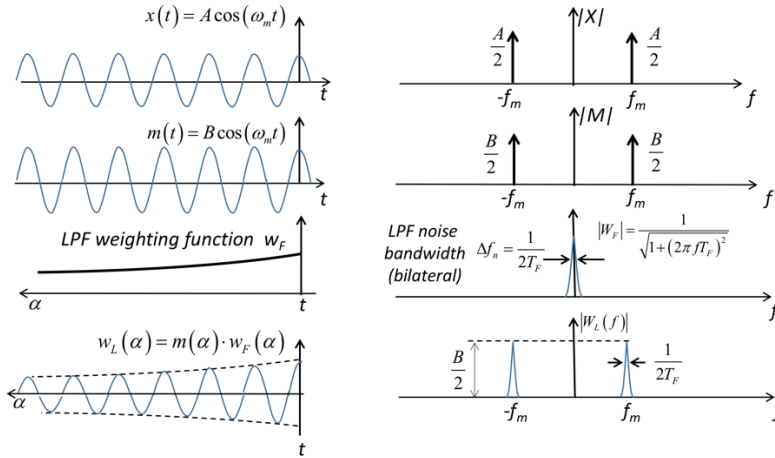
Figure 17.17: Lock-in amplifier weighting function in time domain and frequency domain

We can then calculate the S/N ratio of our example using a LIA. Starting from our modulated signal

$$x(t) = A\cos(2\pi f_m t)$$

and the reference

$$m(t) = B\cos(2\pi f_m t)$$

We find the output signal

$$y_S = 2\frac{A}{2} \cdot \frac{B}{2} = \frac{B}{2}A$$

While for the noise we can calculate it as

$$\overline{n_{yL}^2} = 2\left(\frac{B}{2}\right)^2 \cdot S_{Bb} \cdot \Delta f_n = \frac{B^2}{2} \cdot S_{Bb} \cdot \Delta f_n$$

So, the signal to noise ratio is

$$\left(\frac{S}{N}\right)_L = \frac{y_S}{\sqrt{n_{yL}^2}} = \frac{A}{\sqrt{2S_{Bb}\Delta f_n}}$$

17.6 LOCK-IN amplifier possible issues

In the previous section, we discussed the lock-in amplifier, a solution that can allow us to obtain a good signal to noise ratio also in case of strong 1/f noise thanks to the modulation principle. Let's analyse now two possible issues of this filter:

- Lock in amplifier noise enhancement
- Imperfect modulation
- Phase mismatch

Let's try to compare the case of DC signal filtered with LPF with the case of an AC signal filtered with LIA
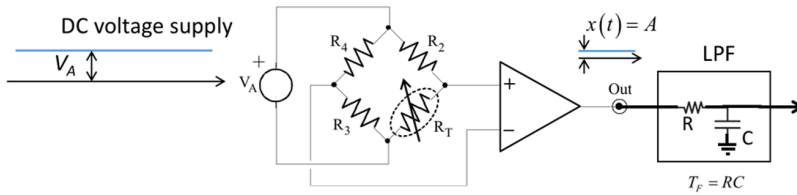
Figure 17.18: Lock-in amplifier with DC signal

Let us consider the set-up of the figure 17.18 (measurement with resistive sensor) now with DC supply voltage V_A equal to the amplitude of the previous AC supply. With a LPF equal to that employed in the previous LIA we obtain:

$$y = A$$

While for the noise

$$\overline{n_{yC}^2} = S_{nu} \cdot f_{Fn}$$

And so

$$\left(\frac{S}{N}\right)_C = \frac{y_C}{\sqrt{n_{yC}^2}} = \frac{A}{\sqrt{S_{nu}f_{Fn}}}$$

This S/N may look better by the factor $\sqrt{2}$ than the S/N obtained with the LIA, but is this conclusion true?

NO, such a conclusion is grossly wrong because the spectral density we are now considering is not the same one of LIA. Let's try to make it clearer showing the difference in the frequency domain. If we look at the figure 17.19 is easy to understand how the two spectral densities are totally different. Since we are integrating the spectral density on a relatively narrow band, we have to compare them locally. In the case of DC components, we are considering a narrow band around the zero frequency where the local value of the spectral density is much higher than the value at f_m.
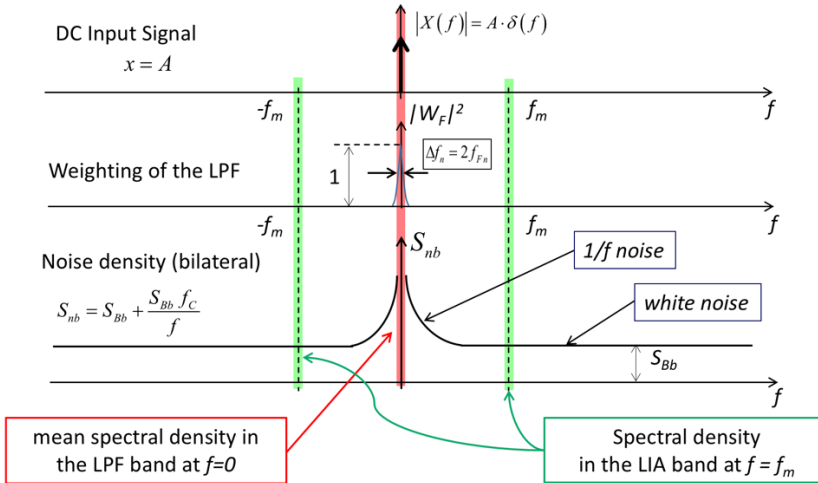
Figure 17.19: Comparison of noise spectral density for DC and AC components

17.6.1 Imperfect modulation

Ideally, the reference waveform should be a perfect sinusoid at frequency f_m with amplitude B_1. Actually, deviations from the ideal case can generate spurious harmonics at multiples $k f_m$ (k = 0, 1, 2 ...) with amplitudes B_k (small $B_k \ll B_1$ in case of small deviations). Moreover, effects equivalent to an imperfect reference waveform can be caused by non-ideal operation (non-linearity) of the multiplier. Since it is

$$|W_L(f)| \cong M(f) * W_F(f)$$

each spurious harmonic component of *M(f)* adds to the LIA weighting function W_L a spurious passband at frequency kf_m with amplitude B_k and shape given by the LPF. A spurious passband at f = 0 is particularly detrimental even with small $B_0 \ll B_1$ because it covers the high spectral density of 1/f noise, is so of fundamental importance to avoid any deviation from perfect balance of positive and negative areas of the reference that produces a DC component with associated passband at f = 0.

17.6.2 Phase mismatch

It is always possible to have a phase mismatch between the modulated signal and our reference. Let's try to understand the effect of this mismatch.

Starting from the definition of modulated signal and reference, we have

$$x(t) = A \cos(2\pi f_S t)$$

$$m(t) = B \cos(2\pi f_m t + \varphi_m)$$

By exploiting the a well-known trigonometric equation

$$\cos \alpha \cos \beta = \frac{1}{2}\cos(\alpha - \beta) + \frac{1}{2}\cos(\alpha + \beta)$$

We can easily obtain the result, since $f_S = f_m$

$$y(t) = x(t) \cdot m(t) = \frac{AB}{2}\cos[\varphi_m] + \frac{AB}{2}\cos[2\pi(2f_S)t + \varphi_m]$$

This means that at the LPF's output there is still a waveform proportional to the slowly varying amplitude of the input sinusoid, but with a modified gain with respect to the previous case. Comparing this situation with the ideal one we have an attenuation of the signal, and so of the S/N of a value $\cos[\varphi_m]$.

CHAPTER 17

Real Lock-In Amplifiers

In this chapter, we will make the transition from ideal to real LIA introducing the problems, trade-offs, and discussing solutions to improve the measurement with this approach. We will study some important cases of LIA operating with square wave reference and sinusoidal or square-wave modulation signal, and then we will compare the results and show all their pros and cons.

18.1 Introduction

In the previous chapter we introduced the Lock-In Amplifier, showing that it is theoretically a simple and very effective solution to extract very small signals potentially buried into very large low-frequency noise. The idea was to modulate the signal at frequencies where the contribution of low-frequency noise like 1/f is negligible. In some cases, we have to handle the modulation, while in some other cases the signal is already modulated. If the reference signal is available, it is possible to make the multiplication of the signal with the reference, and then use a low-pass filter to extract the amplitude of the non-modulated signal.
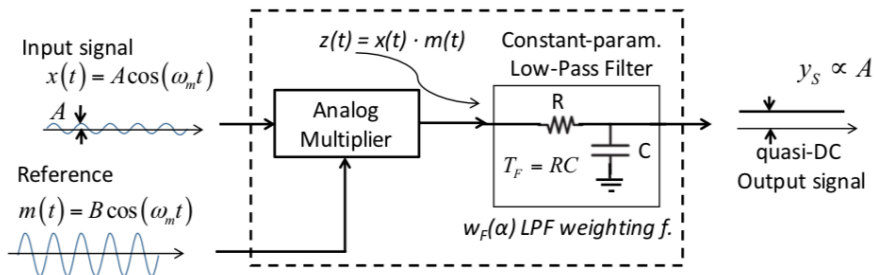
Figure 18.1: LIA with sinusoidal signal and sinusoidal reference employing an analog multiplier.

We saw that by multiplying two sinusoidal signals, one at frequency ω_1 and the other at frequency ω_2, from the trigonometry formulas, we obtain a frequency component at the difference $\omega_1 - \omega_2$ and one at the sum $\omega_1 + \omega_2$.

$$A cos(\omega_1 t) \cdot B cos(\omega_2 t) = \frac{1}{2} AB cos((\omega_1 - \omega_2)t) + \frac{1}{2} AB cos((\omega_1 + \omega_2)t)$$

and so, with $\omega_1 = \omega_2 = \omega_m$:

$$Acos(\omega_1 t) \cdot Bcos(\omega_2 t) = \tfrac{1}{2}ABcos(0) + \tfrac{1}{2}ABcos(2\omega t)$$

We obtain a DC component and a component at high frequency. The purpose of the LPF is to cut the high frequency component and to have at the output just the amplitude of the signal. This is the concept of the Lock-In amplifier.

18.2 From principles to real LIA instruments

To understand all the problems and trade-offs of a real LIA we will do a step-by-step analysis. The modulated input signal is converted by the LIA in a slow demodulated signal, with components from DC to a fairly low frequency limit. This signal must be supplied to a meter circuit that measures its amplitude, i.e. nowadays ordinarily an ADC. The LIA output signal must have scale adequate for the ADC, whereas the LIA input signal is very small: therefore, the LIA must provide high overall gain for the signal.

The idea is to introduce a post-amplifier to increase the signal amplitude just after the LPF. The gain of the post-amplifier doesn't change the SNR because both signal and noise are amplified by the same amount. The bandwidth of the post-amplifier has to be small because the signal at the output of the LPF is DC (normally we don't have a DC signal but a signal slowly varying in time). The real problem is that the post-amplifier introduces low-frequency noise like 1/f noise or drift of the baseline, that is acting directly on our low-frequency signal since it is not filtered by the LIA. One might think that having a high SNR, there is some margin for the added noise. This is not the case because we should not forget that beside having a high SNR, the signal amplitude is very small. So even a small noise introduced by the post-amplifier deteriorates the SNR.

If the demodulated signal is very small, comparable or lower than the baseline drift and noise of the post-amplifier referred to its input, the measurement will be spoiled. The idea is so to increase the signal at previous stages by a pre-amplifier, so the noise introduced by the post-amplifier has less effect. The pre-amplifier is placed at the signal channel, and it obviously acts both on the signal and on the noise. It processes the modulated input signals, hence it is an AC coupled amplifier, either wide-band type including the modulation frequency f_m or narrow-band tuned to f_m, and it receives a signal accompanied by high noise, because it operates before the LIA. Differently from the post-amplifier it may have baseline drift and low-frequency noise, but their role is minor because they are filtered by the LIA (and by the AC-coupled amplifier itself).

It is important to point out an important constraint: in order to obtain the foreseen improvement of S/N, the processing of signal and noise in the LIA must be accurately linear. Deviations from linearity produce detrimental effects (self-modulation of the noise, generation of spurious harmonics, etc.), which irrevocably alter the measure

and degrade the LIA performance, as we saw in the previous chapter. The signal and noise must remain well within the linear dynamic range in every stage involved, particularly in the multiplier (and in the preamplifier).

When a wide-band preamplifier is employed to raise the level of a very small input signal, a problem arises with very small input S/N<<1. The gain required for the signal also amplifies a noise which is much higher than the signal, hence it brings this amplified noise out of the linear dynamic range of the multiplier.

In such cases, for exploiting the required gain it is necessary to reduce the noise received by the preamplifier with a pre-filter. Adequate reduction of the LIA input noise is obtained in many cases with prefilter passband much wider than that of the LIA.

Such a prefiltering would be a useless nonsense in an ideal apparatus, but in real cases it is a necessary feature for avoiding nonlinearity in intermediate stages.

Preamplifiers that incorporate prefiltering are currently available from LIA manufacturers; they are called tuned preamplifiers or selective preamplifiers.

18.3 Lock-in Amplifier with Square-wave Reference

In order to limit some of the overmentioned problem we could remove the analog multiplier replacing it with just switches and amplifiers implementing a modulation with square wave reference *m(t)*.

The noise referred to the input, the linearity and the dynamic range of these circuits are remarkably better than those of analog multiplier circuits (even high-performance types) because they are limited just by the performance of amplifiers and switch-devices.

Therefore, switched linear circuit configurations are often employed as demodulator stage in LIAs in order to avoid the limitations of analog multipliers.
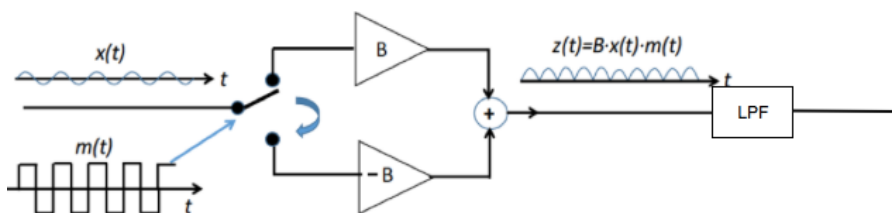

Figure 18.2: LIA with switch-based multiplier.

The weighting function $w_L(\alpha)$ of a LIA is the multiplication of reference waveform $m(\alpha)$ (periodic at frequency f_m) and weighting function $w_F(\alpha)$ of the LPF

$$w_L(\alpha) = m(\alpha) \cdot w_F(\alpha)$$

In frequency domain this corresponds to the convolution of the F-transforms

$$W_L(f) = M(f) * W_F(f)$$

Since the transform $M(f)$ of a periodic $m(\alpha)$ is composed by lines at f_m (fundamental) and integer multiple frequencies (harmonics), and $W_F(f)$ of the LPF has bandwidth much smaller than f_m, the result of the convolution of $W_F(f)$ by any line of $M(f)$ does not overlap the result by any other line (with very good approximation).
We conclude that the $W_L(f)$ is a set of replicas of $W_F(f)$ centered on each line of $M(f)$, multiplied by the line-weight and phase-shifted by the line-phase. So, with very good approximation, the module diagram can thus be obtained simply as

$$W_L(f) \cong |M(f)| * |W_F(f)|$$
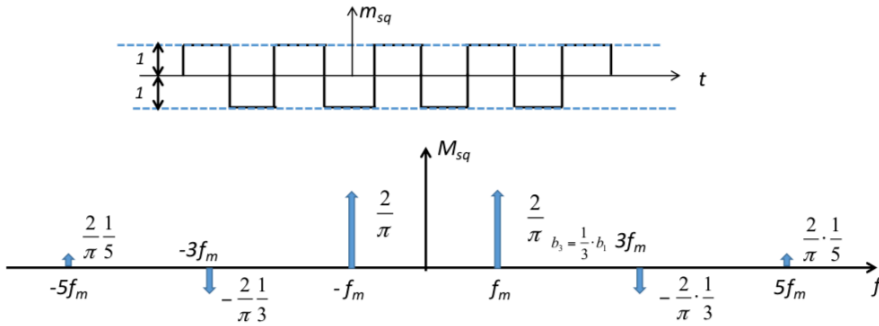
Figure 18.3: Square wave signal with zero mean value and module of the square wave spectrum.

With

$$M_{sq}(f) = \sum_{k=0}^{\infty} \frac{b_{2k+1}}{2} \left[\delta(f - f_{2k+1}) + \delta(f + f_{2k+1}) \right]$$

$$b_{2k+1} = \frac{(-1)^k}{(2k+1)} \cdot \frac{4}{\pi}$$

Having a perfectly symmetrical square wave reference the DC component is completely removed, but this is an ideal case. In real cases, it might happen that the square wave value is not perfectly symmetrical, and so a DC component is still present. During the implementation of the switch driver, for example, we might have some error in the duty-cycle so the negative and positive area might not be exactly the same. But this small DC value is very problematic since it opens a window in frequency where high 1/f noise is collected.
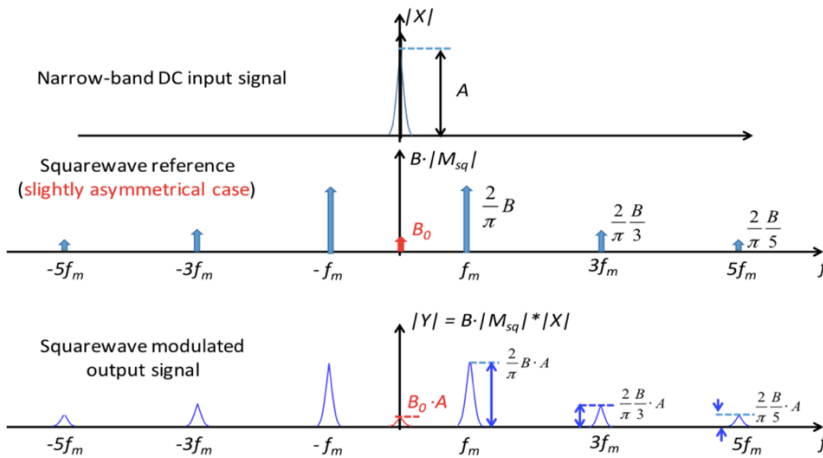
Figure 18.4: LIA weighting function with spurious band at f=0.

18.4 Sinusoidal signal through LIA with square wave reference

Let's apply now our square wave reference to the signal. The effect of the square wave reference with the same frequency and phase on the sinusoidal signal, is the full sinewave rectification. The amplitude of the rectified signal is the product of the amplitude of the sinusoid and the reference. Then this rectified signal passes through the low-pass filter which does the average in time.

Calculation of the average value of the full-rectified sinewave

$$y(t) = \frac{1}{\pi} \int_0^\pi AB\sin(t)dt = \frac{1}{\pi} AB[-\cos(t)]_0^\pi = \frac{2}{\pi} AB$$

Figure 18.5: Input signal and LIA weighting function in frequency domain.

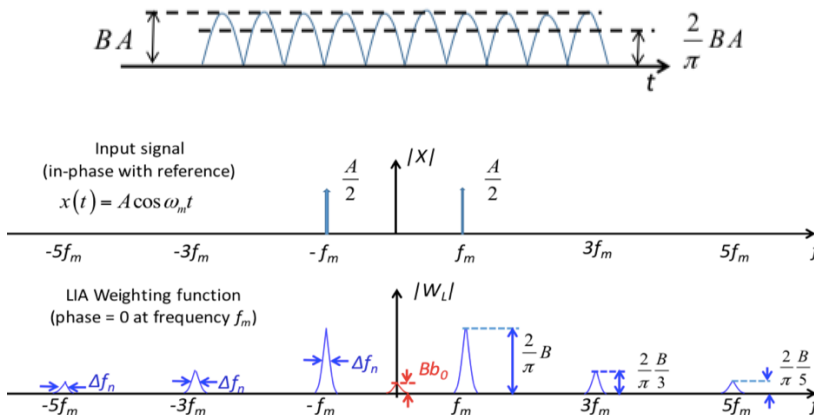In frequency domain, it is immediate to obtain the output signal, it is obtained by calculating the area of the product between the F-transform of the sinusoidal and the square wave. The two deltas at $\pm f_m$ of the sinusoidal multiply with the value in $\pm f_m$ of the squarewave spectrum. Evidently:

$$s_y(t) = 2 \cdot \frac{A}{2} \cdot \frac{2}{\pi} \cdot B = \frac{2}{\pi} AB$$

What we notice in the case of sinusoidal signal and square wave reference, is the fact that the additional windows of the reference don't collect useful signal, but they collect noise.

18.5 Noise through LIA with perfect Square wave Reference

To compute the noise contribution, we make the integral of the product of the noise spectral density with the module-square of the weighting function.

$$\overline{n_{yL}^2} = \int_{-\infty}^{\infty} S_{nb}(f) \cdot |W_L(f)|^2 \, df$$
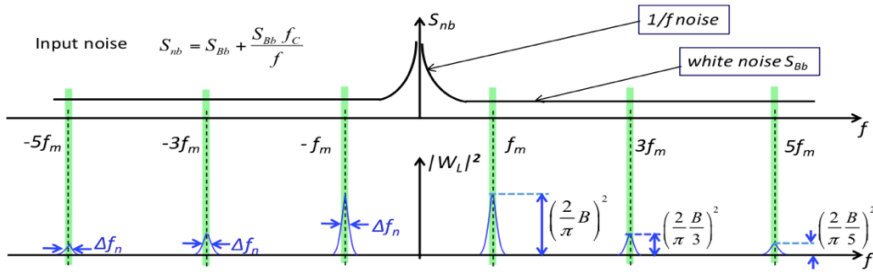
Figure 18.6: Every window of the weighting function in frequency domain collects the input white noise with the proper amplitude.

Every window collects white noise, giving the proper amplitude. The sum of each contribution gives the total noise power

$$\overline{n_{yL}^2} = 2 \cdot S_{Bb} \cdot \Delta f_n \cdot (\frac{2}{\pi} B)^2 \cdot [1 + \frac{1}{3^2} + \frac{1}{5^2} + \dots]$$

where Δf is the window equivalent bandwidth. We can then calculate the enhanced noise due also to the higher passbands at the harmonic frequencies

$$[1 + \frac{1}{3^2} + \frac{1}{5^2} + \dots] = \frac{\pi^2}{8} \cong (1{,}11)^2$$

Using this calculation, we obtain

$$\overline{n_{yL}^2} = 2 \cdot S_{Bb} \cdot \Delta f_n \cdot (\frac{2}{\pi} B)^2 \cdot \frac{\pi^2}{8} = B^2 \cdot S_{Bb} \cdot \Delta f_n$$

So, we can calculate the signal to noise ratio

$$\left(\frac{S}{N}\right)_{sin-sqw} = \frac{s_y}{\sqrt{n_{yL}^2}} = \frac{A}{\frac{\pi}{2}\sqrt{S_{Bb}\Delta f_n}}$$

which in comparison to the result obtained with sinusoidal reference

$$\left(\frac{S}{N}\right)_{sin-sin} = \frac{s_y}{\sqrt{n_{yL}^2}} = \frac{A}{\sqrt{2S_{Bb}\Delta f_n}}$$

$$\left(\frac{S}{N}\right)_{L,sqw} = \frac{\sqrt{2}}{\frac{\pi}{2}}\left(\frac{S}{N}\right)_{L,sin} \cong \frac{1}{1,11}\left(\frac{S}{N}\right)_{L,sin}$$

Therefore, the S/N is reduced with respect to the case of sinusoidal reference, but the reduction is moderate.

18.6 Square wave signal through LIA with square wave reference

Another case which is often met in real life is to have a square wave signal at the input of the LIA. The sinusoidal modulation, in general, could be not easy to be implemented. A simpler way is to use a square wave to modulate the sensor and then use a square wave as a reference.

Let's see how the Lock-In Amplifier performs in the case of square wave signal with square wave reference with the same frequency and phase

The output signal is then

$$s_y(t) = \int_{-\infty}^{\infty} X(f)W(-f)df = 2\left[\frac{2}{\pi}A\frac{2}{\pi}B\right]\left[1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots\right] \cong 2\left[\frac{2}{\pi}A\frac{2}{\pi}B\right]\frac{\pi^2}{8} = AB$$

That is easily verified in time, the output is $A \cdot B$ when the carrier is positive and $(-A) \cdot (-B)$ when the reference is negative.

The noise is the same we calculated in the previous paragraph since it does not depend on the signal.

So, we obtain the signal to noise ratio

$$\left(\frac{S}{N}\right)_{sq-sq} = \frac{s_y}{\sqrt{n_{yL}^2}} = \frac{A}{\sqrt{S_{Bb}\Delta f_n}}$$
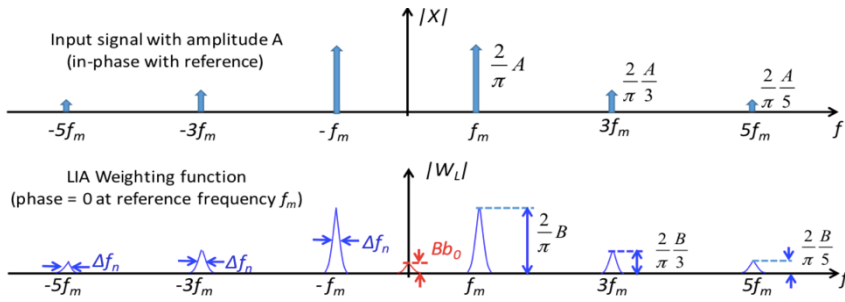
Figure 18.7: Input signal and LIA weighting function with square wave reference and signal

However, for equal amplitude A the square wave signal has double power and correspondingly higher S/N compared to the sinusoidal modulation.

In the table 18.8 we can compare the different combinations of modulated signal and reference.

	SINUSOIDAL Reference	SQUAREWAVE Reference
SINUSOIDAL Signal amplitude A power $P = \frac{A^2}{2}$ A_{min} minimum measurable amplitude (at S/N=1)	$\frac{S}{N} = \frac{A}{\sqrt{2}\sqrt{S_{Bb}\Delta f_n}} = \frac{\sqrt{P}}{\sqrt{S_{Bb}\Delta f_n}}$	$\frac{S}{N} = \frac{A}{\frac{\pi}{2}\sqrt{S_{Bb}\Delta f_n}}$ $= \frac{\sqrt{P}}{\frac{\pi}{2\sqrt{2}}\sqrt{S_{Bb}\Delta f_n}}$
	$A_{min} = \sqrt{2}\sqrt{S_{Bb}\Delta f_n} = 1{,}41\sqrt{S_{Bb}\Delta f_n}$	$A_{min} = \frac{\pi}{2}\sqrt{S_{Bb}\Delta f_n}$ $= 1{,}57\sqrt{S_{Bb}\Delta f_n}$
SQUAREWAVE Signal amplitude A power $P = A^2$ A_{min} minimum measurable amplitude (at S/N=1)	$\frac{S}{N} = \frac{A}{\frac{\pi}{2\sqrt{2}}\sqrt{S_{Bb}\Delta f_n}} = \frac{\sqrt{P}}{\frac{\pi}{2\sqrt{2}}\sqrt{S_{Bb}\Delta f_n}}$	$\frac{S}{N} = \frac{A}{\sqrt{S_{Bb}\Delta f_n}} = \frac{\sqrt{P}}{\sqrt{S_{Bb}\Delta f_n}}$
	$A_{min} = \frac{\pi}{2\sqrt{2}}\sqrt{S_{Bb}\Delta f_n} = 1{,}11\sqrt{S_{Bb}\Delta f_n}$	$A_{min} = \sqrt{S_{Bb}\Delta f_n}$

Table 18.8: Summary of different cases for reference and signal modulation

Photodetector Fundamentals

Photodetectors are pervasively invading our everyday life: they allow the measurement of optical signals that are very useful in many fields, from consumer to transport and industrial fields just to name a few. This chapter introduces some types of existing sensors, their working principle, their applications and it provides technical details on vacuum tubes.

19.1 Introduction

In the field of signal theory and detection, an important class of devices is represented by photodetectors. These devices essentially deal with detecting electromagnetic radiation and converting its characteristics into a measurable electrical signal. This is possible through a series of physical mechanisms that we will analyze in detail, but, first of all, it is necessary a general treatment of the electromagnetic radiation mediating particle: the photon.

19.2 Photon

The photon is an elementary particle mediating one of the 4 fundamental interactions: electromagnetism. It belongs to the group of bosons, that is the class of particles that obey the Bose-Einstein statistics (whole spin particles).
The photon has mass and charge equal to zero and presents both undulating and corpuscular behaviors, a phenomenon called "wave-particle duality" (observable also for electrons). In vacuum, the photon travels at the speed of light, equal to:

$$c_0 = 299.792 \times 10^8 \; \frac{m}{s}$$

and it has unlimited range of action. This means that it can travel indefinitely until it is absorbed (which is why it is still possible to detect the photons emitted during the Big Bang, what is called background fossil radiation). Since photons exhibit a wave behavior, it is possible to define a wavelength λ and a frequency v and the two are linked by the relation:

$$c = \lambda v$$

Where c is the speed of light in a generic medium equal to $c = \frac{c_0}{n}$ (n refractive index).

The set of radiation of the various photons at different wavelengths constitutes the electromagnetic spectrum, which contains the visible spectrum (what we simply call light).

The wavelength of "visible" photons extends from about 390nm up to 700nm, from violet to red (in terms of frequency from 770THz to 430THz).

Range (wavelength)	Region
$\lambda < 400$nm	Ultraviolet
400nm<λ<750nm	Visible
750nm<λ<3000nm	Near infrared (NIR)
3000nm<λ<30000nm	Mid infrared (MIR)
30000<λ	Far infrared (FIR)
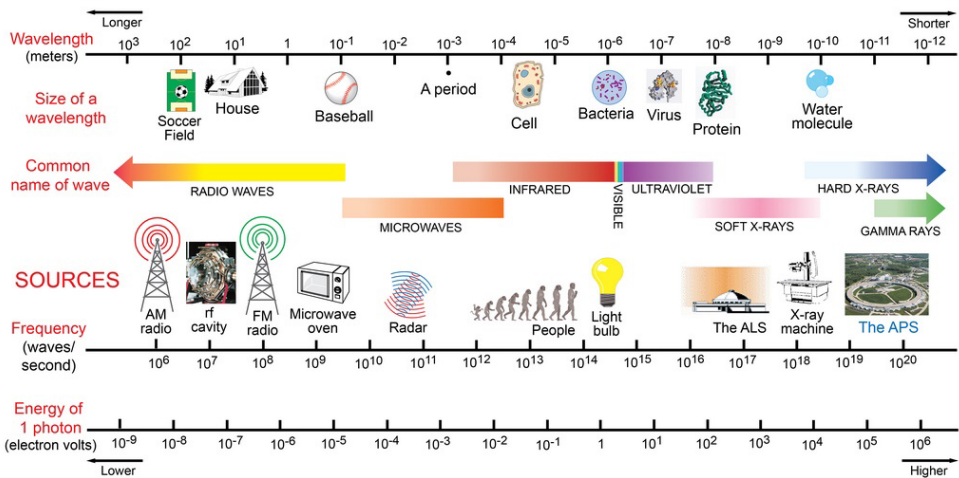
Figure 19.1: Electromagnetic spectrum: it's clear that the visible part of the spectrum is only a small portion of the whole spectrum.

Moreover, photons carry energy equal to:

$$E = h\upsilon \quad [\text{J}]$$

Where h is the Planck constant: $h = 7,6 \times 10^{-34} J \cdot s$

It is precisely this energy that causes the interaction of photons with matter. We can express this energy in electronvolt (eV) instead of Joule, a more versatile unit of measure for a discussion in the field of electronic devices (the semiconductor gaps are of the order of some eV that is why it is convenient to use this unit of measurement).

So, we can divide this energy, expressed in Joule, by the value of the charge of an electron obtaining:

$$E = \frac{h\nu}{q} = \frac{hc}{q\lambda} \qquad [eV]$$

And finally, in a more easy to remember formula:

$$V_p = \frac{hc}{q\lambda} = \frac{1,24\mu m}{\lambda} \qquad [eV]$$

From which we can derive:

400nm $< \lambda <$ 750nm   visible range $\rightarrow$ 3,10eV $> V_p >$ 1,65eV
750nm $< \lambda <$ 3μm   near infrared range $\rightarrow$ 1,65eV $> V_p >$ 0,41eV

19.3 Reflection of photons at interface

At this point, we have in mind the basic concepts regarding photons and their interaction with matter; now we can discuss in details phenomena that occur when radiation hits a surface (which could be the surface of a photodetector). For a simple analysis, we will consider a normal-incident beam hitting the surface; this simple approach still preserves the general value of the concepts here addressed.

The aforementioned refractive index of a material is defined as follow:

$$n = \frac{c_0}{v} = \frac{\dfrac{1}{\sqrt{\varepsilon_0\mu_0}}}{\dfrac{1}{\sqrt{\varepsilon_0\varepsilon_r\mu_0}}} = \sqrt{\varepsilon_r}$$

where c_0 is the speed of light in vacuum and v is the phase velocity of light in the medium.

The refractive index describes how fast light propagates in a certain material; it is referred to the speed of light in vacuum: for example, the refractive index of water is 1.333, meaning that light travels 1.333 times faster in vacuum than in water.

This index is very important because it allows us to define what happens to the radiation when it approaches a discontinuity (i.e. the interface between two different media). In particular, given two materials and a certain incident optical power it's possible to demonstrate that the fraction of reflected power is equal to:

$$R = \left(\frac{n_1 - n_2}{n_1 + n_2}\right)^2 = \frac{P_{reflected}}{P_{incident}}$$

Where n_1 is the refractive index of the material where the light is travelling and n_2 is the one of the material that the light finds on its path. For Silicon n≈3.4: with a direct interface with air (n≈1), we have a reflection coefficient R≈0.3. It has to be said that n is actually a function of the wavelength.

It's clear that the reflection of light due to the interface between two different materials is a big issue in many applications where we want to detect as much light, and so as much signal, as possible, so we need to avoid any loss.

From the conservation of energy, we can compute the fraction of light that is transmitted in the second material, which is equal to:
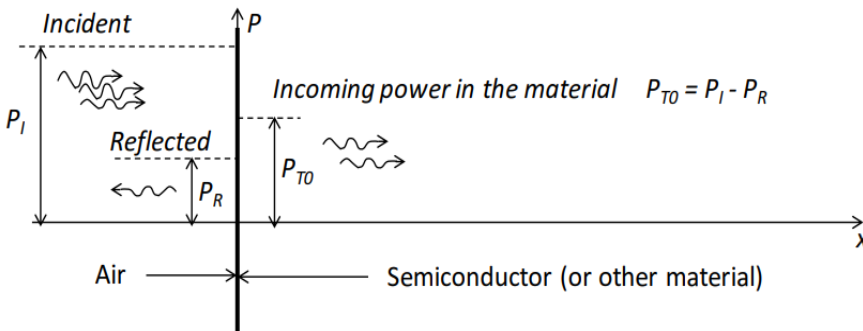
$$T = 1 - R = \frac{P_{t0}}{P_i}$$

Figure 19.2 overall situation of a wave that meet an interface:

Later in this chapter, we will discuss a possible solution to reflection issues.

19.4 Photon absorption in materials

After discussing what happens when the light faces a discontinuity of materials on its path, we will now analyze what happens when the radiation travels through a medium. Absorption can occur when a photon interacts with the matter of the medium it is travelling in; the interaction depends on both the characteristics of the material that composes the medium and the photon wavelength.

Let's consider a flux of photons sent through an absorbing medium; when photons are absorbed, the amount of power travelling is reduced. We can write that the variation of travelling power (and so the absorbed power) is proportional to the travelled depth:

$$\frac{\Delta P}{P} \propto \Delta x$$

It's intuitive that the more the light travels in deep in the material, the more photons are absorbed. Ideally, with a medium of infinite thickness, all photons would be absorbed.

We can now introduce a factor α as proportionality coefficient (optical absorption coefficient); we also introduce a minus sign in the equation because we understood that the travelling power decreases as the travelled path increases:

$$\frac{\Delta P}{P} = -\alpha \Delta x \implies \frac{\Delta P}{\Delta x} = -\alpha P$$

$$passing\ to\ the\ limit \implies \frac{dP}{dx} = -\alpha P$$

And solving this simple differential equation we get:

$$P_t(x) = P_{t0}e^{-\alpha x} = P_{t0}e^{-\frac{x}{L_a}}$$

Where P_{t0} is the optical power actually transmitted in the material right after overcoming the interface and L_a is the *optical absorption depth*.
This equation allows us to compute how much remaining power is travelling through the medium at a certain distance x from the surface. We can derive how much power is absorbed by the medium from 0 (interface) to x:

$$P_a(x) = P_{t0} - P_t(x) = P_{t0}(1 - e^{-\alpha x}) = P_{t0}(1 - e^{-\frac{x}{L_a}})$$

The results are summarized in the following plot:
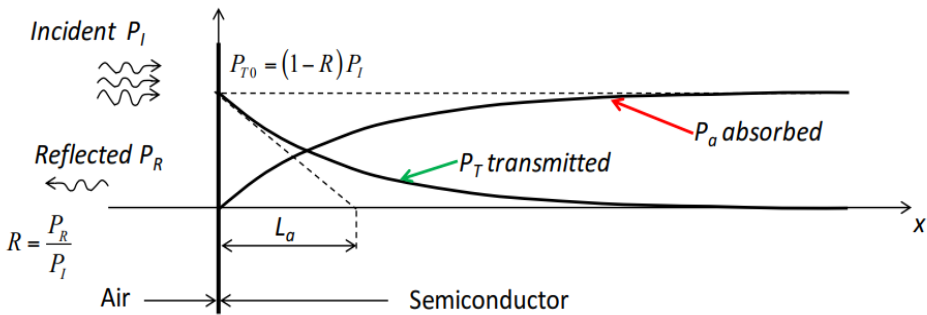
Figure 19.3 absorbed and transmitted power

We have to underline that the optical absorption of materials strongly depends on the wavelength of the radiation. To have an idea, in Silicon we have that the absorption depth at λ=400nm (violet) is ≈100nm, at λ=800nm (red) is ≈10μm, as can be seen in the following plots.
This means that after 5μm, red light is attenuated less than half (60% is still there) while the violet one is attenuated by a factor 1200 (just 0.079% of the incoming violet light is still travelling, it was almost all absorbed!).

Figure 19.4: Absorption coefficient of intrinsic silicon

Figure 19.5: Absorption depth in silicon

19.5 Thermal Photodetector Principles
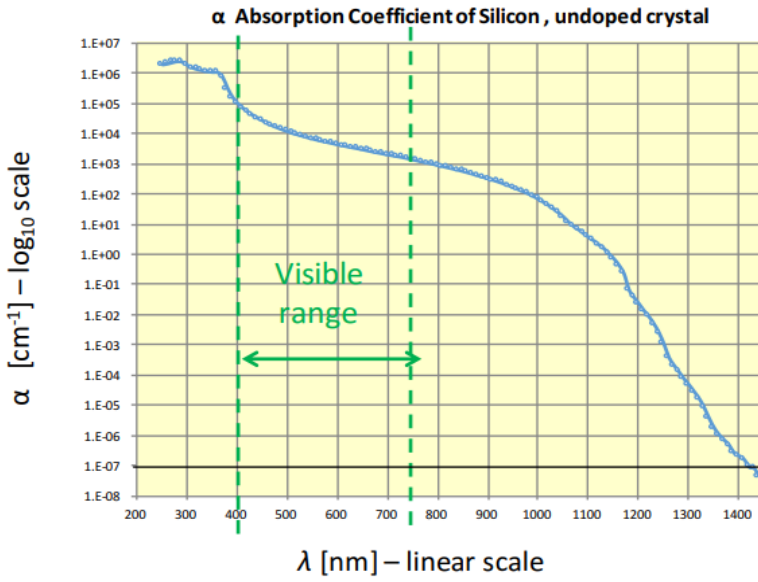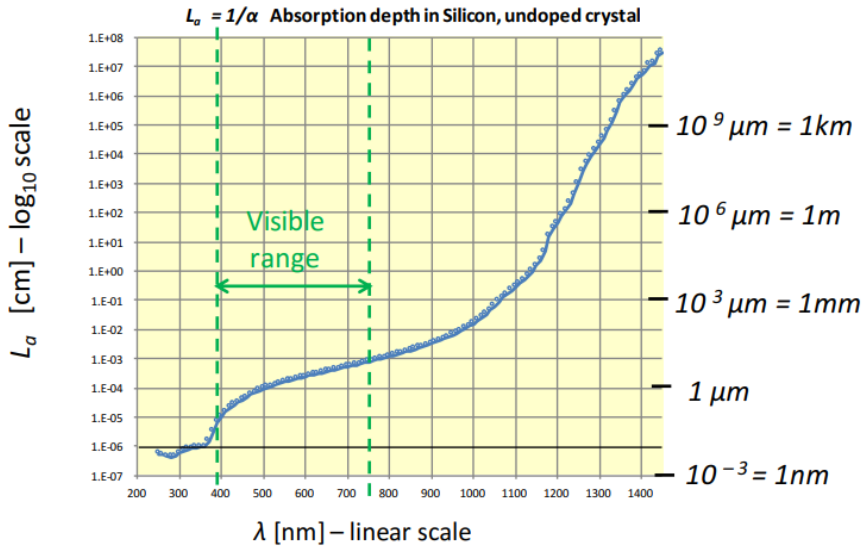
One of the physical principles that can be used to detect light signals is the heating of a medium caused by the absorption of optical power. This approach relies on the measurement of the relative ΔT temperature variation (in a direct or indirect way). Sensors that exploit this principle are called "thermal sensors". One of the advantages of these sensors is that they are sensitive to a very wide light spectrum since it is enough that the photon is absorbed to be detected. Unfortunately, these sensors have a very low sensitivity: to have a significant ΔT it is necessary that many photons are absorbed; for example, if we want to increase the temperature of a $\Delta T = 0.1K$ of a drop of water of 1mm, we need 10^{15} photons with a wavelength $\lambda = 475nm$ (light blue). Moreover, the dynamic behavior of these sensors is inherently slow: they require a long time interval to go from 10% to 90% of the final value; therefore, thermal photodetectors are used for stationary light radiation measurements.
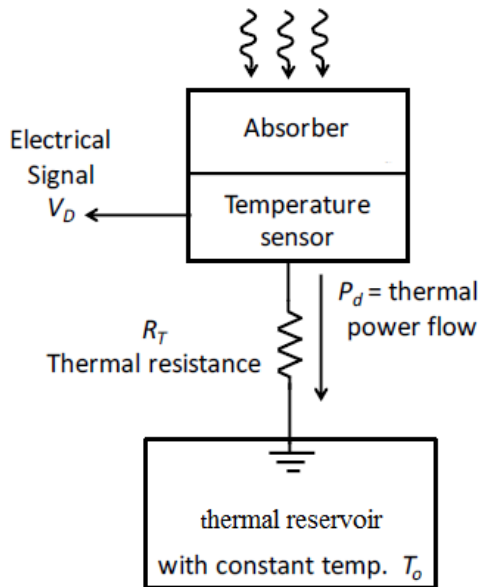
Figure 19.6: Conceptual scheme of a thermal photosensor

Regardless of the physical principle that is used to transform the ΔT of the temperature sensor into an electrical signal, it is possible to create a model for the behavior of any thermal photosensor. The conceptual scheme, shown in the figure 19.6, consists of a thin absorber layer (black body) that is hit by light radiation; the absorbed power leads to an overall increase in temperature of the absorber, which features a specific heat capacity c_a and a mass m_a; the absorber heats the substrate with respect to the reference temperature T_0 of the thermal reservoir, to which it is linked by a thermal resistance R_T. ΔT is converted into a proportional electrical signal

by means of a temperature sensor; it may be useful to find the transfer function between the optical power P_P and the temperature variation of the sensor ΔT.

We can define the optical power as the number of photons that arrive in one second n_p and the energy carried by each photon E_p

$$P_p = n_p \cdot E_p = n_p \cdot qV_p$$

At the same way we can define the substrate heat capacity C_a as the ratio between the heat ΔQ added to (or removed from) the thermal reservoir and the resulting temperature change ΔT

$$C_a = c_a \cdot m_a = \frac{\Delta Q}{\Delta T}$$

The thermal resistance quantifies how much an object or material resists a heat flow. It is defined as the ratio between the temperature variation and the absorbed power:

$$R_T = \frac{\Delta T}{P_d}$$

Using the principle of conservation of energy, it is possible to define an energy balance equation; the optical energy must be equal to the sum of the energy accumulated in the substrate and the exchanged energy of the resistance with the thermal reservoir.

$$P_p \cdot dt = \frac{\Delta T}{R_T} dt + C_a d(\Delta T)$$

$$\frac{d(\Delta T)}{dt} = \frac{P_p}{C_a} - \frac{\Delta T}{R_T C_a}$$

Going into the Laplace domain and remembering that the time signals are ΔT and P_p, we obtain:

$$s\Delta T = \frac{P_P}{C_a} - \frac{\Delta T}{R_T C_a}; \quad \rightarrow \quad \Delta T = P_p R_T \frac{1}{1 + sR_T C_a}$$

The dynamic response is a single-pole low-pass filter with characteristic time constant $\tau_a = R_T C_a$ and $f_t = \frac{1}{2\pi R_T C_a}$, whose bandlimit decreases with R_T. That highlights how the response (ΔT) of the system to an input (P_p) is intrinsically slow. The steady state response ($\Delta T = P_p R_T$ obtained with $s = 0$) increases as the thermal resistance R_T is increased.

It is useful to find an electrical equivalent to the above model, in which the substrate, with a thermal capacity C_a, is replaced by an electrical capacity in parallel with a resistance R_T, that models the thermal resistance, linking the thermal reservoir (the electrical ground) and the substrate. The optical input power can be seen as a current generator. Between gain and band of the system there is therefore an opposite

dependence on R_T. To overcome this problem, you can act on the substrate, choosing one with a smaller c_a, or reducing the mass m_a, that is using smaller devices.
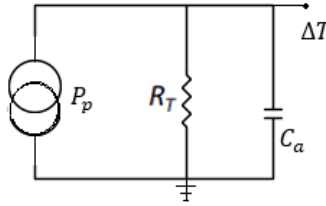

Figure 19.7: Equivalent electrical circuit for a generic thermal photosensor

Thermal detectors typically transduce the optical power P_P in an electrical output signal V_D of the temperature.
The basic quantitative characterization of the performance of a photodetector is given by the *"Radiant Sensitivity"* (also called Spectral Responsivity) S_D, defined as:

$$S_D = \frac{output\ voltage\ [Volt]}{optical\ power\ on\ the\ detector\ sensitive\ area\ [Watt]}$$

For a given optical power absorbed by the detector, the temperature variation does not depend on the wavelength of the radiation λ. Therefore, uniform S_D would be obtained at all λ if the reflection and absorption were constant, independent of λ. Obviously, λ influences the amount of radiation absorbed and reflected by the active layer. The variations in absorption and reflection as a function of λ are kept at moderate levels thanks to continuous research in this sector, achieving a fairly uniform S_D over fairly wide wavelength ranges, extended well into the infrared spectral region.

19.6 Quantum Photodetector Principles

A completely different principle for detecting an electromagnetic signal exploits photoelectric effects to directly generate a current inside the sensor. The energy of the absorbed photons is used to generate carriers inside the device, thus obtaining a current variation. Sensors that exploit this principle are called *Quantum Photodetectors* or *Photon Detectors* and can be vacuum tubes or semiconductor devices. The basic quantitative characterization of the performance of quantum photodetectors is given by the *Quantum Detection Efficiency* (or Photon Detection Efficiency) η_D defined as

$$\eta_D = \frac{\text{number of photogenerated electrons}}{\text{number of photons reaching the detector}} = \frac{N_e}{N_p}$$

We have also seen that a figure of merit of the performance of any kind of photodetector is the *Radiant Sensitivity S_D*, that allows us to compare very different kinds of photodetectors. In this case S_D is equal to:

$$S_D = \frac{\text{electrical output current [Ampere]}}{\text{optical power on the detector sensitive area [Watt]}} = \frac{I_D}{P_L}$$

It is easy to find a link between these two quantities. Photons of wavelength λ, arriving with a rate n_p, have an optical power P_L

$$P_L = n_p \cdot h\upsilon$$

While the electrons (or electrons-holes pairs) photogenerated in the detector with a rate n_e produce a current

$$I_D = q \cdot n_e$$

So, the Radiant Sensitivity is

$$S_D = \frac{I_D}{P_L} = \frac{n_e}{n_p} \cdot \frac{q}{h\upsilon} = \frac{n_e}{n_p} \cdot \frac{q \cdot \lambda}{hc}$$

And since the ratio between the number of carriers photogenerated in one second and the number of photons collected by the detector in one second is η_D, we can write

$$S_D = \eta_D \cdot \frac{q \cdot \lambda}{hc} = \eta_D \cdot \frac{\lambda[in\ \mu m]}{1.24}$$

We see that the Radiant Sensitivity of a quantum detectors intrinsically depends on the wavelength λ, even with constant quantum efficiency η_D. This occurs because a given optical power P_L corresponds to a different photon rate n_p at different wavelengths λ.

19.7 Photo-Tubes

In vacuum tubes, the photoelectric effect is used to emit an electron, also called photoelectron, from a metal or a semiconductor. We have seen that photons have an energy depending on the frequency of light $E = h\upsilon$; if the energy of the photon is equal or greater than the energy necessary to promote an electron either from the conduction band (in the case of metals) or from the band of valence (in the case of semiconductors) to the free electron level, the carrier will no longer be subject to the electrical potential of the material that emitted it. In a vacuum tube, an electrode (cathode K), enclosed in a vacuum cell, absorbs the incident photons, which increase the kinetic energy of the electrons. Once they diffused towards the surface of the cathode, they are emitted into the vacuum. The generated carriers are accelerated towards the other electrode (the anode) placed at a greater potential; in this way a current signal is obtained between cathode and anode proportional to the number of absorbed photons. If no amplification effects are present, currents of a few microamperes are obtained.
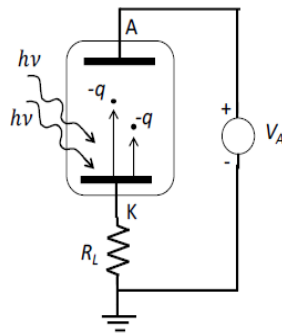
Figure 19.8: Working principle of a Photo tube

A cesium-antimony cathode is very sensitive to the region of violet and ultraviolet. Cesium with oxidized silver provides a cathode that is more sensitive, instead, to infrared and red light. The frequency response is generally limited by the transit time of the electrons from the cathode to the anode.

19.8 Photo-Diodes

Nowadays photodiodes are the most common type of photosensors.
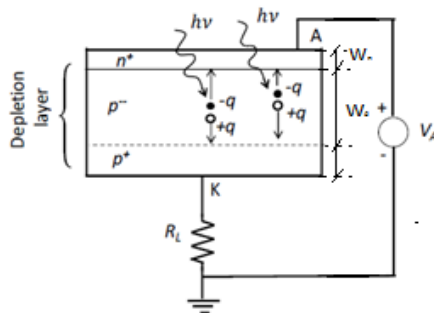

Figure 19.9: Working principle of a Photodiode

In this case, any photon impinging on a reverse-biased p-n junction diode can promote an electron from valence to conduction band of the semiconductor, generating a free electron-hole pair. The free carriers generated in the depletion layer are drawn by the junction electric field, obtaining a current signal proportional to the number of photons absorbed in depletion region. Even without any light hitting the device, a finite current flows in a reverse-biased p-n junction. It is called *Dark Current* in photodiodes and reverse current in ordinary diodes and it is due to spontaneous generation of free carriers in the depletion region due to thermal effects. Semiconductor photodiodes internal noise is typically negligible compared to the circuit noise, but much higher than noise in vacuum phototubes. This fact has significantly limited the size of the active of semiconductor detectors for very low-noise operation; unfortunately, the lower the area, the lower the number of collected

photons. In other words, Vacuum tube sensors can be employed for operating at low noise levels without tight limitations to the sensitive area: this is a clear advantage over semiconductor photodiodes.

19.9 Photon shot noise

As we have seen, optical radiation is composed of quantum particles called photons, which can only be detected in whole numbers. This discretization phenomenon implies that a set of repeated collections of photons will surely suffer from random fluctuations in the number of detected photons. This means that in every photodetector there will be a shot noise independent from the technology. In most cases photon noise can be properly modelled using Poisson statistics, having a variance equal to:

$$\sigma_p{}^2 = \overline{N_p}$$

where $\overline{N_p}$ is the average number of particles that hit the target. Since each photon has energy hv, the total optical power arriving at the detector is $P_p = n_p h v$, being n_p the arrival rate of the photons. Therefore, the power spectral density of the noise will be:

$$S_p = 2hvP_p = 2\frac{hc}{\lambda} P_p$$

In the previous expression an analogy with the electrical shot noise can be easily observed, being $S_I = 2qI_e$ the power spectral density of electrical shot noise. Electrical shot noise, as is well known, comes from the fluctuations of the number of electrons that trespass a potential barrier (i.e. a p-n junction), so it is a consequence of the discretization of the charge. However, photons have a further degree of freedom with respect to electrons, since S_p depends both on the number of particles and on the energy of the particle itself, while S_I depends only the current, so on the number of electrons.

19.10 Shockley-Ramo theorem

The main working principle behind all photodetectors relies in the generation of free carriers once the electromagnetic radiation interacts with the sensitive part of the detector. Under the effect of an electric field, kept by means of an external power supply, the charge moves towards the electrodes of the device, inducing a charge of opposite sign on the electrodes themselves. As the charge between the electrodes moves, the induced charge changes and, if the electrodes are part of a closed form circuit, this variation translates into a current signal. Hence the main issue is the calculation of the induced charge, which can be very difficult if a purely electromagnetic point of view is used. In fact, to calculate the charge we would need to compute several electric fields, each corresponding to a different position of the external charge, through the application of Gauss' law. However, there is a way more convenient method to get to the same result without involving too complex electromagnetics calculations: the Shockley-Ramo theorem. This method was

independently developed by William Shockley and Simon Ramo in the late 1930s specifically for vacuum tubes, but it is valid in general for any kind of detector. In short, the Shockley-Ramo theorem states that the current i_c on a selected electrode induced by a moving charge q is given by:

$$i_c = qE_v \bullet v_c$$

Where $\bullet$ stands for scalar product, v_c is the velocity of the free carrier moving toward the electrode and E_v is called *weighting electric field*. This last element is crucial to understand how the theorem works: E_v is the field that would exist at the position of the charge if the following conditions were verified: the selected electrode is at unitary potential, all other electrodes are grounded, all external charges are removed. Therefore, the weighting field does not correspond to the actual electric field, which is instead hidden in the velocity of the carrier, since drift velocity is usually dependent on it.
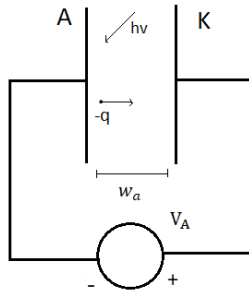
19.11 Application of the theorem: Vacuum Phototube

**Figure 19.10:** Polarization of a Phototube with planar geometry

To illustrate how the theorem can be practically used, a vacuum phototube with planar geometry, shown in figure 19.10, is considered. Our goal is to calculate the *single electron response* (SER) using the Shockley-Ramo theorem. Since the bias generator keeps the voltage between the two electrodes constant, there will be a forcing electric field of value $E_d = \frac{V_A}{w_a}$ between the two conductors. Knowing both the charge of the carrier and the electric field, we can compute the acceleration of the particle, starting from the electrostatic force to which the electron is subject:

$$qE_d = F \qquad a_c = \frac{F}{m_e} = q\frac{E_d}{m_e} = q\frac{V_A}{w_a}$$

Knowing the acceleration, the velocity of the carrier can be derived by simply integrating a_c over time:

$$v_c(t) = \int a_c \, dt = q\frac{E_d}{m_e}t + v_0$$

v_0 is the initial velocity of the carrier, which can be reasonably considered null. Let us summarize the results obtained up until now. The following graphs show respectively $E_d(x)$, $V_A(x)$ and $v_c(t)$:
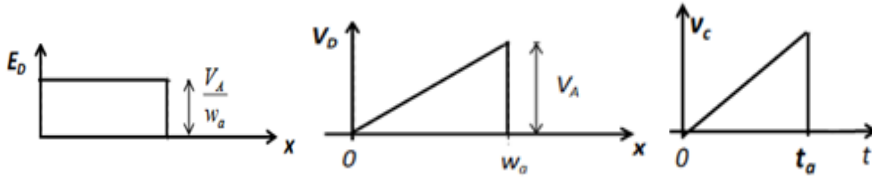
Figure 19.11: Electric field E_d, Voltage in the vacuum tube V_A, velocity of the carrier v_c

As we can see, the velocity of the carrier interrupts after a time interval t_a. Although this last term does not appear explicitly in the formula of $v_c(t)$, it will certainly correspond to the time at which the particle hits the anode (so it reaches $x = w_a$). The value of this time interval will be computed later, for now let us focus on the application of the theorem. In fact, once computed the velocity, the only element that is missing in order to correctly apply the Shockley-Ramo theorem is the weighting field; to calculate it we refer to the three conditions that were mentioned in the definition of the theorem. Hence, the free charge between the conductors must be removed, the electrode toward which the electron was moving has to be set at unitary potential (i.e. 1 Volt) and the other electrode at ground potential. Figure 19.12 shows how the system has been modified after the three conditions are applied:
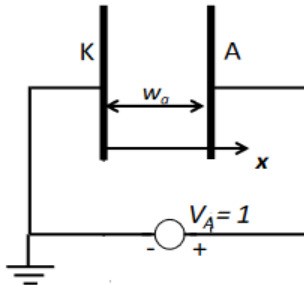
Figure 19.12: Vacuum tube without charge, with an electrode at unitary potential and the other grounded.

Thanks to the particularly easy geometry taken into consideration the computation of the weighting field is immediate since once again inside the detector the electric field is unique and constant, and its value is:

$$E_v = \frac{1}{w_a}$$

Both v_c and E_v were computed in terms of absolute value instead of vectors; however, due to the planar shape of the detector, their directions coincide, so the scalar product

between the two reduces to the basic product of the absolute values (note that there would be a minus sign since the directions are opposite but it simplifies with the charge -q, which is negative for an electron). In the end, the current induced in the electrodes will be:

$$i_c(t) = qE_v v_c = \frac{q^2 V_A}{m w_a^2} t$$

At first, we can observe that it increases linearly with time, as velocity does, so it is reasonable to think that the current pulse will have a triangular shape, since the velocity of the carrier abruptly interrupts once the electron reaches the anode. Therefore, the duration of the pulse coincides with the transit time of the electron between the two electrodes, which can be computed making simple cinematic considerations. We start from the expression of the travelled distance of the carrier, obtained by integrating velocity over time:

$$x(t) = \int v(t)dt = q\frac{E_d}{2m_e}t^2 + x_0$$

In our reference system $x_0 = 0$. The transit time t_a corresponds to the moment when the electron reaches the anode, so $x(t_a) = w_a$, resulting:

$$q\frac{E_d}{2m_e}t_a^2 = w_a$$

Therefore, we get:

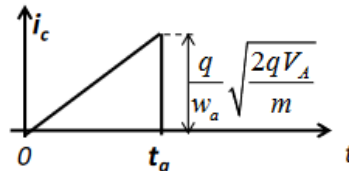$$t_a = \sqrt{\frac{2m_e w_a}{qE_d}} = \sqrt{2\frac{m}{q}}\frac{w_a}{\sqrt{V_A}}$$

Figure 19.13: The graph shows the form of the current pulse

By the time we have a better idea on the shape of the SER, we can also make some considerations on the frequency response of our system. The high frequency cut-off of the detector is, as first approximation, inversely proportional to the duration t_a. Since we want the system to respond as quickly as possible, we need to reduce t_a at minimum. The parameters on which we can intervene to achieve this goal without changing the system are w_a and V_A. Typical values for these quantities are about 1cm for the phototube length and around 100V for the bias voltage, which lead to a transit time of approximately three nanoseconds, translating into frequencies of some hundreds MHz. We can't keep increasing V_A and reducing w_a in order to increase the

frequency response, since not only it will lead to an increase of costs related to a more complex technology, but it would also produce very high levels of electric fields which can be difficult to handle. The following paragraph shows a clever method to increase the frequency response of the phototube without incurring in the drawbacks that were mentioned.

19.12 Application of the theorem: Phototube with screened anode

As explained in the introductory paragraph for the Shockley-Ramo theorem, an electrode instantly reacts to the presence of a moving charge, starting a pulse current signal that propagates in the rest of the circuit. However, as shown in the previous paragraph, the frequency response of a vacuum tube photodetector yet depends on the transit time of the electron between the two metal plates. Our goal is hence to decrease the transit time as much as possible without further increasing the electric field. To do so we modify the vacuum tube by simply inserting a metal wire grid in front of the anode, as shown in figure.
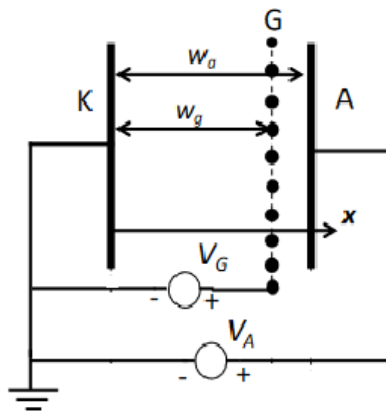
Figure 19.14: Vacuum tube with screened anode

The grid acts as a controller for the device; in fact, by modifying the voltage V_G applied between the grid itself and the cathode, we can basically control the flow of electron inside the device. A very similar structure was developed in 1906 by Lee De Forest: the inventor of "Audion", the first vacuum triode which was able to perform an amplification of a signal. In order to better understand the working principle of the structure, let us see how the circuit behaves when different voltages are applied to the grid. In the case $V_G < 0$, the grid repulses the electrons coming from the cathode, thus some of them will not be able to reach the anode. Viceversa, if $V_G > 0$, some of the electrons will be "captured" by the grid and yet they fail to reach the anode. This mechanism of modifying the number of electrons that reaches the anode obviously leads in change of current through the circuit and in particular it is used in triodes which basic functioning is the application of a signal to the grid itself that

translates into a large variation of current. However, for the purpose of photodetectors, the voltage of the grid has to be set at a particular value, the one for which it does not modify the overall electric field between the two electrodes. Taking figure 19.14 as a reference, we know that in the absence of the grid the electric field inside the device would be constant and equal to $\frac{V_A}{w_a}$, on the other hand the electric field between the cathode and the biased grid would be $\frac{V_G}{w_g}$. Since we are willing to get the same electric field in the overall system, with both generators on, we must equate the two fields:

$$\frac{V_G}{w_g} = \frac{V_A}{w_a} \rightarrow V_G = V_A \frac{w_g}{w_a}$$

From a graphic point of view, we can see that with this particular value for V_G the shape of the overall potential across the two plates has not changed, as shows in the following graph:
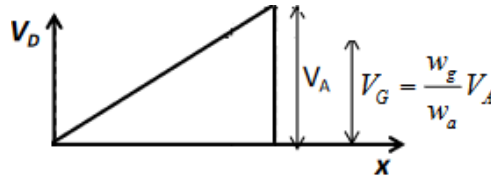

Figure 19.15: The shape of the overall potential across the two plates

Since the electric field inside the detector has not changed with respect to the case shown in the previous paragraph, also the motion of the electron will remain unchanged. Therefore, let us recall the expression of the velocity of the electrons that we previously obtained:

$$v_c(t) = q \frac{E_d}{m_e} t u_x$$

Where $E_d = \frac{V_G}{w_g} = \frac{V_A}{w_a}$ is the unchanged electric field and u_x is the unit vector of the x-axis. Although everything seems to be the same as the previous case, the crucial difference is that, as we will show, the grid does not allow the traveling electron to induce a charge on the anode. Once again, the proposed one is a complex electromagnetic problem that would be very difficult to solve without the Shockley-Ramo theorem. However, before applying the theorem, let us focus for a while on the physical situation we are handling. The electron traveling from 0 to w_g is certainly inducing a charge on the grid which compensate the effect of the electron. However, on the anode plate there cannot be any induced charge until the electron actually reaches the grid, otherwise the charge balance of the system could not be respected. Let us verify our assumption by computing the weighting field in the present situation, once again recalling the three conditions under which it has to be computed:

electron removed, anode raised at unitary potential and all other electrodes grounded. The following picture depicts the situation:
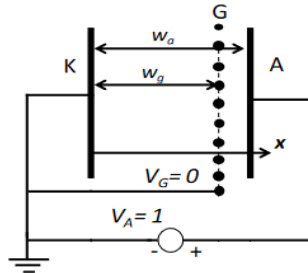

Figure 19.16: Vacuum tube with screened anode in the hypothesis of the Ramo Theorem

Since the cathode and the grid are in a short-circuit, the voltage drop between them has to be zero, so also the electric field is necessarily zero in that region. Between the anode and the grid instead there is a voltage drop of 1Volt, therefore the field, due to the geometry of the system, will be:

$$E_v = \begin{cases} 0 & \text{if } 0 < x < w_g \\ \dfrac{1}{w_a - w_g}(-u_x) & \text{if } w_g < x < w_a \end{cases}$$

Obviously, $w_a - w_g$ is the distance between the anode and the grid, we can visualize the overall behaviour of the electric field in the following:
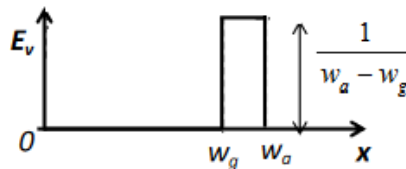
Figure 19.17: Overall behaviour of the electric field in the device

The field has a rectangular shape much thinner than in the case without the grid and its value is way larger than in the previous case. Since now we have all the ingredients that are needed to apply the Shockley-Ramo theorem; we can perform the scalar product between the velocity of the electron v_c and the weighting field E_v, once again noticing that they are both directed in the same way (the direction of the x-axis), so the scalar product reduces to the classical product of their absolute values except for a minus sign, which will be compensated by the fact that the electron has negative charge. In short, the SER current pulse is:

$$i_c(t) = qE_v v_c(t)$$

Since the field has rectangular shape, smaller in width than the one of the triangular shape of the velocity, we expect the current to have a trapezoidal shape which width has to be smaller than the transit time of the electron inside the device (t_a). As we said before, the anode is able to notice the presence of the charge only when the electron reaches the grid: this happens after a time interval t_g, which can be computed in a similar way of t_a in the previous paragraph, thus resulting:

$$t_g = \sqrt{\frac{2m_e w_g}{qE_d}} = \sqrt{2\frac{m}{q}}\frac{w_g}{\sqrt{V_G}}$$

As the pulse starts in t_g and it lasts until t_a, that is the time in which the electron hits the anode. Now that we have the overall picture of the situation, we can sketch the graph of the resulting current pulse as it follows:
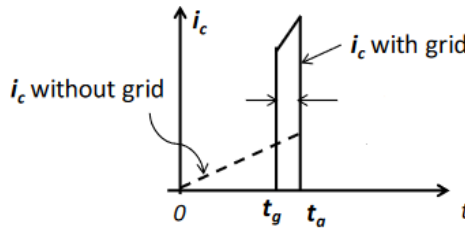
Figure 19.18: Difference between the current pulse in the case of screened anode and without the grid

The duration of the pulse obtained in the new configuration is certainly shorter than the previous one; we can compute it by simply subtracting the two quantities:

$$t_a - t_g = \sqrt{2\frac{m}{q}}\frac{w_a}{\sqrt{V_A}} - \sqrt{2\frac{m}{q}}\frac{w_g}{\sqrt{V_G}} = \sqrt{2\frac{m}{q}}\left(\frac{w_a}{\sqrt{V_A}} - \frac{w_g}{\sqrt{V_G}}\right)$$

But since $V_G = V_A\frac{w_g}{w_a}$:

$$t_a - t_g = \sqrt{2\frac{m}{q}}\left(\frac{w_a}{\sqrt{V_A}} - \frac{w_g}{\sqrt{V_A\frac{w_g}{w_a}}}\right) = \sqrt{2\frac{m}{q}}\frac{w_a}{\sqrt{V_A}}\left(1 - \frac{\sqrt{w_g}}{\sqrt{w_a}}\right) = t_a\left(1 - \frac{\sqrt{w_g}}{\sqrt{w_a}}\right)$$

Since our initial goal was to increase the frequency response of the detector, we see that in order to have $t_a - t_g$ as short as possible we must put the grid as close as we can to the anode.

CHAPTER 19

Phototubes

Phototubes are photodetectors that feature high sensitivity, superior temperature stability, wide dynamic range, large photosensitive area. They are widely used in applications such as chemical and medical analysis and laser measurement.

20.1 Introduction

Phototubes are high-sensitivity devices also known as photoemissive cells. They are one of the oldest and most elaborate ways to transduce light into electricity. They basically are vacuum or gas-filled tubes containing two opposite electrodes: a photocathode and an anode. Thanks to the photoelectric effect, incoming photons hitting the photocathode knock electrons out of its surface and they are attracted to the anode. Thus, current is dependent on the frequency and intensity of the incoming photons. Unlike photomultiplier tubes (as we will see in the next chapters), no amplification takes place, so the current through the device is typically in the order of few microamperes. The light wavelength range over which the device is sensitive depends on the material used for the cathode.

Some basic features of the phototubes are of our interest: their high-speed response (as phototubes they have low capacities, so their rise-time characteristics are fast) and large photosensitive area compared to semiconductor sensors.

20.2 Phototube Device Structure

The structure of a phototube basically features an evacuated glass tube which contains two metal electrodes: a photosensitive cathode that emits electrons when illuminated and an anode for collecting the emitted electrons. The tube is connected to a power supply that controls the potential difference V between the two electrodes.
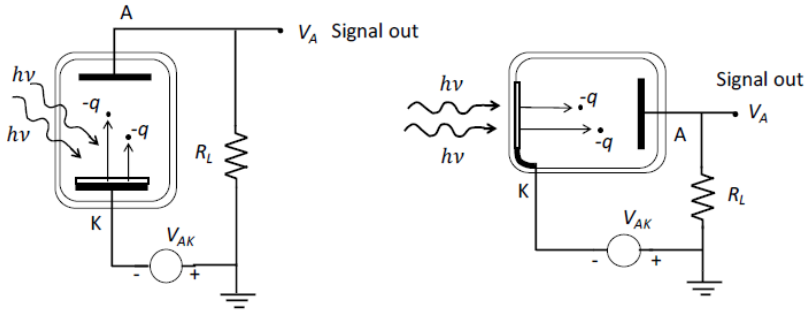
Figure 20.1: Side Window Tube and End Window Tube

Figure 20.1 shows the two main structures of a phototube: on the left, we have a photocathode made of a thick opaque layer deposited on a metal-support electrode, and a side window of the glass tube which allows the light to enter the device. This structure has an unfavorable geometry and the collection of light on the photocathode is uneasy and not very efficient. On the right-side of figure 20.1, we can see an end-window tube: the photocathode is made of a thin transparent layer deposited on the interior of the glass-tube end. This structure has a favorable geometry and the collection of light is easy and efficient.

20.3 Stationary I-V Characteristics

A typical current-voltage characteristic for a phototube is shown in figure 20.2.
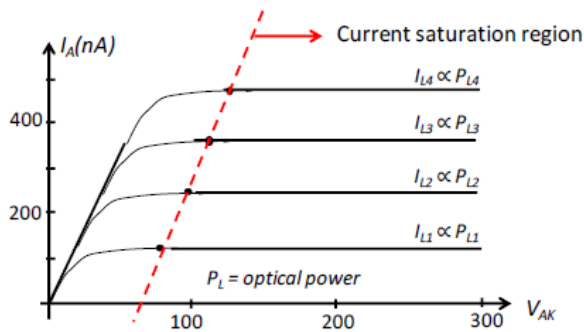
Figure 20.2: I-V characteristics of phototube

At low V_{AK} voltage, the photocurrent collected at the anode side is limited by the electron space-charge effect. As V_{AK} is increased, the higher electric field reduces the space charge and a higher number of emitted electrons is collected: as a consequence, the current increases. As V_{AK} exceeds a saturation value V_{AKS}, all photoelectrons are

collected, and the current is constant vs. V_{AK}. The saturation value V_{AKS} increases with the optical power P_L hitting the detector. Phototubes are typically operated biased in the current saturation region.

20.4 Electron Photoemission and Photocathode technology

As mentioned before, the working principle of phototubes is based on the photoelectric effect, i.e. the photocathode can emit electrons when photons strike the surface. The emitted electrons will travel across the evacuated tube and reach the other end, where they will be collected by the anode.

Photoelectrons are emitted when a single photon (quanta) of energy hv is absorbed by the material, where h is Planck's constant and v the frequency of the light used. The energy of the photon must be larger than the energy separation between the top of the valence band and the vacuum level.

The design of photocathodes is based on the 3-step model. In this approximation, the photoelectron emission process is divided into three separate steps:

1. Photon absorption: free electrons are generated by photon absorption.
2. Transport of excited electrons to the surface; the electrons generated are diffused in photocathode layer.
3. Passage through the surface and escape into the vacuum

To maximize the total efficiency of the device, equal attention must be given to each one of these steps. The excited electrons may lose energy during transport to the surface and so their flux is attenuated. The attenuation is lower for materials where inelastic electron scattering is hindered, like insulators or wide band-gap semiconductors, that is why the most suitable materials are semiconductors. Metals are unsuitable due to their high reflectivity, small diffusion length and low escape probability (high potential step from inside up to the vacuum level).

20.4.1 Ordinary Photocathodes with positive E_a

We have different types of photocathodes and first of all we are going to consider the basic ones, usually referred to as ordinary photocathodes. Ordinary photocathodes are characterized by positive electron affinity (E_a). If the energy of the absorbed photon is larger than the energy separation between the top of the valence band and the vacuum level, electrons can escape into the vacuum level. Otherwise, if this energy is not enough, the electron does not escape and in a few ps it thermalizes down to the bottom of the conduction band. At the bottom of conduction band an electron can diffuse further a few μm and about 100ps before recombining (i.e. getting down to valence band) but it cannot escape any more into vacuum.

A fundamental monitor of photocathode performance is the quantum detection efficiency (QE), which is defined as the percentage of ejected electrons per incoming photons.
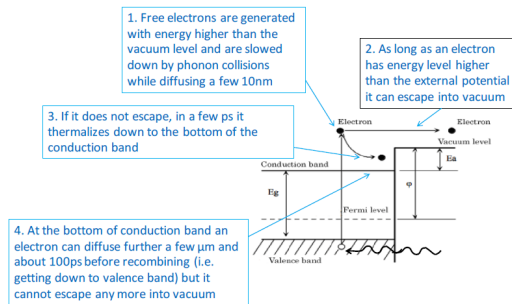
Figure 20.3: Photo-emission on photocathodes with positive Ea

In order to provide good quantum detection efficiency, the photocathode material must fulfill some basic requirements.

1. The inside-to-vacuum energy barrier $E_g + E_a$ must be smaller than the photon energy E_p. In the visible range $1{,}6$ eV $< E_p < 3{,}1$ eV and $E_g \approx 1$eV for semiconductors; therefore, the electron affinity must be limited.

$$E_a \leq 1eV$$

2. In general, there is always a tradeoff between the need to have a cathode material which strongly absorbs light, but sufficiently thin that electrons can be extracted. The light penetration in the cathode is strongly wavelength dependent. For short wavelengths in a multialkali cathode, for example, this is on the scale of 30nm for wavelengths below 400nm.

    Furthermore, electrons generated in deep layers are not emitted, since they are not able to reach the surface. Escape probability is high only for electrons generated in a surface layer that is very thin, about a diffusion length L_{eh} of high-energy electrons. For a significant absorption in this layer, the optical penetration length L_a must be anyway NOT much higher than L_{eh}. For a high absorption the two should be comparable with each other:

$$L_a \approx L_{eh}$$

In conclusion, the thickness of the photocathode layer contributing to the electron emission is intrinsically limited to about L_{eh} in any case. That is, the active layer is very thin, independent from the total thickness of the photocathode.

20.4.2 SemiTransparent Photocathodes

The active layer of the previous photocathodes was always very thin, also for thick cathodes deposited on a metal electrode, because the thickness of the cathode layer contributing to the electron emission is limited.
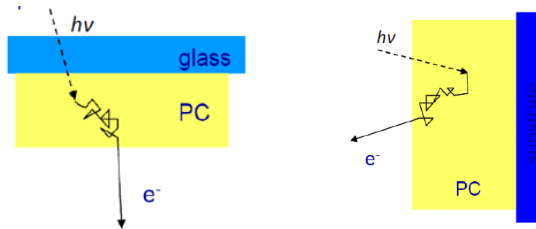
Figure 20.4: Semi transparent and opaque photocathode

Due to this fact, a new type of photocathode was developed. It is a thin photocathode (with thickness about $\approx L_{eh}$) deposited on the interior of the glass tube in the end-window of the detector. They are called *semitransparent cathodes*.

They are illuminated on the outer side through the glass window end emit photoelectrons from the inner side. They make it possible and easy to achieve a much better optical collection with respect to the side-window geometry.

20.4.3 Photocathode Types

The light wavelength range over which the phototube is sensitive depends on the material used for the cathode.

Classifications of Photocathode types are made by industrial standard committees. Most widely used is that by JEDEC (Joint Electron Devices Engineering Council US), which denotes cathode types S1, S2, ... and classifies them by spectral responsivity type (rather than by chemical composition or fabrication recipe).

Let's see some key types:

- S1 photocathode was introduced in the '30s and it is still in use. The QE is low (peak $\eta_D \approx 1\%$) but it covers a wide spectrum in the IR. It is a matrix of Cesium oxide that includes silver microparticles and it's currently denoted as Ag-O-Cs.

Highly efficient photocathodes for the visible range were introduced in the '50s and progressively developed employing compounds of alkali metals (Na, K, Cs, which have low work functions) and Antimony (Sb). Main types are:

- S11 (SbCs) is one of the earliest photocathodes with a spectral response covering the ultraviolet and visible range, offering better blue response and lower dark current. It ranges from 300nm to 600nm and has a maximum quantum detection efficiency of 15% at 450nm.
- S20 ranges from 300nm to 800nm, has a maximum $\eta_D \approx 20\%$ at 350nm; multi-alkali halide Na-K-Sb-Cs.

- S25 extends the range up to 800nm, peak $\eta_D \approx 5\%$ at 600nm; multi-alkali Na-K-Sb-Cs like S20, but with a thicker layer that gives higher sensitivity in the red, at the cost of lower sensitivity in the blue-green.

In figure 20.5 it is possible to see the radiant sensitivity graph of different types of cathodes.
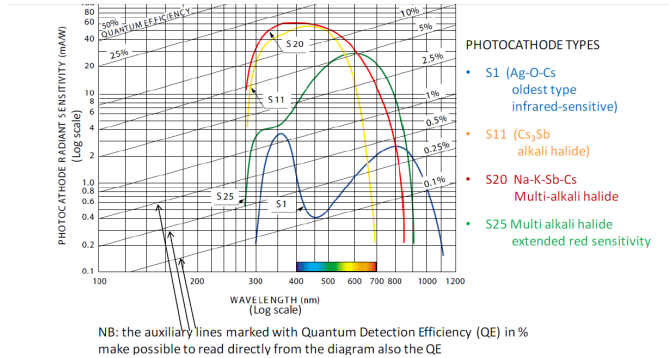

Figure 20.5: Photocathode Radiant Sensitivity

20.4.4 PhotoCathodes with negative E_a

Progress in semiconductor physics and technology led in the '70s to devise a new class of photocathodes, called photocathodes with Negative Electron Affinity (NEA) Starting from a GaAs crystal substrate, a few atomic layers of Cesium Oxide (*Cs-O*) are deposited and activated, thus forming a very thin positive charge layer of Cs^+ ions.

The electric field generated at the surface curves downward the energy bands: the vacuum potential level is now lower than the bottom of conduction band, i.e. the electron affinity E_a is negative. Electrons can now escape into vacuum also when thermalized at the bottom of conduction band and so the QE is enhanced.

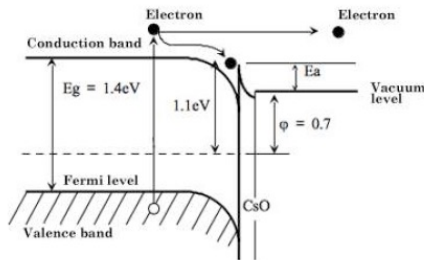Photoelectron emission is obtained also with photons with lower energy E_p, down to the GaAs energy gap E_g

Figure 20.6: p-doped GaAs crystal activated to achieve negative electron affinity

20.5 Detector Dark Current and Noise

A finite current is emitted by any photocathode even when kept in the dark, i.e. without any light reaching it; this a spontaneous emission of carriers due to thermal effects (phonon-electron interactions in the cathode); it is called *Dark Current*.

The dark current density j_B (per unit area of cathode) depends on the cathode type and on the cathode temperature. Typical values at room temperature are reported in the following Table:

PhotoCathode type	Dark Current density j_B in A/cm²	Dark Electron Rate density n_B in electrons/s·cm²
S1	$\approx 10^{-13}$	$\approx 10^6$
S11	$10^{-16} - 10^{-15}$	$10^3 - 10^4$
S20 and S25	$10^{-19} - 10^{-16}$	$1 - 10^3$
GaAs NEA	$10^{-18} - 10^{-16}$	$10 - 10^3$

The total Dark Current is

$$I_B = j_B A_D$$

where A_D is the area of the photocathode. The shot noise of I_B is the photodetector unavoidable internal noise, with effective power density (unilateral):

$$\sqrt{S_B} = \sqrt{2qI_B} = \sqrt{2qj_B}\sqrt{A_D}$$

Typical values of $\sqrt{S_B}$ are reported in the table below:

PhotoCathode type	Dark Current density j_B A/cm²	Shot Noise Effective density $\sqrt{S_B}$ $pA/\sqrt{Hz}\sqrt{cm^2}$
S1	$\approx 10^{-13}$	$\approx 10^{-4}$
S11	$10^{-16} - 10^{-15}$	$\approx 10^{-5}$
S20 and S25	$10^{-19} - 10^{-16}$	$\approx 10^{-7} - 10^{-6}$
GaAs NEA	$10^{-18} - 10^{-16}$	$\approx 10^{-6}$

We know that for low-noise operation a high impedance sensor must be connected to a preamplifier with high input impedance and low input noise. The best available preamplifiers have current noise at room temperature:

$$\sqrt{S_i} \approx 0.01 pA/\sqrt{Hz}$$

When dealing with phototubes, the circuit noise $\sqrt{S_i}$ is always dominant and the detector internal noise $\sqrt{S_B}$ plays in practice no role with any phototube, even for detectors with S1 photocathodes (that have the highest noise) and even with very wide sensitive area (up to many square centimeters). In fact, for producing shot noise with power density higher than typical circuit noise, the phototube dark current should be $I_B > 300$pA, corresponding to an emission rate $n_B > 10^9$ electrons/s.
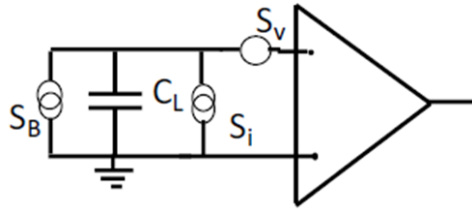

Figure 20.7: Equivalent circuit of high impedance sensor

Vacuum tube sensors can thus be employed for operating at low noise without stringent limits to the sensitive area. As we will see, this is a definite advantage over semiconductor photodiodes, which are dominated by the internal shot noise.

20.6 Low- noise Preamplifiers for Phototubes

Phototubes are high-impedance sensors; hence for low-noise operation it is mandatory to employ a preamplifier with high input resistance $R_{iA} \to \infty$
A typical configuration used in this case consists of a voltage buffer featuring a high input impedance and a low-noise amplifier.
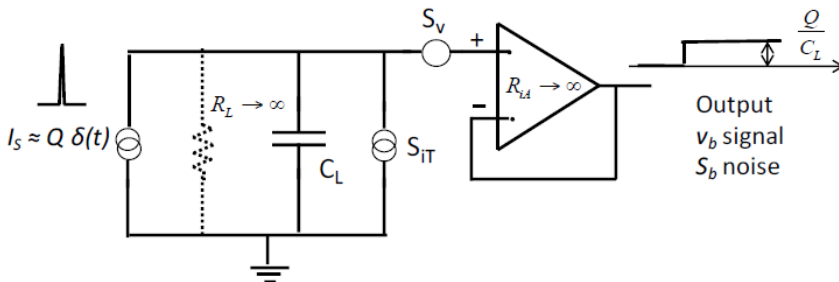

Figure 20.8: Voltage buffer based on high input impedance preamplifier

With this configuration, shown in figure 20.8, we are simply integrating the output current of our device on the capacitor C_L.

Considering

- C_L total load capacitance = C_D (detector cap.) + C_{iA} (amplifier cap.) + C_S (connection cap.)
- R_L total load resistance → ∞
- S_v amplifier voltage noise
- S_{iT} total current noise = S_{iD} detector noise + S_{iA} amplifier noise (+ S_{iR} load resistor noise)

We will obtain

$$v_b(t) = \frac{Q}{C_L}\, 1(t)$$

And, for the noise spectrum

$$S_b = S_v + S_{iT}\, \frac{1}{\omega^2 C_L^2}$$

The buffer configuration has some noteworthy drawbacks.

- The signal amplitude Q/C_L is ruled by the total capacitance $C_L = C_D + C_{iA} + C_s$, whose value is not very small and not well controllable, particularly in cases where long sensor-preamplifier connections contribute to a remarkable C_s. C_L may be different from sample to sample of the amplifier, even of the same amplifier model.
- With signals in high-rate sequence, the superposition of voltage steps may build-up and produce a significant decrease of the phototubes bias voltage. This may change the operating conditions of the sensor and consequently the parameters and performance of the detector, particularly if the phototube is biased not much above the saturation voltage.

20.6.1 Charge Preamplifier or Transimpedance Preamplifier

An alternative configuration for a charge amplifier is an operational amplifier in integrator configuration based on a low-noise amplifier with high input impedance.
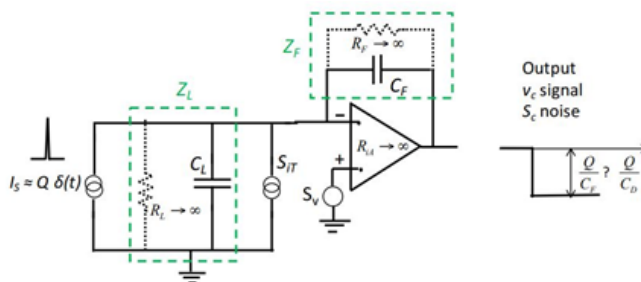

Figure 20.9: Transimpedance pre amplifier

In this case we have

- C_F capacitor in feedback. The C_F *value can be very small and is accurately set* by the capacitor component, because the inherent stray capacitance between output and input pins of the amplifier is negligible. Therefore, *one can work with $C_F << C_L$*
- R_F feedback resistor $\rightarrow \infty$
- C_L total load capacitance $= C_D$ (detector cap.) $+ C_{iA}$ (amplifier cap.) $+ C_S$ (connection cap.)
- R_L total load resistance $\rightarrow \infty$
- S_v amplifier voltage noise
- S_{iT} total current noise $= S_{iD}$ detector noise $+ S_{iA}$ amplifier noise ($+ S_{iR}$ load resistor noise)

We can then calculate the output signal, both in frequency domain

$$V_c = -QZ_F = -\frac{Q}{j\omega C_F}$$

And in time domain

$$v_c(t) = -\frac{Q}{C_F}(t)$$

With respect to the buffer, the amplitude is greater by the gain factor

$$G_c = C_L/C_F >> 1$$

$$|v_c| = \frac{Q}{C_F} = \frac{C_L}{C_F}\frac{Q}{C_L} = \frac{C_L}{C_F}|v_b| = G_c|v_b|$$

We have seen that the higher the signal, the lower is the impact of the noise of the following circuits, so the advantage of this stage is its amplification. Furthermore, the signal amplitude is ruled by the well-controlled and stable C_F, no more by the other capacitance C_D, C_{iA}, C_S. Finally, the detector terminal is connected to the virtual ground of the amplifier, hence it stays at constant bias voltage even with signals coming in high-rate sequence.

If we consider now the noise of the configuration, denoting by Z_L the load impedance and by Z_F the feedback impedance we have

$$S_C = S_V \left|1 + \frac{Z_F}{Z_L}\right|^2 + S_{iT}|Z_F|^2$$

In our case $Z_L \sim 1/j\omega C_L$ and $Z_F \sim 1/j\omega C_F$ so that:

$$S_C = S_V \left|1 + \frac{C_L}{C_F}\right|^2 + S_{iT}\left(\frac{1}{\omega^2 C_L^2}\right) = \left(\frac{C_L}{C_F}\right)^2 \left[S_V(1 + \frac{C_F}{C_L})^2 + S_{iT}(\frac{1}{\omega^2 C_L^2})\right]$$

If $\frac{C_F}{C_L} \ll 1$, with good approximation it is

$$S_C \sim \left(\frac{C_L}{C_F}\right)^2 \left[S_V + S_{iT}(\frac{1}{\omega^2 C_L^2})\right] = \left(\frac{C_L}{C_F}\right)^2 S_b = G_c^2 S_b$$

With respect to the buffer, the signal and noise thus benefit of the same gain G_C: therefore, the attainable S/N is the same with the charge preamplifier as with the voltage buffer preamplifier (as expected).

20.8 NEP and Detectivity

Photodetector sensitivity is a convenient, even necessary, metric by which the performance of a particular photodetector can be quantified and compared with other detectors. However, in some cases this parameter could be not enough. The *Noise Equivalent Power* (NEP) is a figure of merit that takes into account the photon detection efficiency and the detector dark-current noise, but not the preamplifier noise. We can define it as *the input signal power that results in a signal-to-noise ratio (S/R) of 1 in a 1 Hz output bandwidth*. Essentially, the NEP expresses the minimum detectable power per square root bandwidth of a given detector; in other words, it's a measure of the weakest optical signal that can be detected if we are limited by the internal noise of the detector and not by the electronic circuit noise. This is NOT the case with phototubes but we will see that it is the case with photomultiplier tubes.
NEP was devised as a figure of merit for comparing objectively the intrinsic quality of different detectors.
Let's consider a photocathode with an area A_D, signal current Ip and Dark Current I_B with area density j_B. Employing a filter with bandwidth (unilateral) Δf we have noise:

$$\sqrt{\overline{i_n^2}} = \sqrt{2qI_B \Delta f} = \sqrt{2qj_B}\sqrt{A_D}\sqrt{\Delta f}$$

And we would have a signal- to- noise ratio:

$$\frac{S}{N} = \frac{I_P}{\sqrt{\overline{i_n^2}}}$$

The minimum measurable current signal $I_{p,min}$, which corresponds to S/N=1, is:

$$I_{p,min} = \sqrt{\overline{i_n^2}} = \sqrt{2qj_B}\sqrt{A_D}\sqrt{\Delta f}$$

For illumination with optical power P_P at a given λ the Detector Responsivity is:

$$S_D = \frac{I_P}{P_P} = \eta_D \frac{\lambda}{hc/q} = \eta_D \frac{\lambda[\mu m]}{1,24}$$

As mentioned before, NEP is defined as the minimum input optical power $P_{P,min}$ corresponding to the minimum measurable signal:

$$NEP = P_{P,min} = \frac{I_{P,min}}{S_D} = \frac{\sqrt{\overline{i_n^2}}}{S_D} = \frac{\sqrt{2qj_B}\sqrt{A_D}}{S_D}$$

However, the NEP is not a fully objective figure of merit for assessing and comparing the quality of photocathodes: in fact, cathodes of equal quality have different NEP if they have different area.
A different figure named *Detectivity D** was therefore derived from the NEP by

a) considering the NEP value normalized to unit sensitive area ($A_D = 1\text{cm}^2$)
b) defining the Detectivity D* as the reciprocal of the normalized NEP

The advantage of using specific detectivity D* is that it allows the comparison of detectors that have different active areas.

$$D^* = \frac{\sqrt{A_D}\sqrt{\Delta f}}{NEP}$$

Which would be:

$$D^* = \frac{S_D}{\sqrt{2qj_B}} = \eta_D \frac{\lambda[\mu m]}{1,24} \frac{1}{\sqrt{2qj_B}}$$

To select the optimal photodetector for a certain application, a number of different factors need to be considered. Above all, the wavelength range, the detector bandwidth, the conversion gain, and the detector size should match the requirements of the intended measurement.

Semiconductor Photodiodes

In this chapter a particular kind of sensor is presented: the photodiode. We will describe the physical structure and behavior of this device. We will study how to deal with noise in the system and to quantify its efficiency.

21.1 Introduction

The physical structure of a semiconductor photodiode is based on a pn junction: we will draw the characteristic of this sensor and find a suitable stationary equivalent circuit for each working point. We will discuss the main features and limits of this detector: starting from the explanation of how electron-holes pairs are generated, we will introduce the photon detection efficiency of this sensor; at the same time, we will face issues related to dark current and noise. By applying the Shockley-Ramo theorem we will finally study the current flowing in the external circuit.

With an equivalent circuit we will determine the dynamic response: we will see that the sensitive area has to be limited to get a reasonable time constant. Some unwanted effects connected to the diffusion of carries in neutral regions will be considered.

21.2 PhotoDiode (PD) devices

In this section, we will mainly consider pn junction-based photodiodes because of their advantageous dimensions and good performance. The working principle is the same as phototubes, but since they are based on semiconductor material they can be easily integrated and employed in optoelectronics applications such as optical communications.

Figure 21.1 shows the internal structure of a pn junction photodiode. As you can see, it can be both side-illuminated (on the left) or front-illuminated (on the right). The side-illuminated solution is adopted for specific purposes, e.g. in microsystems with integrated waveguides for on chip optical connections. Front-illuminated photodiodes, instead, are more widely employed since the design of the active area

(that collects the light) is more flexible and the size can be wider. Illumination usually happens through a window defined by an annular electrode.
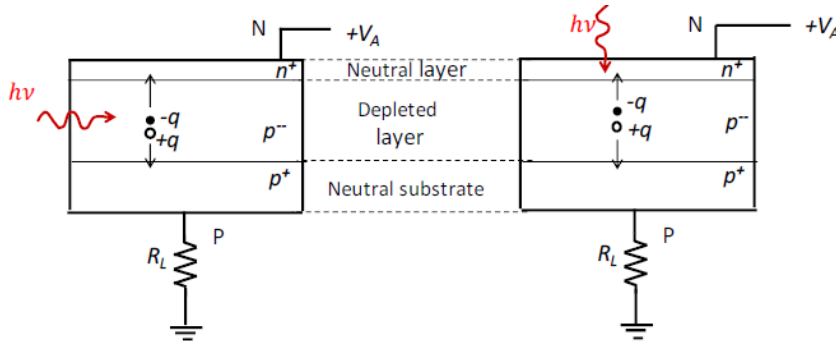
Figure 21.1: Basic device structure of a photodiode.

For the sake of simplicity, we choose to study the structure in figure 21.1 as a 1-dimension device, that is to neglect the curvature effects. The most used material is Silicon with different doping on the two sides. Referring to figure 21.1, the top region is doped with a concentration N_d of donors, whereas the bottom region has a doping concentration N_a of acceptors. Let's consider the case where N_a and N_d are homogeneous in their regions and let's consider a p^+n junction, which means that $N_a \gg N_d$. The p^+ region is typically very thin (thinner than 1 μm), so it is usually formed by planar diffusion into the n-layer. As the n-region and the p-region are put in contact, the majority carriers of each side diffuse towards the opposite side. At the equilibrium we have two neutral layers at the opposite ends and a depleted layer DL around the junction. In the DL there are positive exposed charges, ionized donors, on the n side and negative exposed charges, ionized acceptors, on the p side. The number of free carries here is negligible with respect to N_d and N_a.

21.3 Carrier generation

When light impinges on the photodiode, we can have two different scenarios:
- *the photon is absorbed in the depleted region*: the photon overcomes the neutral input region without being absorbed and the absorption event happens in the depleted region in which a pair is generated. A carrier in the depleted layer induces opposite charges in the conductive electrodes (neutral semiconductor layer and metal contact to the external circuit). The value of the induced charge on a given electrode depends on the carrier distance from the electrode. If the carrier moves, the charge induced on the electrode varies, hence current flows through the contact.
- *the photon is absorbed in the neutral region*: the absorption of the photon takes place in the neutral region and here the electron-hole pair is generated. A carrier in the neutral region is surrounded by a huge population of other free carriers. When a

carrier moves, the distribution of free carriers swiftly rearranges itself to electrically screen any effect of the carrier motion to the external circuit.
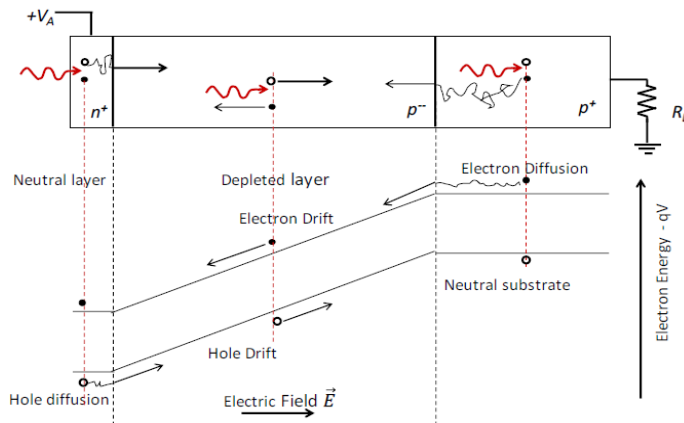
Figure 21.2: Carrier motion in PD. For sake of simplicity the band diagram is linear, instead of quadratic.

To summarize, if the electron-hole pairs is generated in the depleted region, it feels the high electric field and causes current to flow through the metal contact to the external circuit, whereas if the carriers are generated in the neutral region they do not feel any electric field and do not generate any current.

21.4 Carrier diffusion effects

Now we want to study more in detail the effects of free carriers diffusing in the neutral regions. Let us first consider the current that flows in the external circuit as a consequence of the photogeneration of a single free carrier.

According to what we have previously explained, a carrier generated in the DL is immediately drifted by the electric field, so the response is a prompt current pulse (Figure 21.3, case a). In alternative a carrier can be generated in a neutral region and it random-walks by diffusion: in this case, the carrier can either attains the DL or not; in the second case, it recombines to a trap state according to the Shockley-Hall-Read theory. The recombination rate is
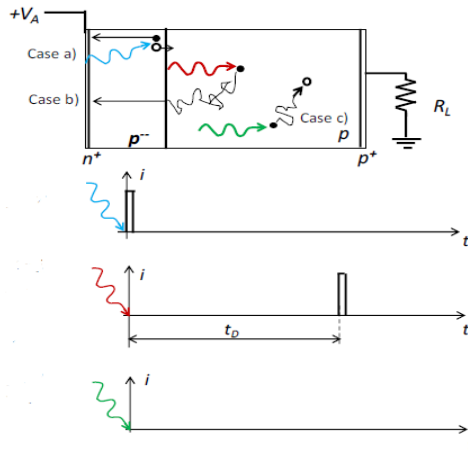
$$R = \frac{\Delta n}{\tau}$$

Figure 21.3: Single-carrier responses.

where Δn is the minority carrier excess with respect to the equilibrium condition and τ is their lifetime.

When a carrier generated in the neutral region reaches the DL, it gives rise to a pulse with a random delay t_D equal to the time needed by the particle to reach the DL itself (figure 21.3, case b); in this case, the signal lasts as long as the time needed by the electron or the hole to cross the DL.

It is obvious that in case the electron and hole recombine before they reach the DL, no current pulse can be measured by the external circuit (figure 21.3, case c).

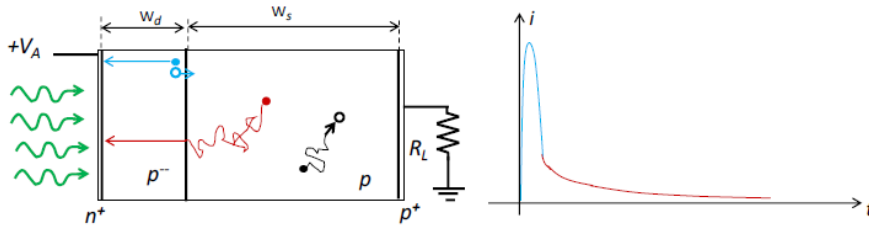Now we can describe the response to a multiphoton pulse (figure 21.4).

Figure 21.4: Response to a multiphoton pulse.

The blue part of the graph represents the short main pulse, which is the response to the photons absorbed in the depleted layer. The longer and slower "tail" is due to the photons absorbed in the neutral substrate. The shape and the size of this red part are related to the device geometry, to material properties such as the diffusion coefficient and the minority carrier lifetime and to the space distribution of the absorbed photons, hence to the wavelength. The presence of the "diffusion tail" improves the photon detection efficiency, since it brings to the output the contribution of photons absorbed

in the neutral region, but it degrades the dynamic response since the "tail" is usually far longer than the main pulse. The span of the red part increases with the width of the neutral regions and with minority carrier lifetime, which is longer at a lower doping level. In Si-photodiodes the "tail" can be quite significant, ranging from a few 100ns with thick layer (ws >100μm) and low doping ($\approx 10^{14}$/cm³) to a few 100ps with thin layer (ws≈1÷2μm) and moderately high doping ($\approx 10^{16}$/cm³)

21.5 I-V characteristics of PhotoDiodes

A silicon photodiode can be modelled with a current source in parallel to a capacitor (figure 21.5). The current source represents the current generated by the incident radiation, and the capacitor (C_D) represents the p-n junction. In addition, a *shunt resistance* (R_j) is typically in parallel to the other components, while a *series resistance* (R_D) is connected in series with all components in this model.
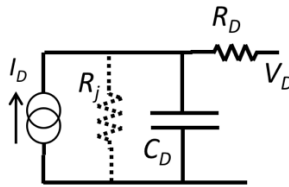

Figure 21.5: Equivalent circuit of a photodiode

The I-V characteristic is equal to the one of a normal diode but it is shifted vertically by the additional photogenerated contribution as we can see in figure 21.6.
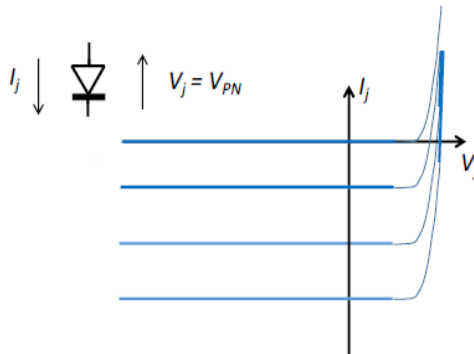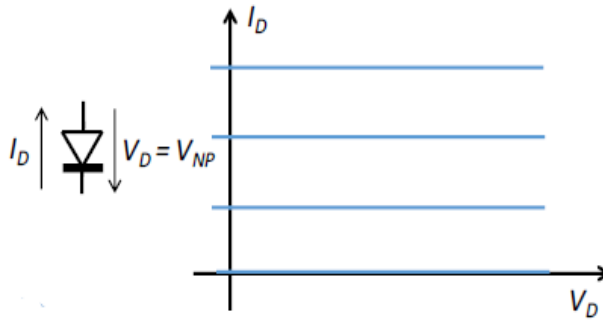

Figure 21.6: I-V characteristic of a photodiode

$$I_j = I_0 \exp\frac{qV_j}{kT} - I_0 - I_p$$

With:
- V_j: direct bias voltage
- I_p: photocurrent (photogenerated carriers)
- I_0: reverse current (thermally generated carriers)

Figure 21.7: Reverse I-V characteristic

The characteristic above is simply obtained flipping on both axis the graph in figure 21.6.

$$I_D = I_p + I_0 - I_0 \exp\left(-\frac{qV_D}{kT}\right)$$

In case of $V_d \gg V_{th}$ the reverse current I_d will be almost equal to the photogenerated current and hence to the impinging light.

$$I_D = I_p + I_0 \approx I_p$$

The photogenerated current will be simply:

$$S_D = \frac{I_{ph}}{P_L} \longrightarrow I_{ph} = S_D \cdot P_L$$

Where S_D is the radiant sensitivity and P_L is the input optical power.

The signal current can be basically read in two *operation modes*: with passive load or with active load. The main purpose of the load is to convert the signal from current into voltage. The active solution is typically preferred in order not to modify the bias of the photodiode and hence its performance. From the SNR point of view the two configurations are obviously the same.

Semiconductor photodiodes can be operated also without a bias voltage source (V_A). As outlined below, the short-circuit current is measured in the *photoconductive mode* and the open-circuit voltage in the *photovoltaic mode*. These configurations have modest sensitivity and slow response, but their simplicity is attractive in some practical cases, e.g. for monitoring a steady light over a wide dynamic range.
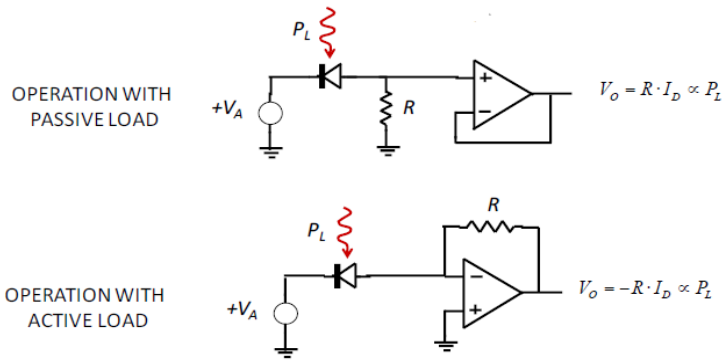
Figure 21.8: Stationary operation of PD

In the photoconductive mode the feedback fixes the bias voltage at zero so the photodiode basically works on the y-axis of the I-V characteristic. All the current will flow through the feedback resistance and the output voltage will be proportional to it.

In the photovoltaic mode, instead, the infinite input resistance of the amplifier makes nil the current. We are now working on the x-axis. The problem of this configuration is that the variations of the voltage are small and often difficult to detect.

PHOTOCONDUCTIVE MODE
PD in short-circuit $V_A = 0$
Linear output scale

$$I_D = I_p + I_0 \approx I_p$$

$$V_o = -R \cdot I_D$$
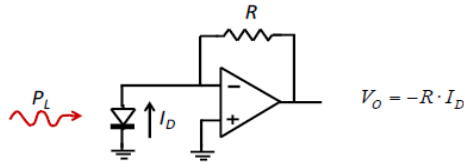
PHOTOVOLTAIC MODE
PD in open-circuit $I_D = 0$
Logarithmic output scale

$$V_j = \frac{kT}{q} \ln\left(1 + \frac{I_p}{I_0}\right) \approx \frac{kT}{q} \ln\left(\frac{I_p}{I_0}\right)$$

$$V_o = V_j$$
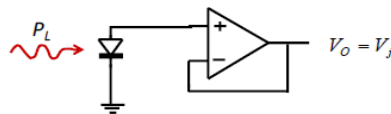
Figure 21.9: Operation modes of PD without bias voltage

21.6 Photon Detection Efficiency η$_D$

The photon detection efficiency or quantum detection efficiency (η_D) in photodiodes is strictly related to the probability (P_d) of a photon to generate a free electron-hole pair in the depletion layer, which is the product of three terms:

- The probability of a photon to not being reflected at the surface of the photodiodes
- The probability of a photon to not being absorbed in the top neutral layer w_n
- The probability of a photon to being absorbed in the depletion layer w_d
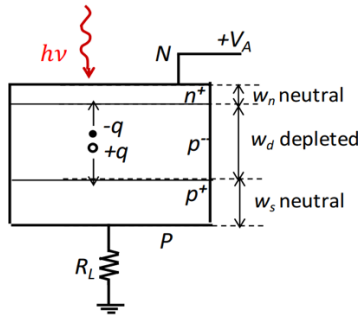
Figure 21.10: Basic device structure of a photodiode.

and so P_d is given by

$$P_d = (1-R) \cdot e^{-\alpha w_n} \cdot \left(1 - e^{-\alpha w_d}\right)$$

where R is the reflectivity, which describes the probability of reflection (this is why we have the term 1-R), α is $1/L_\alpha$ where L_α is the optical absorption depth.

Well, also in the neutral region there is photogeneration of carriers, but in the majority of photodiode structures the probability that those carriers reach by diffusion the depletion layer is low, so in these cases the photon detection efficiency is simply

$$\eta_D = P_d = (1-R) \cdot e^{-w_n/L_a} \cdot \left(1 - e^{-w_d/L_a}\right)$$

Obviously in the structures where the carriers photogenerated in the neutral region have significant probability of reaching the depletion region, we have to take into account other contributions to η_d , but we are not dealing with those cases here.

It is clear that the photon detection efficiency depends on the actual material properties, on the photodiode structure and on the light wavelength λ which cause surface reflection, absorption in the neutral layer and the incomplete absorption in the depletion layer.

The reflection at vacuum-semiconductor's surface is quite strong due to the high step discontinuity in the refractive index n, which is high in semiconductors. For example, in silicon n>3.5 in the visible spectrum and it rises at shorter λ, and so does the reflectivity, which is R>30% in the visible spectrum and it rises at shorter wavelength. We can reduce these reflection losses by applying layers of anti-reflection (AR) coating consisting of alternating layers of differenta index material. With this technique we can obtain a reduction down to R<<10%. For example, in silicon photodiodes, a simple AR coating made of S_iO_2 (n ≈ 2) can reduce R down to 10%.

For the loss caused by the absorption in the neutral input layer, η_d has basically a cutoff at short wavelength, because photons are all absorbed in this layer before they reach the depleted layer. As we can see from the second term of η_d

$$\eta_D \;=\; P_d \;=\; (1-R)\cdot e^{-\frac{w_n}{L_a}}\cdot\left(1-e^{-\frac{w_d}{L_a}}\right)$$

the probability for the photon to escape the neutral layer is ruled by w_n/L_α. In silicon L_α is small at short λ; for example, at $\lambda = 400$nm we have L_α around 100nm.

The last source of η_d loss we deal with is, as we mentioned before, the incomplete absorption in the depletion layer, which gives a cutoff at long wavelengths.

The term of η_d involved in the absorption is the third one; it is ruled by w_d/L_α, and with

$$w_d/L_\alpha \ll 1$$

we get

$$\left(1-e^{-w_d/L_a}\right)\approx w_d/L_a$$

Also, in actual Si-PD structures the depth of the depletion layer (w_d) can go from one to tens of μm: combining this with the λ-dependence of L_α, the λ cutoff goes from about 900nm to 1100nm.

Current Si-PD can provide high efficiency in the visible spectrum; if we want to extend the range to longer wavelength, we have to use other semiconductors; for example, with Germanium the limit is up to 1500nm and we can reach 2000nm with InGaAs.

21.7 Dark Current

As seen in the previous chapter, dark current (I_B) is a relatively small electric current that flows through photosensitive devices, like photodiodes, even when no photons are entering the device; it consists of charges generated in the detector when no outside radiation is entering the detector. It is referred to as reverse bias leakage current in non-optical devices and it is present in all diodes. Physically, dark current is due to the random generation of electrons and holes within the depletion region of the device, and it is caused by thermal effects (actually, in device structures with high electric field it is caused also by tunnel effects).

Various physical phenomena take part in the carrier generation and recombination, with different relevance in the various cases, with different materials, device structures and operating conditions (bias voltage, temperature, etc.). For example, silicon has very favorable properties for achieving low generation rate; otherwise materials for IR detectors (Ge, InGaAs) have smaller energy gap and therefore inherently higher noise, since all generation processes are favored by a smaller E_G.

In this case we are dealing with both direct transitions and trap-assisted transitions; in this last case the electron in transition between bands passes through a new

energy state (localized state) created within the band gap by an impurity in the lattice.
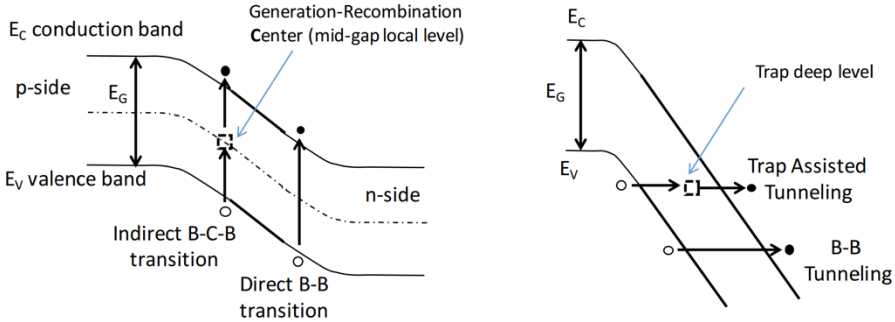
Figure 21.11: Thermal transitions (on the left) and transitions assisted by high electric field (on the right)

The localized state can absorb differences in momentum between the carriers, and so this process is the dominant generation and recombination process in silicon and other indirect bandgap materials.

The transition due to tunnel effect is possible only in high electric field because the conduction and valence bands have to be bended a lot, in order to have a transition with the tunnel effect.

The dark current contributes to the total system noise and it produces random fluctuations around the average photocurrent; it therefore manifests itself as shot noise on the photocurrent. The shot noise of I_B is the photodiode internal noise, with effective unilateral power density equal to

$$\sqrt{S_B} = \sqrt{2qI_B}$$

The internal noise of photodiodes with microelectronic size (with a sensitive area < 1mm^2) is typically much lower than the input noise of even the best high-impedance preamplifiers. In the applications of microelectronic photodiode, the circuit noise is dominant, just like for vacuum phototubes.

However, semiconductor PDs have dark current density j_B much higher than vacuum phototubes; this fact significantly limits the size of the active area of semiconductor detectors that can be employed for very low-noise operation.

In silicon devices, it is proven that the main contribution to dark current in reverse biased junctions with moderate electric field intensity is due to thermal generation of carriers in the depletion region. The thermal generation rate in the depletion layer has a volume density n_G given by

$$n_G = \frac{n_i}{2\tau}$$

where n_i is the intrinsic carrier density, equal to 1.45×10^{10} cm^{-3} at room temperature, and τ represents the minority carrier lifetime, which is strongly dependent on the device technology, like the material used and the fabrication process. The carrier lifetime is in the order of μs for ordinary Si technology in integrated circuits, ms for detector devices and in the range of 1-10 seconds for the best available Si technology for detector devices.

The last thing we discuss in this paragraph is the active area of Silicon Photodetectors. A Si-PD with circular active area of diameter D (so the area is A = $\pi D^2/4$) and depletion layer thickness w_d has dark generation rate of $n_B = n_G A w_d$; if we set a limit $n_B < n_{Bmax}$ the diameter of the active area has to be limited accordingly and we obtain

$$A < Amax \frac{n_{Bmax}}{n_G w_d} = \frac{2\tau n_{Bmax}}{n_i w_d}$$

$$D \leq Dmax \sqrt{\frac{8\tau n_{Bmax}}{\pi n_i w_d}}$$

21.8 Current Signal in PDs

The carriers drifting in depleted regions induce current at PD terminals, whereas carriers diffusing in neutral regions do NOT. The Shockley-Ramo (S-R) theorem is still valid in presence of space charge.

The Shockley–Ramo theorem allows us to easily calculate the instantaneous electric current induced by a charge moving in proximity of an electrode. It is based on the concept that current induced in the electrode is due to the instantaneous change of electrostatic flux lines which end on the electrode, not the amount of charge received by the electrode per second.

We saw the S–R theorem states that the instantaneous current i induced on a given electrode due to the motion of a charge is given by

$$i = E_v \, q \, v$$

where q is the charge of the particle, v is its instantaneous velocity and E_v is the component of the electric field in the direction of v at the charge's instantaneous position, under the following conditions: the charge is removed, the given electrode is raised to unit potential, and all other conductors are grounded. So, knowing the actual velocity v_c of a drifting carrier, the current induced at the PD terminals can be computed by the S-R theorem.
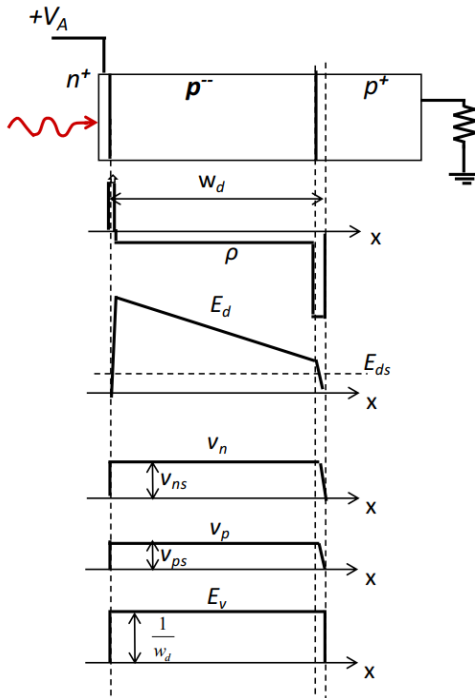
Figure 21.13:

Representation of some parameters used in this analysis, in relation to the cross-section of a typical PD structure

Figure 21.13a:

Space charge density ρ in the depleted region

Figure 21.13b:

Electric field E_d, which as we can see is greater than the saturation field E_{ds} over almost all w_d

Figure 21.13c:

Electron and hole drift velocity v_n and v_p, which are $\approx v_{ns}$ and $\approx v_{ps}$ over almost all w_d

Figure 21.13d:

Reference electric field E_v for S-R theorem, which is equal to $E_v = 1/w_d$
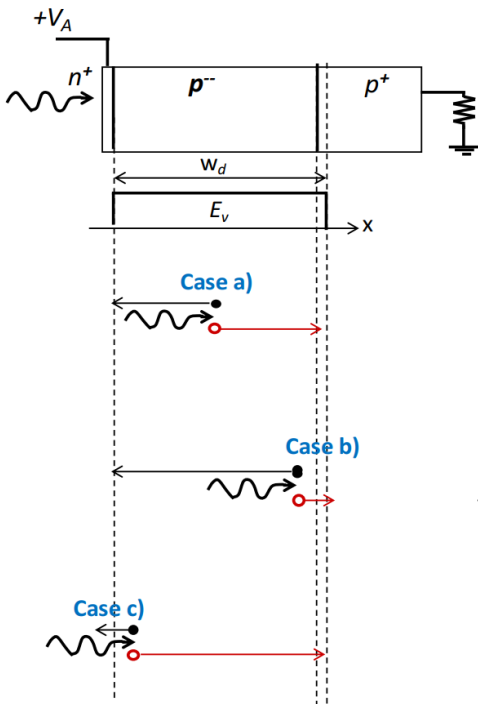
Figure 21.20:

Representation of the electron and hole current (in time) with different generation positions.
The current is $i_c = E_c q_c v_c$ for the S-R theorem, as we mentioned before

The motion of carriers in a semiconductor with electric field E_d is different from that in vacuum with equal E_d, because carriers suffer of scattering on the lattice and dissipate most of the energy received from the field due to the collisions. So, the drift velocity v_c has to be a function of the field E_d, and not the accelera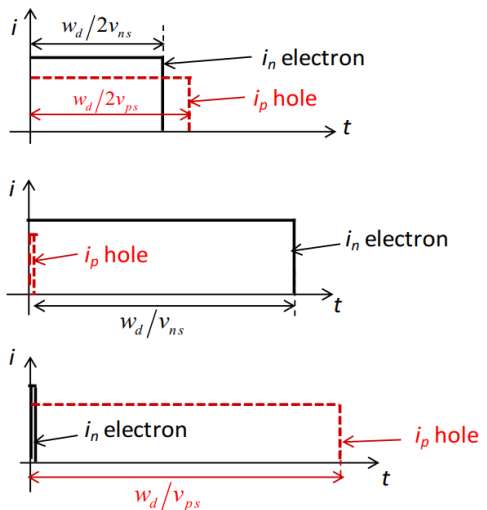tion. Moreover, in Silicon (and other materials) the motion of electrons is different from the motion of holes: at low field $E_d < 2$ kV/cm = 0,2 V/μm the regime is Ohmic and we have $v_c = \mu_c E_d$ (with electron mobility of $\mu_n \approx 1500$ cm^{2}V^{-1}s^{-1} and holes mobility of $\mu_p \approx 450$ cm^{2}V^{-1}s^{-1}). As E_d increases above 2kV/cm, the velocity rises progressively slower, at $E_{ds} \approx 20$kV/cm = 2V/μm the velocity saturates at the scattering-limited values of $v_{ns} \approx 10^7$ cm/s for electrons and $v_{ps} \approx 8 \times 10^6$ cm/s for holes, which are almost equal to the thermal scattering velocity $v_{th} \approx 10^7$ cm/s.

There is not much to say about what we see in the Figure 21.20: the currents in time clearly depend on how much the carrier has to move and, on its velocity,, so the duration of a single-carrier pulse is given by the transit time T_t of the carrier in the depleted region. At saturated velocity it is quite short; in silicon the carrier travels at $\approx$10ps/μm, and with $w_d = 1 \div 100$μm we have $T_t = 10$ps $\div$ 1ns.

The single-carrier pulse duration thus depends on the position of carrier generation. Strictly speaking, the waveform of the current due to a fast multi-photon pulse is not the convolution of the optical pulse with a standard carrier response, but it is a more complex computation that depends on the spatial distribution of absorbed photons. However, convolution with a suitable standard single-carrier response gives the waveform with approximation adequate for most cases, at least for times longer than the carrier transit time. A simplifying and conservative approximation currently employed for silicon PDs assumes as standard the response to an electron that crosses all the depletion layer. A finite width of response implies a low-pass filtering in light-to-current transduction, it's a mobile-mean over time $T_t = w_d/v_{sn}$, with upper band-limit $1/2T_t = v_{sn}/2w_d$.

It's worth noting the there is a trade-off on the choice of w_d: a long w_d is required for high quantum efficiency at long wavelength λ, while a short w_d is needed for ultrafast time response. Remark, however, that this is valid for front-illuminated junction and not with side illuminated junction.

CHAPTER 21

Photomultiplier Tubes

Among photodetectors, Photomultiplier Tubes play a main role as they are able to detect single photons impinging on their surface. In this chapter Photomultipliers main characteristics are analysed, with the aim of clarifying their working principle and features. A closer analysis will be carried out concerning their Single Electron Response, their gain and the minimum measurable signal.

22.1 Introduction: the rationale behind Photomultipliers

Low-power light measurements are the essence of several applications, such as spectroscopy, nuclear physics and biomedical diagnostics. However, using a photodiode or a phototube can be challenging that is because often the noise of the electronic circuit is dominant with respect to the noise of the detector, i.e. the dark current, and it limits the minimum detectable signal. (figure 22.1)
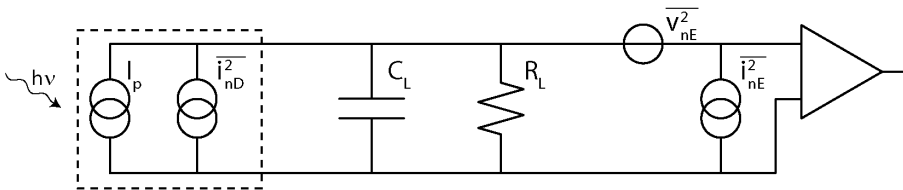
Figure 22.1: Vacuum Tube Photodiode signal and noise model

To enhance the Signal to Noise Ratio (SNR), a possible solution would be to amplify the signal, before it sums up with the circuit noise. Such a gain cannot be provided by an electronic circuit, as we would incur in the aforementioned problem, but we should change the structure of the photodetector itself. (figure 22.2)
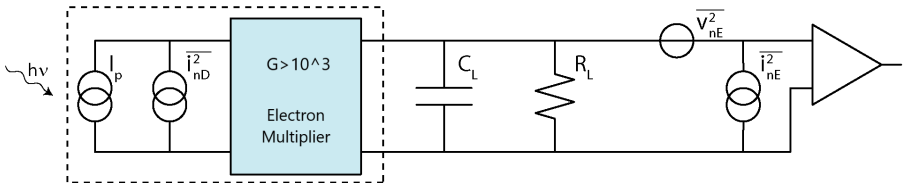
Figure 22.2: Principle of Photomultiplier tubes: through an electron multiplier a gain is added to the phototube without increasing the noise of the electronics.

The core idea is to multiply the primary electrons, i.e. the electrons emitted by the photocathode due to the incoming light: in this scenario, the output anode current would be higher than the photocathode current.

Ideally, following this approach, it would be possible to detect a weak light down to a single photon. In the next sections an overview of the physical principle is provided as well as a more detailed analysis concerning statistical description of Single Electron Response and gain. Moreover, the achievable Signal to Noise ratio is derived and further possible improved structures are presented.

22.2 Physical principle and Device structure

To increase the anode current without further increasing the noise of the electronics, an electron multiplier should be used.

An electron multiplier is a chain of dynodes, which are electrodes coated with a material having a high secondary emission coefficient. Secondary emission is a phenomenon where primary incident particles of sufficient energy induce the emission of secondary particles, when hitting a surface.

To be emitted, a secondary electron must have:

- $E_s \geq E_g + E_a$

    where E_s is the energy of the secondary electron, that should be sufficient to overcome the potential barrier, composed by the energy gap E_g and the electron affinity E_a ;
- A diffusion length larger than the distance from the point of generation to the surface

In case of a single photon reaching the cathode, a primary electron is emitted with a low kinetic energy ($E_e < 1eV$). The electron is driven in vacuum applying a high voltage to the dynode and impacts with high energy on its surface (figure 22.3). Energy is transferred to the electrons in the dynode: the ones that gain enough energy to overcome the potential barrier can escape and reach the anode.
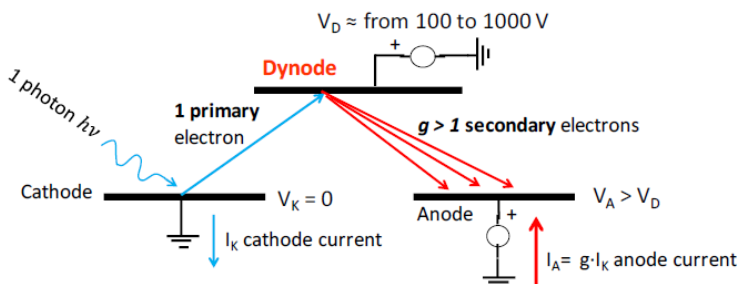
Figure 22.3: Basic structure of an electron multiplier, composed by a cathode, a dynode and an anode.

The secondary emission yield g is the number of emitted secondary electrons per primary electron. As the current is proportional to the number of charges, the anode current is the cathode current multiplied by the yield

This basic principle can be extended to a chain of dynodes, further enhancing the gain. (figure 22.4)
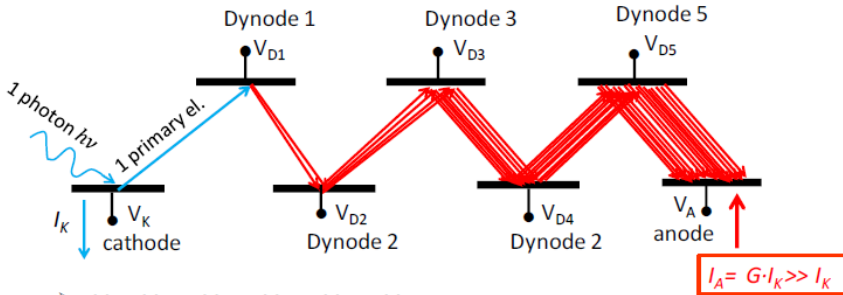

Figure 22.4: Electron multiplier involving five dynodes to achieve higher gain

Each dynode must be biased at a higher potential than the previous one in the chain, to drive the electrons towards it. The bias voltage is in the order of hundreds of volts, to ensure that the dynode is hit with energy suitable for multiplication. The product of dynodes gain is the overall multiplier gain G:

$$G = \prod_{k=1}^{N} g_k$$

Being the current proportional to the number of electrons, the anode current is:

$$I_A = G \cdot I_k$$

The yield of the multiplier is strictly related to the properties of the dynode. Hence, it is crucial to better understand the structure of these electrodes.

22.2.1 Dynode Materials

Dynodes are electrodes coated with materials that have a secondary emission yield higher than one: materials such as Magnesium Oxide (MgO), Beryllium Oxide (BeO), Cesium Antimonide (Cs_3Sb) are coated onto a substrate electrode made of nickel, stainless steel or copper-Beryllium alloy.

Moreover, semiconductor materials with negative electron affinity, as Gallium Phosphide, have been developed for this purpose, which make it possible to achieve higher coefficients. That is because a negative electron affinity decreases the potential barrier that secondary electrons need to overcome in order to be emitted.

The electron multiplication factor depends on the energy of the photoelectron, which in turn is determined by the voltage difference between each electrode: the choice of the bias voltage is therefore crucial.
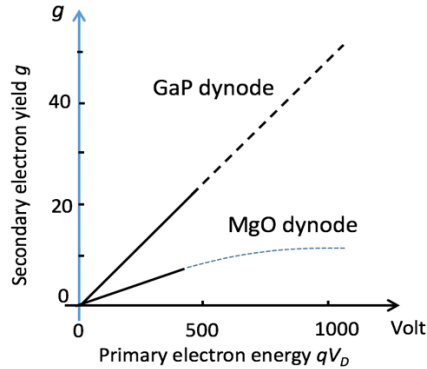
Figure 22.5: Secondary electron yield g versus accelerating voltage

Figure 22.5 shows the secondary emission yield g as a function of the accelerating voltage for the primary electrons V_D. Accelerating voltage can be substituted with primary electron energy, as it is defined as qV_D. In the plot is it possible to distinguish between ordinary yield, achieved by MgO dynode, and high yield, reached by GaP dynode.

For V_D up to 500 V, there is a linear relationship between the emission yield g and the accelerating voltage:

$$g = k_s V_D$$

As the voltage increases, the energy is transferred to electrons in deeper layers in the material, which have lower probability to escape in vacuum, and g saturates. The bias point is chosen to work in the linear range: ordinary emitters work with g values from ≈ 5 to ≈ 7, while GaP dynodes have g values from ≈ 5 to ≈ 25, but are more expensive and delicate. Moreover, GaP dynodes require special care in operation and over long operation times their yield tends to decrease.

Considering for example 12 dynodes and a yield $g = 3.2$ the overall gain G is 10^6, which is a typical value for a PMT.

Now let us consider the device structure of Photomultiplier Tubes

22.2.2 Photomultiplier Tubes device structure

The first Photomultiplier Tube was developed by RCA Laboratories in 1937: it was built with discrete dynodes and electrostatic-focusing. In the following years different companies contributed to improve these dectectors but the essentials elements are the same:

- a *photocathode* converting the incident light into electron flux;
- an *electron-optical input system* which focuses and accelerates the electron flux;

- an *electron multiplier* whose basic structure was described in previous sections;
- an *anode* collecting the electron flux.

Among photocathodes we can distinguish two main kinds:

- Opaque cathodes: electrons are emitted from the illuminated side.

- Semi-transparent cathodes: electrons are emitted on the opposite side to the incident light

These two kinds are used for different input windows and PMTs can be distinguished in Side-on Photomultiplier Tubes (figure 22.6) and End-on Photomultiplier Tubes (figure 22.7).
In Side-on multipliers, incident light enters from one side of the device; an opaque photocathode is therefore used to generate the signal, because the dynode chain is placed on the opposite side.
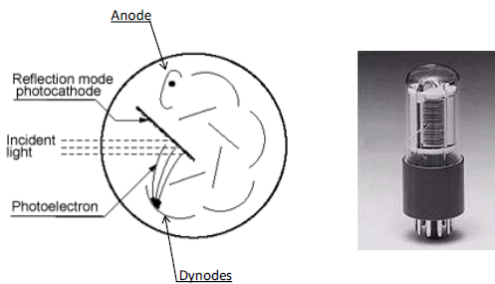

Figure 22.6: Side-on Photomultiplier tube

On the contrary, as shown by figure 22.7, End-on multipliers have an input window on one end of the tube: a semi-transparent photocathode is used as the dynode chain is placed behind the illuminated side of the cathode.
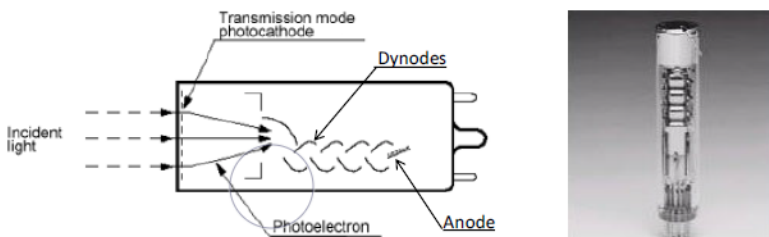

Figure 22.7: End-On Photomultiplier tube

Thanks to the huge dimensions of the photocathode, PMTs can offer a higher SNR than smaller solid-state detectors; indeed, when increasing the area the device collects a higher numer of photons N_p (if the incoming light is not spatially limited) and as a consequence the signal is increased by a factor N_p while the noise by a factor of $\sqrt{N_p}$ due to photon Poissonian distribution.

PMTs also have some drawbacks: first of all, they are quite fragile as the tube is made of glass, and vibrations need to be controlled; their fragile and bulky nature prevents them to be used in a wide range of applications where more robust and small detectors are needed. Moreover, mechanical designing is extremely challenging due to the high number of dynodes close together; finally the magnetic field has a very detrimental effect on the operation of the device so they cannot be exploited in some environments. Nevertheless, PMTs are still used in many fields.

Once the basic structure of the device is understood, let us focus on the dynamic response of the detector, through the analysis of its single electron response (SER).

22.3 Photomultiplier Dynamic Response: Single electron response pulses

The output current pulse for a single photoelectron is called the Single Electron Response or SER of the PMT. Figure 22.8 shows the current pulse given by a single photon hitting the cathode
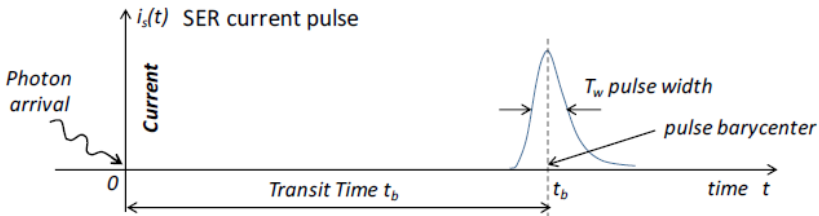
Figure 22.8: SER current pulse characterized by its transit time t_b and its pulse width T_w

Some observations can be made from the graph. As dynodes electrostatically screen the anode, only an electron emitted from the last dynode can reach the anode and generate a current pulse: a delay is therefore present and it is expressed as the delay of the pulse barycenter, called transit time (t_b).

This intuitive explanation is verified by the Shockley-Ramo Theorem, that states that the current i_c flowing into the output anode can be computed as:

$$i_c = q\overrightarrow{E_v} \cdot \vec{v}_c$$

Where $\overrightarrow{E_v}$ is the reference electrostatic field obtained grounding all electrodes except from the anode (at 1V) and $\vec{v}_c$ the speed of the electron.

Due to the dynode electrostatical screen, the reference electric field is nihil until the last dynode is reach, and so is the current.

It can be observed that transit time t_b is a statistical variable, that fluctuates from pulse to pulse. It is possible to define a transit time jitter T_j, as the full width at half-maximum FWHM of the probability density $p_b(t_b)$ (figure 22.9).

Figure 22.9: Transit time probability density characterized by transit time mean value $\overline{t_b}$ and its jitter T_j

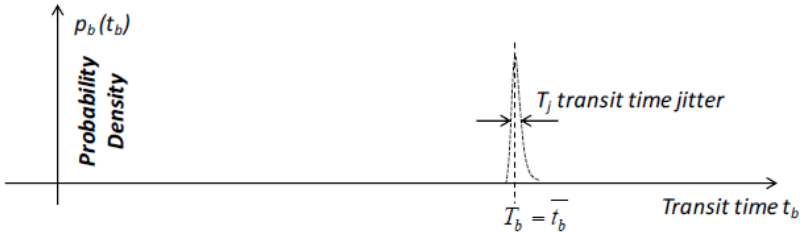Such fluctuations originate from the dispersion of transit time from one electrode to the following in the chain.

There are several causes of time dispersion, but let us have an insight on the main ones:

- Electrons are emitted from different positions on the electrode: as the distance from the electrode and the dynode is not the same from every point of the electrode, there are different transit times, depending on the single trajectory.
- Electrons have random initial velocity: even starting from the same position the transit time can fluctuate, depending on the initial speed of the emitted electrons.

To minimize transit time jitter, the geometry and the potentials of the electrodes should be carefully designed. Another key parameter is the SER pulse width T_w computed as the FWHM of the output current. Although it experiences some fluctuations, they are very small and practically negligible.

From figure 22.8 and figure 22.9 it can be inferred that the SER pulse width is 5 to 10 times wider than the transit time jitter.

Once the single electron response is known, it is possible to compute the PMT response $i_p(t)$ to a multi-photon δ-like light pulse.

While all the photons arrive at the same time at the photocathode, the secondary electrons reach the anode after a transit time that fluctuates. The total current at the output is the convolution of the transit time distribution and the SER:

$$i_p(t) = \int_0^\infty p_b(t_b) i_s(t - t_b) dt_b = p_b * i_s$$
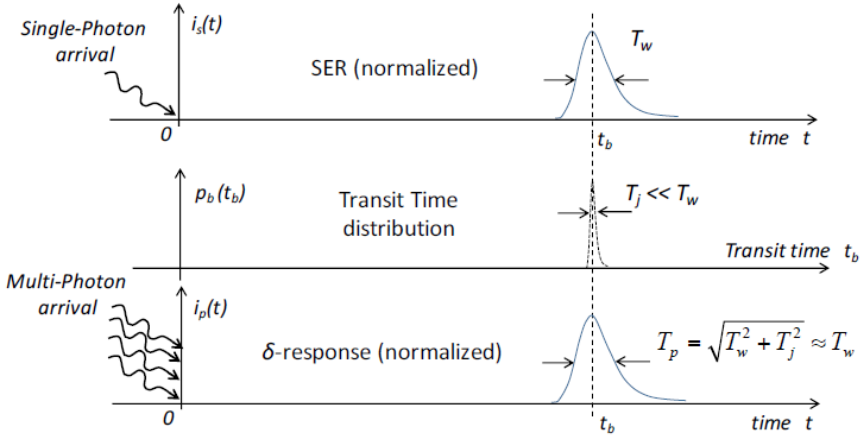
Figure 22.10: δ-response for multi-photon arrival

As the δ-response to a multi-photon arrival is the convolution of p_b and i_s, its FWHM T_p is the quadratic sum of the FWHMs, T_j and T_w, of the two terms:

$$T_p = \sqrt{T_w^2 + T_j^2}$$

Gathering T_w^2 and taking it out of the square root, we get:

$$T_p = T_w \sqrt{1 + \left(\frac{T_j}{T_w}\right)^2}$$

Remembering that $\sqrt{1+x} \approx 1 + \frac{x}{2}$ for Taylor Series and being $T_j \ll T_w$, $\frac{T_j}{T_w}$ is small (from 0.1 to 0.2) we obtain that the width of the δ-response is practically equal to the SER current pulse width:

$$T_p \approx T_w \left[1 + \frac{1}{2}\left(T_j/T_w\right)^2\right] \approx T_w$$

Being T_w finite, PMT employed as a current amplifier has a finite bandwidth f_p:

$$f_p = \frac{1}{k_a T_w}$$

The coefficient k_a depends on the SER pulse waveform (from ≈ 3 to ≈ 10).

22.4 Gain in-depth analysis

As explained in the previous sections, the core idea behind a PMT is the presence of an internal gain, enabling low-power light measurement until single photon detection.

Consequently, the gain parameter is highly critical: indeed, its variations can impair accurate measurements.

There are two main reasons underneath the fluctuations of the gain:

1) Electrical fluctuations: the gain is very sensitive to power supply variations

2) Intrinsic fluctuation: the emission of electrons both from cathode and from dynodes is not a deterministic function, but rather a statistical one, thus implying a random variation of the overall gain.

It is therefore of crucial importance to understand how to choose a proper voltage supply to stabilize the gain and how statistical fluctuations may impact on the device performance and Signal to Noise ratio.

22.4.1 Power supply requirements

Concerning gain regulation, the guideline to follow in order to attain a stable gain is the adoption of a remarkably stable power supply.

This principle can be inferred by looking at how gain and voltage supply are related. Let us consider for instance a single power supply V_A (typically from 1500 to 3000 V), which is partitioned in n voltages by means of a voltage-divider resistor chain. (figure 22.11)
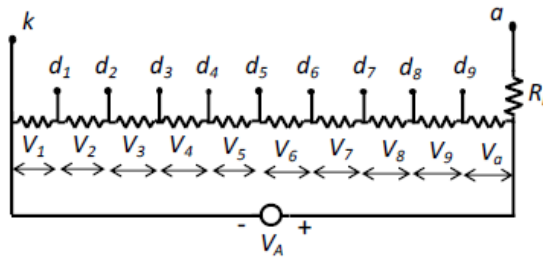
Figure 22.11: Bias circuit of the PMT: anode (a) and cathode (k) terminals are connected to the device, while each dynode is connected to an intermediate node of the voltage divider.

Each dynode is attached between two terminals and is biased at $V_j = f_j V_A$, where f_j is a pre-set fraction of the total voltage. The gain of a single dynode is proportional to the applied voltage and can be expressed as:

$$g_j = k V_j = k f_j V_A$$

Multiplying the gain of every single dynode we obtain the total gain G, which increases with the frequency in a non-linear way:

$$G = g_1 g_2 g_3 \dots g_n = (k f_1 V_A)(k f_2 V_A)(k f_3 V_A) \dots (k f_n V_A)$$

Gathering all the constants we obtain:

$$G = K \, V_A^n$$

Finally calculating the relative variation of the total gain with respect to the voltage:

$$\frac{dG}{G} = n \frac{dV_A}{V_A}$$

We can observe that little variations of the power supply voltage are amplified by the number of stages, thus causing higher gain variations. For instance if we are interested in a gain stability better than 1% we would require a voltage supply more stable than $1\%/12 = 0.08\%$.

22.4.2 Statistical fluctuations of gain

Let us focus now on the impact of statistical fluctuations on the signal and noise amplification of the photomultiplier.

When experimentally evaluating the SER of a Photomultiplier tube, a particular phenomenon arises: the pulses are different one from the other, in the sense that they present the same shape but with randomly varying amplitude (figure 22.12).
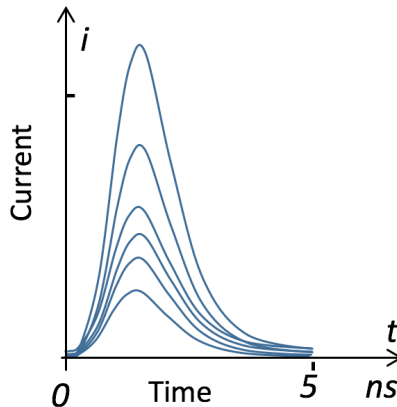
Figure 22.12: Single Electron Response experimental measurement: current shape is measured versus time; depending on gain statistical fluctuations, the amplitude of the SER randomly varies in different measurements.

As every response is generated by a single electron, we may derive that the changing parameter in this situation is the overall gain of the dynode chain. In order to get the distribution of this gain, it is possible to measure each pulse amplitude and classify them in a histogram, which is the discrete version of the gain statistical distribution (figure 22.13).
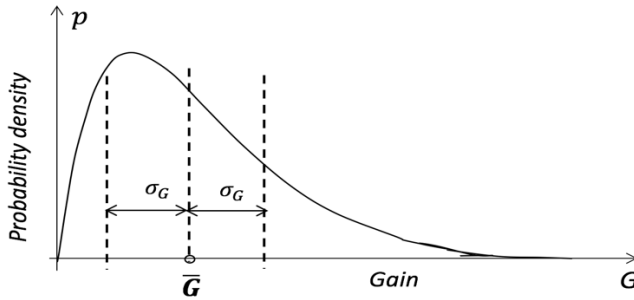
Figure 22.13: Gain statistical distribution: the graph represents the probability density of obtaining a certain gain in a single measurement versus the possible gain values.

The distribution is different changing the bias condition and it is not gaussian (as it is not symmetric with respect to the mean value) but skewed towards high gain values. The main parameter needed to describe this process are the mean gain $\overline{G}$, the gain variance σ_G^2 and the relative variance $v_G^2 = \frac{\sigma_G^2}{(\overline{G})^2}$.

To get a deeper insight on how the gain distribution affects the signal and the noise, we can divide the PMT operation in two phases:

1) Initially the photocathode emits primary electrons as in a phototube; the number of electrons emitted is described by means of a Poisson distribution, i.e. a statistical distribution with mean value N_p, variance $\sigma_p^2 = N_p$ and relative variance $v_p^2 = \frac{\sigma_p^2}{N_p^2} = \frac{1}{N_p}$.

2) It is then followed by a cascade of secondary impact emissions. The overall statistical distribution of these secondary emissions is the one previously depicted in figure 22.13.

In order to describe the statistics related to the number of output electrons, it is now useful to define another statistical variable, with mean value N_u, variance σ_u^2 and relative variance $v_u^2 = \frac{\sigma_u^2}{N_u^2}$. The aim is therefore to find a proper expression for N_u and σ_u^2.

N_u can be easily obtained considering that the two emission processes are independent, so that we can multiply the mean values:

$$N_u = N_p \ \overline{G}$$

For the variance σ_u^2 it is not mathematically correct to directly multiply the process variances, as we are dealing with statistical variables. It is consequently useful to resort to Laplace theory of probability generating functions, stating that: the relative

variance v_u^2 of the output of a cascade is the sum of the relative variance of each stage, divided by the product of the mean values related to all the previous stages.

$$v_u^2 = v_p^2 + \frac{v_G^2}{N_p} = \frac{1}{N_p} + \frac{v_G^2}{N_p} = \frac{1}{N_p} (1 + v_G^2)$$

Moreover, considering that:

$$v_u^2 = \frac{\sigma_u^2}{N_u^2}$$

and equating the two expressions, we get:

$$\sigma_u^2 = v_u^2 \, N_u^2 = \frac{1}{N_p} (1 + v_G^2) \, N_p^2 \, \overline{G}^2 = N_p \, \overline{G}^2 (1 + v_G^2) = \sigma_p^2 \, \overline{G}^2 (1 + v_G^2)$$

where $F = 1 + v_G^2$ is called *Excess Noise Factor*.
In conclusion, we obtain the subsequent formula, meaning that the variance of the noise due to primary emission (the same present in a phototube) is enhanced not only by a factor $\overline{G}^2$ as expected, but also by a factor F.

$$\sigma_u^2 = \sigma_p^2 \, \overline{G}^2 \, F$$

In most common cases the value of F is almost equal to 1.5-2, as will be further demonstrated below.

22.5 Operating mode and Signal to Noise Ratio

Once the statistical distribution of the output electrons has been investigated, it is of high interest to calculate the SNR and the minimum detectable signal.
It is useful to split the analysis in two separate discussions as the photomultiplier can be operated in Analog mode or in Single Photon Counting mode (also known as Digital mode).
Analog mode is commonly used with higher light level, where many photons are impinging on the photocathode, thus generating a high number of photoelectrons and output pulses which are no more separable one from the others.
Concerning the Digital mode, it operates at lower light level when single pulses can be identified and counted.

22.5.1 Analog mode

When dealing with Analog mode, the sensor and read-out electronics can be represented by the subsequent model (figure 22.15), where the signal is amplified by means of a preamplifier and then filtered with a low pass filter.
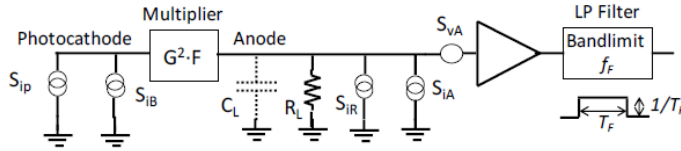
Figure 22.15: Model of PMT sensor and its read-out electronics when operating in Analog mode.

The elements considered in the model are:
Signal:

$$I_p = n_p q \text{ where } n_p \text{ is the photoelectron rate at photocathode}$$

Noise sources:

- Due to the physics of the device:

  o $S_{i_p} = 2qI_p = 2q^2 n_p$, shot noise of the photocurrent, increasing with signal
  o $S_{i_B} = 2qI_B = 2q^2 n_b$, shot noise of background, independent from signal
  o $S_{i_D} = 2qI_D = 2q^2 n_d$, shot noise of dark current, independent from signal

- Due to the preamplifier:

  o S_{V_A}, voltage noise spectral density of the preamplifier
  o S_{i_A}, current noise spectral density of the preamplifier

- Due to the load resistor R_L:

  $$S_{i_R} = \frac{4kT}{R_L}, \text{thermal noise of resistor}$$

Load resistor R_L, typically includes cable resistance, device resistance and externally added resistor.

Load capacitor C_L, taking into account the capacitance of the device, the capacitance of the preamplifier and cable stray capacitances.

Multiplier G^2F, indicating the intrinsic gain of the PMT (note that signal is increased by a factor G, instead noise variance by a factor G^2F,)

Low pass filter after the preamplifier: a gated integrator was chosen as an example, thus implying a bandwidth limited to $f_F = \frac{1}{2T_F}$.

A first expression of the Signal to Noise Ratio may be now derived:

$$\frac{S}{N} = \frac{I_p G}{\sqrt{\left[2q\left(I_p + I_B + I_D\right)G^2 F + S_{i_A} + \frac{S_{v_A}}{R_L^2} + S_{i_R}\right]f_F}} = 1$$

Since the gated integrator bandwidth is much smaller than the $R_L C_L$ counterpart, i.e. $\frac{1}{2\pi R_L C_L}\frac{\pi}{2} \gg f_F = \frac{1}{2T_F}$, the low pass filtering effect of $R_L C_L$ has been neglected for noise. Concerning signal, it is supposed to have a bandwidth much smaller than the one of the $R_L C_L$ parallel, so that the pulse shape is not affected.

The minimum signal is reached when $SNR = 1$, i.e. when signal is comparable to noise. Therefore if it is not possible to establish a priori which noise contribution is dominant, a second order equation needs to be solved in order to derive the minimum signal. Anyhow, it is often possible to neglect preamplifier and resistor noise, as the intrinsic gain makes device noise dominant with respect to the other contributions.

The expression is consequently simplified to:

$$\frac{S}{N} = \frac{I_p G}{\sqrt{2q\left(I_p + I_B + I_D\right)G^2\,F\,f_F}} = 1$$

In order to further simplify computations two practical cases may be taken into account:

1) Background and dark current noise are negligible with respect to photocurrent noise:

$$\frac{S}{N} = \frac{I_p G}{\sqrt{2q\,I_p\,G^2\,F\,f_F}} = 1$$

$$I_{p-min} = 2q\,F\,f_F$$

$$N_{p-min} = \frac{I_{p-min}T_F}{q} = \frac{2q\,F\,f_F T_F}{q} = F$$

As obtained, the minimum number of detectable photoelectrons on a time window T_F depends on F, which is a really small number $1.5 - 2$ compared to the performance achieved by a phototube. It is therefore confirmed the possibility to detect lower light values leveraging a PMT.

2) On the other hand photocurrent noise could be negligible:

$$\frac{S}{N} = \frac{I_p G}{\sqrt{2q\ (I_B + I_D)\ G^2\ F\ f_F}} = 1$$

$$I_{p-min} = \sqrt{2q\ (I_B + I_D)\ F\ f_F}$$

In the intermediate case the second order equation needs to be solved:

$$I_p^2\ (G^2) + I_p(-2qG^2\ F\ f_F) - 2q\ (I_B + I_D)G^2\ F\ f_F = 0$$

Once the minimum current has been calculated it is possible to derive the minimum power by means of:

$$P_{p-min} = \frac{I_{p-min}}{S_D} \quad \text{and} \quad S_D = \eta_D \frac{\lambda[\mu m]}{1.24[V \mu m]}$$

Where S_D is the radiant sensitivity of the device, λ the operating wavelength and η_D the quantum efficiency which can be found on photocathode datasheets as for a phototube.

22.5.2 Single Photon Counting mode

Let us focus now on Single Photon Counting mode, which can be performed exploiting the conceptual scheme visible in figure 22.16
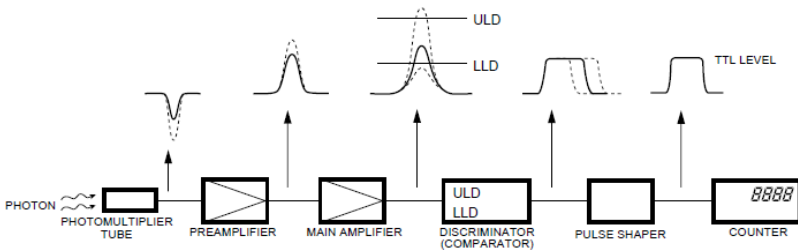

Figure 22.16: Photon counter system

Photons impinging on the photocathode generate single pulses which are amplified by the preamplifier and if necessary, by the main amplifier. These output pulses are then directed into a comparator featuring two thresholds: pulses higher than upper

threshold (ULD) and pulses lower than lower threshold (LLD) are eliminated. The pulse shaper cleans the pulses, allowing the counter to count them.

Concerning Signal to Noise Ratio, in this situation the photomultiplier gain can be set so high that noise coming from the preamplifier is surely negligible.

As we are performing digital counting, it results:

$$\frac{S}{N} = \frac{N_p}{\sqrt{N_p + N_B + N_D}}$$

where $N_p = n_p T_F$, $N_B = n_b T_F$ and $N_D = n_d T_F$ are respectively the number of signal, background and dark current electrons and T_F is the measurement time.

As discussed for the Analog mode, background and dark current noise may be negligible, thus implying one single event can be detected:

$$\frac{S}{N} = \frac{N_p}{\sqrt{N_p}} = 1$$

$$N_{p-min} = 1$$

On the other hand, in case background noise is not negligible, it is possible to take advantage of background subtraction, meaning that measurement is executed twice: the first time with signal and the second one without signal (background measurement involving only N_D and N_B). Once both measurements have been acquired, the number of pulses of the background and dark current are subtracted from the total measurement.

In this way the Signal to Noise ratio is slightly modified, as at the denominator we have to take into account that, when the baseline is subtracted, the variances of the noise acquired in the two measurement are not subtracted, but added as the two noises are uncorrelated.

Consequently, we need to add a correcting factor:

$$\frac{S}{N} = \frac{N_p}{\sqrt{N_p + (N_B + N_D)\left(1 + \frac{T_F}{T_{FS}}\right)}}$$

where T_{FS} represents the background measurement time. If $T_{FS} = T_F$ background noise is doubled, if $T_{FS} \gg T_F$ the doubling effect is negligible and the correcting factor can be omitted.

22.6 Progress in Photomultiplier structure

As already mentioned, the main drawbacks concerning PMTs are their dimensions, fragility and sensibility to vibrations. A possible structure that solves those issues is the Continuous Channel Multiplier (CCM), represented in figure 22.17.

A CCM is a special glass capillary tube, featuring a diameter smaller than $1mm$ and acting both as voltage divider and as electron multiplier. Its inner surface is chemically treated and converted in a semiconductor layer, providing high resistivity and secondary electron emission yield $g \approx 1.2 - 3$.
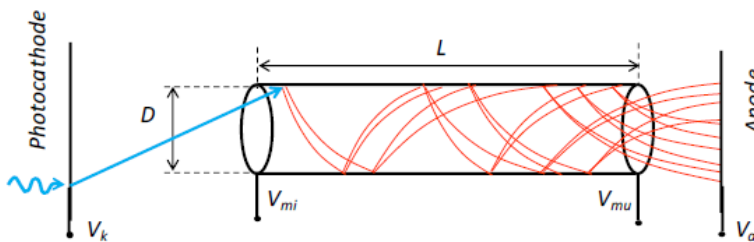

Figure 22.17: Continuous channel multiplier CCM structure, featuring a diameter D and a length L

When a photon strikes the photocathode, an electron is emitted and focused to the inner surface of the tube. Due to multiple impacts, many secondary emissions are triggered, giving rise to an electron multiplication.

For a fixed applied voltage the gain is dependent on the ratio L/D, as increasing this ratio the number of impacts increases. However, as far as the number of impacts is augmented, the electron energy decreases, thus implying a decrease in the yield.

A maximum gain G of $10^5 - 10^6$ is attained with $L/D \approx 50$ and an applied voltage in the range of $2000 - 3000\ V$.
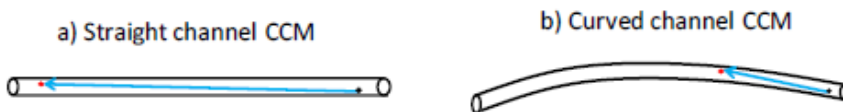

Figure 22.18: Straight channel CCM (a) and curved channel CCM (b)

It is not possible to adopt a CCM with the basic structure previously described. Indeed, Ion Feedback effect may cause a positive feedback enhancing current amplification in an uncontrolled way and generating a self-sustaining breakdown current.

This phenomenon consists in the generation of free heavy ions in the last part of the channel, where energy is higher, by collision between electrons and residual gas molecules. These free ions drift in the electric field causing a strong electron emission by impacting on the capillary wall. If this happens in the first part of the channel, the electrons undergo all the channel multiplication.

In order to avoid this effect, it is possible to bend the glass capillary as in figure 22.18. The free ions are therefore no more able to reach the channel input, as their trajectory is almost straight.

The main features of a CCM can be resumed as follows:

- Dark current is lower than in common PMTs
- The excess noise factor F is bigger than 2, which is significantly higher than dynode-PMTs due to the greater statistical dispersion in CCM electron multiplier;
- The amplification of a pulse leaves a charge on the multiplier surface, near to the output. As a consequence, a high charge modifies the electric field, thus impairing the amplification of the following pulses during a long recovery transient. This effect can be controlled, limiting the product of the multiplier gain and input pulse charge and/or repetition rate.
- Space charge can give rise to non-linearities for high pulses and high gain.

For all these reasons CCM-PMTs are suitable to detect pulses with moderate repetition rate and small size, down to single photons, while they offer worse performance for many-photon pulses and for stationary light intensity.

In order to solve part of these issues, Micro-Channel Plate Multiplier (MCP) were developed (figure 22.19). Their basic structure is composed by an array of thousands of multiplier microtubes, embedded in beehive structure into a plate. The two faces of the plate are biased with a high voltage V_D through metal electrodes, thus all channels are in parallel.

Thanks to their planar geometry they can be matched to a planar end-window photocathode. Photoelectrons are focused from cathode to multiplier input by simply applying a high voltage.
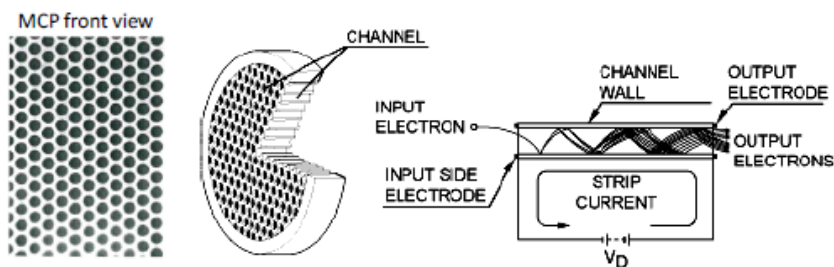

Figure 22.19: Micro-Channel Plate Multiplier (MCP) front view (on the left) and basic scheme

MCP features a small diameter $D \approx 5\mu m - 50\mu m$ and an active area, which is the sum of the channel input sections, almost equal to $50 - 60\%$ of the total plate area. As every channel works as an individual miniaturized CCM, the optimum gain is still achieved with a ratio $L/D \approx 50$

Concerning Ion feedback, it is avoided by mounting in series MCPs with channels having different inclination of their axis, thus forming an angle as in figure 22.20.
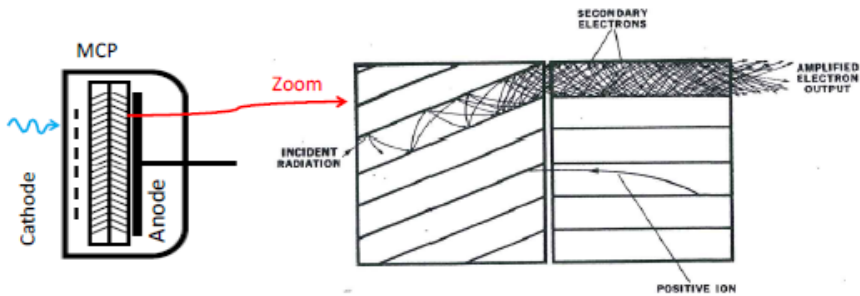

Figure 22.20: MCP in series to avoid Ion Feedback

Due to the presence of multiple channels, MCPs structure allows to relax some limitations that were introduced by CCMs.

Firstly, this is due to the fact that electrons emitted from the same position in the photocathode do not enter all in the same microchannel, but they are distributed over a group of facing channels in the MCP. Moreover, the perturbation of the voltage distribution in a channel affects the multiplication and collection of electrons just in that channel and in the closest ones.

For these reasons some differences can be noticed with respect to CCMs. First of all, the limit to the output mean signal current is much higher than in CCMs and it is a small percentage of the total bias current of the MCP, not of a single microchannel.

Furthermore many-photon optical pulses can be processed correctly, since the pulse photoelectrons do not enter a single channel, but they are processed in parallel, in different microchannels.

On the other hand MCPs statistical gain distribution is similar to that of CCMs with an excess noise factor F higher than 2. Additionally, MCPs have a remarkably superior dynamic response than dynode PMTs: indeed, transit time and jitter are reduced down to $T_b \approx 1ns$ and $T_j \approx 10ps$. The SER pulse-width T_w is shorter, around few hundreds ps.

CHAPTER 22

CHAPTER 23

Avalanche Photodiodes

Avalanche photodiodes (APDs) are highly sensitive semiconductor electronic devices that can not only exploit the photoelectric effect to generate a current signal from light, but they also provide a built-in gain through avalanche multiplication. To this aim, a relatively high reverse voltage has to be applied to enhance the electric field and let the carriers gain sufficient kinetic energy to impact-ionize.

23.1 Introduction

We have seen in the previous chapter that one of the best features of PMTs is their capability to provide an internal gain, i.e. photogenerated carriers are multiplied within the detector, giving rise to an amplification of the current before it mixes with electronic-circuit noise. The internal gain is clearly advantageous for the SNR. Following the same idea, a smaller, less bulky and less fragile alternative to PMTs was developed: the Avalanche Photodiode (APD).

APDs basically have the same structure of photodiodes (PDs) but different behavior. Let's recall that PDs are essentially p-n junctions operated in reverse-bias conditions, meaning that the n-doped region is at a higher potential than the p-doped region. As
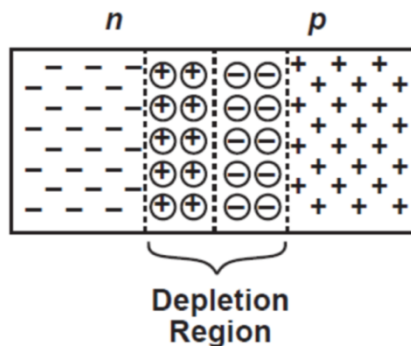
Figure 23.1: Depletion region of a pn junction

a result, a depletion region is formed at their interface, as shown in figure 23.1. If a photon is absorbed in the depletion region, the free electron-hole pair that is generated

gives rise to a current at the detector terminals; otherwise, a photon that is absorbed in the neutral region typically does not cause any current flow.

The main difference between a PD and an APD is that in the first one each photon can generate at most a single electron-hole pair; on the contrary, the absorption of a photon in an APD can trigger an avalanche multiplication process. The design of an APD starts from the basic structure of a PD, but some advanced modifications have to be introduced in order to overcome some issues or simply to achieve high performance. For these reasons, several structures have been proposed in the literature, such as p-i-n APDs, which feature an intrinsic region between the n- and the p-doped region or the Reach-through APDs (RAPDs), which are p-i-n structures with an additional small p-type layer between the n-doped and the intrinsic region.

APDs can be fabricated exploiting different materials: p-n junctions are usually implemented using silicon, but some alternatives can be useful for different reasons, e.g. to modify the wavelength range, the spectral response, the speed, or the energy gap properties. Thus, as well as in the implementation of LEDs and lasers, junctions fabricated using Ge, GaP, GaAs, InGaAs or other composites will be considered.

23.2 Impact Ionization

Considering the easiest implementation, the APD consists of a simple p-n junction reversely biased as depicted in Figure 23.1. The absorption of a photon in the depletion region generates an electron–hole pair (EHP): the electric field here separates the two carriers and drifts them towards n- and p-side, respectively. A free electron drifting in the field gains kinetic energy $\Delta E_n = E_n - E_c$. In an APD, the electric field in the depletion region is high so the carrier can gain a high energy. In particular, if $\Delta E_n > 1.5 E_g$ a ionizing collision can occur, meaning that another EHP is released.

As represented in Figure 23.2, this generation is a cascade process: in fact, the secondary electrons will be still accelerated by the high electric field and they can impact-ionize as well, generating other EHPs (internal gain mechanism). The cascade of ionizing collisions produces the **avalanche multiplication** of carriers.

It is worth noting that, not only electrons, but also holes participate in the impact-ionization process, meaning that an inherently-positive feedback is present. We will see that an important design guideline would be to guarantee the multiplication of only one carrier (either electrons or holes), to avoid the positive feedback. Indeed, even if this mechanism would increase the gain of the device, it is undesirable for several reasons: it is a random process and therefore increases the device noise and it can be unstable, thereby causing avalanche breakdown.

23.3 Ionization Coefficients

Even though the above described avalanche is an intrinsically discrete statistic process (in fact the generation of an EHPs is a phenomenon that can happen or not), to easily describe it we would like to address this topic in terms of continuous probabilities. In order to do that we have to make the following assumption: the width

of the multiplication region (where the high field is present) must be larger than the mean path covered by both electrons and holes between each collision.
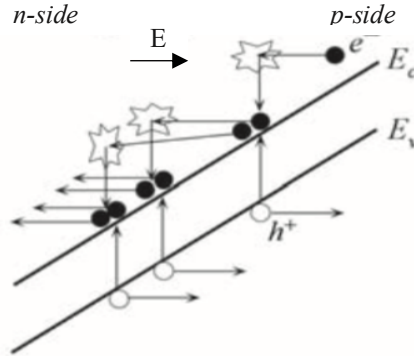

Figure 23.2: Impact ionization process

Under this assumption, we can define α and β coefficients (sometimes in literature they are also called α_n and α_p respectively) as the ionization probabilities per unit length (rate of ionization, 1/cm) for electrons and holes respectively (α and β are thus probability densities). The inverse coefficients, $1/\alpha$ and $1/\beta$, represent the average distances between consecutive ionizations. The ionization coefficients increase with the depletion-region electric field (since it is responsible for the acceleration) and decrease with the device temperature. If we increase the temperature, in fact, we have more thermal agitation, hence more collisions, reducing the opportunity that a carrier gains enough energy to ionize. The following relationships hold:

$$\alpha = \alpha_0 \exp\left(-\frac{F_{no}}{F_e}\right) \qquad \text{for electrons}$$
$$\beta = \beta_0 \exp\left(-\frac{F_{po}}{F_e}\right) \qquad \text{for holes}$$

In Silicon α_0 = 3,8*10^6 cm^{-1}, F$_{no}$= 1,75*10^6 V/cm; β_0 = 2,25*10^7 cm^{-1}, F$_{p0}$ = 3,26*10^6 V/cm.

Finally, the ratio k = β /α is defined: it is a very important parameter for the characterization of an APD. Its value indicates which is the carrier who contributes the most in the avalanche: as already said, we would like to have k proximate to 0 if we want electrons multiplication or k close to infinite if we want holes multiplication. Instead if the value of k is around 1, it means that both the carriers are equally contributing to the build-up of the avalanche.

23.4 Ionization Integral and Breakdown Voltage

First of all, we will derive the basic ionization integral which determinates the breakdown voltage condition; we can easily understand that, to achieve an infinite

current due to the avalanche process, we need to increase the reverse voltage, thus the electric field, up to a limit value, which is called breakdown voltage.

Let's assume that a current I_{po} is flowing at the left-hand side of the depletion region (as represented with j_i in figure 23.3); let's consider $x = 0$ at the beginning of the depletion region at the p side and $x = w$ at the end of the depletion at the n side, being w the length of the space charge region.
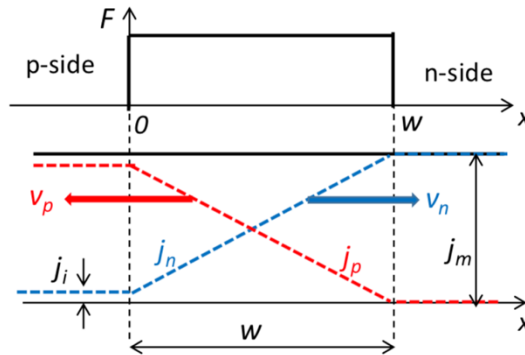

Figure 23.3: Electric Field across the junction and current densities

If the electric field in the depletion region is high enough to generate EPHs by the impact ionization process, the hole-current I_p will increase with distance along the depletion region and reach a value $M_p I_{po}$ at $x = w$.

Similarly, the electron current I_n, will increase from $I_n(w) = 0$ to $I_n(0) = I - I_{po}$, where the total current

$$I = I_p - I_n$$

is constant at steady state. The **continuity equation** (carrier balance in dx) is

$$\frac{dj_n}{dx} = -\frac{dj_p}{dx} = \alpha(x)j_n(x) + \beta(x)j_p(x)$$

Taking into account that $j_m = j_n + j_p$ we obtain the equation for j_n

$$\frac{dj_n}{dx} - [\alpha - \beta]j_n = \beta j_m$$

The functions $\alpha(x)$ and $\beta(x)$ are known: they are obtained by computing the field profile $F_e(x)$ and employing the known $\alpha(F_e)$ and $\beta(F_e)$.

The equation can then be integrated with $j_n(w) = j_m$ and with $j_n(0) = j_i$ as boundary condition.

In the simplest **case α = β** (e.g. in GaAs) the equation is simply

$$\frac{dj_n}{dx} = \alpha \, j_m$$

By integration we get

$$j_m - j_i = j_m \int_0^w \alpha(x)\,dx$$

and finally

$$j_m = \frac{j_i}{1 - \int_0^w \alpha(x)\,dx} = \frac{j_i}{1 - I_i}$$

$$I_i = \int_0^w \alpha(x)\,dx$$

This is called *ionization integral* and it has a clear physical meaning: it is the probability for a carrier to have an ionizing collision in the path from x=0 to x=w. The current j_m is the primary current j_i amplified by the multiplication factor M

$$M = \frac{j_m}{j_i} = \frac{1}{1 - I_i}$$

In **cases with** $\alpha \neq \beta$ the equation can still be integrated with $j_n(w) = j_m$ and with $j_n(0) = j_i$ as boundary condition. The results can still be written in the form

$$M = \frac{j_m}{j_i} = \frac{1}{1 - I_i}$$

but the ionization integral I_i is now the integral of an effective ionization coefficient α_e

$$\alpha_e = \alpha \exp\left[-\int_0^w (\alpha - \beta)\,d\xi \right]$$

so that in this case

$$I_i = \int_0^w \alpha_e(x)\,dx = \int_0^w \alpha \exp\left[-\int_0^w (\alpha - \beta)\,d\xi \right] dx$$

The ionization integral I_i in any case strongly depends on the applied bias voltage V_a and on the temperature T. I_i is nil until the field F_e produced by V_a attains level sufficient for impact ionization then rises with V_a, with a slope that strongly depends

on the actual field profile *F(x)*. The rise of I_i is always steep and produces in all cases a steep rise of M with V_a. Computations and experiments show that the slope of M gets steeper and steeper as the high-field zone gets wider. This is quite intuitive, since a wider zone corresponds to a higher number of collisions, which enhances the effect of the increased impact ionization probability due to an increase of the electric field. When the applied bias voltage V_B reaches a characteristic value V_B, the Ionization Integral $I_i \rightarrow 1$ and, according to the equation, M $\rightarrow \infty$ and $j_m \rightarrow \infty$: V_B is called **Breakdown Voltage.** It is a characteristic feature of the diode, ruled by the distribution of the electric field F_e and by the dependence of α and β on the electric field F_e and on the temperature T.

In particular, V_B increases with the temperature T. The dependence is different in devices with different field profiles. It is anyway strong, some 0,1% per K degree. For Si, it is about ≈30 mV/K in devices with V_B = 30 V and ≈900 mV/K in devices with V_B = 300 V.

Actually, the breakdown current is not divergent. We have seen that the current is self-sustaining, because of the positive feedback intrinsic in the avalanche ionization process. However, what keeps finite the avalanche current is the feedback effect due to the mobile space charge. The effect is negligible for $V_a < V_B$ (hence it is not taken into account in the former equations), but it is enhanced by the current rise at $V_a > V_B$ and reduces the electric field that acts on the carriers. The multiplication thus stabilizes itself at the self-sustaining level.

For $V_a > V_B$ the avalanche current increases linearly with V_a, so that an avalanche resistance R_a can be defined:

$$R_a = \Delta V_a / \Delta i_a$$

In fact, ΔV_a produces a proportional increment of the electric field, which increases the impact ionization probability, hence the avalanche current. In turn, the current rise produces an increase of the space charge, which counteracts the effect of ΔV_a. The current thus rises until it brings back to self-sustaining condition the avalanche multiplication; that is, the current increase ΔI_a is proportional to the voltage increase ΔV_a.

To summarize, an Avalanche PhotoDiodes (APD) is basically a photodiode biased at V_a below the breakdown voltage V_B but close to it where it provides a linear amplification of the current by exploiting the avalanche carrier multiplication.

The amplification gain is the multiplication factor M, which can be adjusted by adjusting the bias voltage V_a with respect to V_B. Since V_B strongly depends on the diode temperature T, variations of T have effect equivalent to significant variations of the bias V_a. Therefore, for having a stable gain M, the temperature of the APD must be stabilized.

A very steep increase of M with V_a is unsuitable for accurate and stable control, since small variations of V_a produce large variations of M. A gradual increase of M with V_a is preferable.

The actual dependence of M on V_a can be fitted fairly well by an empirical equation

$$M = \frac{1}{1 - \left(V_a / V_B\right)^u}$$

with exponent u that depends on the field profile (and on the type of semiconductor); it varies from 3 to 6, with higher values corresponding to wider high-field zone.

23.5 Reach-through APD (RAPD)

To improve the basic structure of the APD, we can think about significantly increasing the Electric Field over the active area so the electrons would gain consequently a greater kinetic energy, aiming to a better multiplication efficiency.

The Reach-through APD (RAPD) solution is to add a very thin p layer between the p- and the n-doped regions, so that it will be completely depleted, reaching the so-called reach-through condition. A simplified scheme of a Si reach-through APD is shown in Figure 23.4: close to the illumination window (on the left), we have a thin n^+ layer and then three p-type layers of different doping levels, to suitably modify the field profile across the diode. The first layer is a thin p-type one, the second one is a thick and slightly p-type doped (almost intrinsic) layer - which is indicated as π - and the third one is a heavily doped p^+ layer. The diode is reverse biased to increase the field in the depletion regions. The net space charge distribution across the diode is shown in Figure 23.5. When a sufficiently-high reverse bias is applied, the depletion region in the p layer widens to reach through to the π layer (that is why it is called the reach-through). The electric field is given by the integration of the net space charge density. The variation in the field across the diode is shown in Figure 23.5: E is maximum at the n^+p junction interface, then it decreases slowly through the upper p layer, that will be the avalanche region.
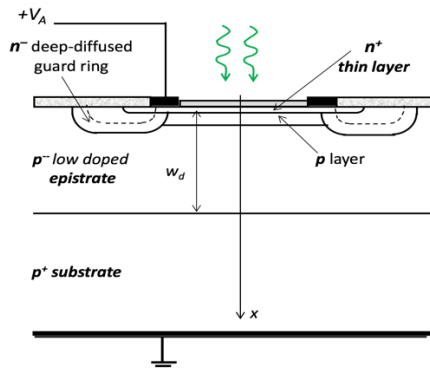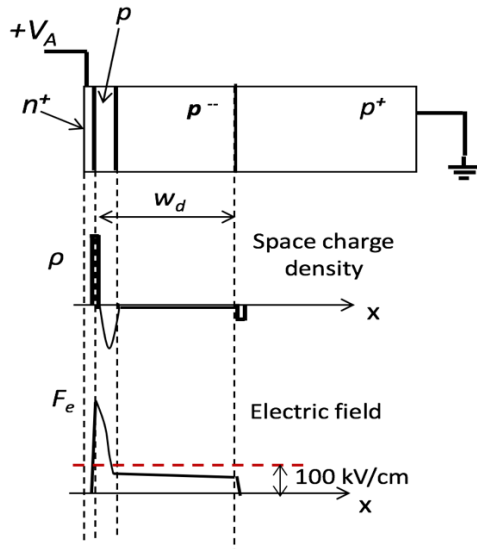

Figure 23.4: Simplified scheme of a Si reach-through APD

Figure 23.5: The net space charge distribution and electric field across a Si reach-through APD

Through the π layer, the field decreases only slightly as long as the net space charge density here is small. The absorption of photons takes place mainly in the long π layer. The reason for keeping the photogeneration within the π region and reasonably separate from the avalanche p region is that, avalanche multiplication is a statistical process and hence leads to carrier generation fluctuations, which lead to a noise in the photocurrent.

This is minimized if impact ionization is performed by the carrier with the highest impact ionization efficiency, which in Si is the electron, which means that the generation of electrons is dominant. For this reason, the structure in Figure 23.4 allows only the photogenerated electrons to flow and reach the avalanche region but not the photogenerated holes.

The total depletion layer width of Si RAPDs in most cases is from 10 to 30μm, in order to obtain high detection efficiency up to 800-900nm wavelength (NIR edge) but the width of the multiplication region (where F exceeds the ionization threshold) is much thinner, from 1 to a few μm. Thanks to this design, a moderately-steep rise of M with the bias voltage is obtained: the RAPD gain can thus be reliably controlled. The dependence of M on the device temperature is still remarkable also in this structure and it must be taken into account.

Also with RAPD, the highest M obtained with Si-APDs is much lower than the gain level currently provided by PMTs. In the best cases M values up to about 500 are obtained; attaining M=1000 is out of the question.
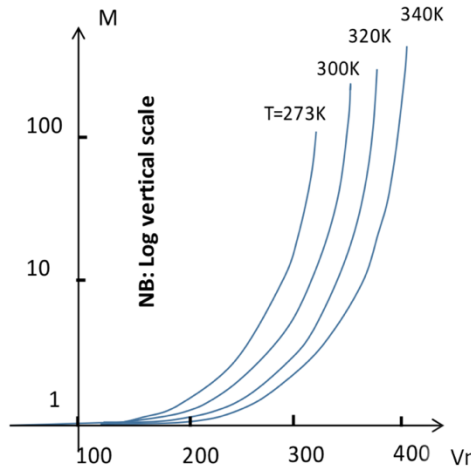
Figure 23.6: Multiplication factor vs. reversed voltage and temperature variation

23.6 Multiplication Noise in APDs

Avalanche multiplication is a statistical process so the APD gain has random fluctuations that has to be taken into account in noise calculation. Let us denote by

M the mean multiplication gain
σ_M^2 the gain variance
$v_M^2 = \sigma_M^2 / M^2$ the gain relative variance (defined as [variance]/[mean]^2)

Input: primary carriers with mean number Np
variance $\sigma_P^2 = Np$ (Poisson statistics)
Output: multiplied carriers with
mean number $Nu = M\,Np$
variance σ_u^2

In the multiplication, the fluctuations of the number of primary charges are not only amplified by M^2; they are further enhanced by a factor $F>1$ called Excess Noise Factor (like for PMTs). At the multiplier output, the relative variance is higher by the factor F than the input relative variance

$$\sigma_u^2 = F\,M^2\sigma_p^2 = F\,M^2 N_p$$

and so relative variance

$$v_u^2 = \sigma_u^2 / N_u^2 = F\,v_p^2 = F/N_p$$

Primary carrier generation and avalanche multiplication are processes in cascade.

The output relative variance v_u^2 is expressed by an equation (obtained from the Laplace probability generating function of cascaded processes) stating that v_u^2 of the cascade is the sum the relative variance of each stage divided by the mean gain of the preceding stages. In our case

$$v_u^2 = v_p^2 + \frac{v_M^2}{N_p} = \frac{1}{N_p} + \frac{v_M^2}{N_p} = \frac{1}{N_p}\left(1 + v_M^2\right)$$

Hence

$$\sigma_u^2 = N_p^2 M^2 v_u^2 = N_p M^2 \left(1 + v_M^2\right) = \sigma_p^2 M^2 \left(1 + v_M^2\right)$$

The excess noise factor thus is ruled by the relative variance of the gain, like for PMTs

$$F = 1 + v_M^2$$

However, there is a fundamental difference with respect to PMTs. APDs have relative variance v_M^2 definitely higher than PMTs already at low M level; furthermore, the situation gets progressively worse as M is increased, because also v_M^2 increases, thereby increasing F.

This behavior limits the highest APD gain M that can be usefully exploited. The physical processes exploited for multiplying electrons in PMTs and in APDs are remarkably different and the detector gain has remarkably different features.

In PMTs:

- the accelerated electron that hits a dynode is lost and the number of emitted secondary electrons fluctuates in a set of values that includes zero. The resulting mean number of carriers coming from the dynode is just the mean number of emitted secondary electrons and is definitely *higher than unity*.
- the gain is produced by a unidirectional sequence of events, the cascade of statistical multiplications at the various dynodes. Cascaded statistical processes can be well analyzed by known mathematical approaches (as the Laplace probability generating function)

In APDs:

- the accelerated electron that undergoes an ionizing impact is not lost, it remains available for further impacts; the generation of a further electron (plus a hole) is statistical and the mean number of generated electrons is definitely *lower than unity*. The resulting mean number of electrons after the impact is *one plus the mean number of generated* electrons.
- the statistical process is much more complicated than a simple cascade because of the intrinsic *positive feedback* in the impact-ionization. Rather than a cascade, it is a complex of interwoven feedback loops, each one originating from the other type of carrier (the hole in our case) generated in the impact.

Let's try to analyze a standard case: a Silicon APD with electric field of moderate intensity, just above the ionization threshold. In this situation the ratio of ionization coefficients is very small $k = \beta/\alpha < 0,01$ so the probability of impact ionization by holes is much lower than that of electrons.

The mean number μ of secondary electrons generated by the impact of an electron is small $\mu \ll 1$, the positive feedback in the multiplication is so small that it can be neglected. The multiplication can thus be analyzed as a unidirectional cascade of electron impacts. By employing the Laplace probability generating function and giving a progressive index to the impacts in sequence we get

$$v_M^2 = v_{m1}^2 + \frac{v_{m2}^2}{m_1} + \frac{v_{m3}^2}{m_1 m_2} + \ldots$$

and with equal impacts (i.e. uniform electric field)

$$v_M^2 = v_m^2 + \frac{v_m^2}{m} + \frac{v_m^2}{m^2} + \ldots = v_m^2 \left[1 + \frac{1}{m} + \frac{1}{m^2} + \ldots \right] \approx v_m^2 \frac{1}{1 - 1/m} = v_m^2 \frac{m}{m-1}$$

that is

$$v_M^2 = v_m^2 \frac{1+\mu}{\mu} = \frac{1}{1+\mu}$$

Since $\mu \ll 1$ it is $v_M^2 \approx 1$ and we get

$$F = 1 + v_M^2 = 1 + \frac{1}{1+\mu} \approx 2$$

$F=2$ is the lowest possible F for Si-APDs and is achieved at low gain level. The conclusion is confirmed by experiments on carefully designed APD devices operating at $M<50$. For comparison, recall that ordinary PMTs routinely offer $F < 2$ at very high gain $M>10^5$.

Silicon with electric field just above the ionization threshold is a particularly favorable case. In all other cases, the positive feedback in the avalanche process is remarkable, it cannot be neglected and has detrimental effect on the variance of the APD gain. The fluctuation of the electrons generated in an impact is not only amplified by the further electron impacts in the subsequent multiplication path. The holes that are generated in the impact travel back and re-inject the fluctuation in a previous step of the multiplication path. This back-injection of fluctuations enhances the excess noise factor F, with an efficiency that increases with the k factor (the relative ionization efficiency of holes versus electrons). So, in Silicon the k factor markedly increases as the field is increased. Therefore, F markedly increases as the bias voltage of the APD is raised for increasing the gain.

The primary photocurrent (injected and amplified in the high field zone) can be carried by electrons (as in the device examples previously shown) or holes (as in devices with inverted polarity of the semiconductor layers). At a given gain M, the *F* value is lower if the forward branch is given by the most efficient carriers (electrons) and the backward branch in the loop is given by the less efficient carriers (holes). In conclusion: the best results with Silicon APDs are obtained with primary current carried by electrons.

A thorough mathematical treatment of the avalanche multiplication is quite complicated and beyond the scope of this course. We will just comment some results of treatments reported in the literature.

With some simplifying assumptions (uniform electric field; constant k value), it has been shown that the excess noise factor *F* with primary current of electrons is

$$F \approx M\left[1-(1-k)(1-1/M)^2\right]$$

In cases with negligible positive feedback k=0, the equation confirms the result of the previous approximated analysis.

In cases with full positive feedback (i.e. equally efficient carriers, as in GaAs and other III-V semiconductors) it is k≈1 and F increases as M

$$F \approx M$$

Finally, in cases with intermediate feedback level it is 0<k<1 and the equation specifies how F increases with M with rate of rise that increases with k.

The gain *M* of the APD is intended to bring signal (and noise) of the detector to a level higher than the noise of the following circuits, with the aim of attaining better sensitivity than a PN photodiode (which is limited by the circuit noise). However, when the voltage is raised to increase *M* also the variance of the gain fluctuations increases. At some level M_{max} the effect of the gain fluctuations becomes greater than that of the circuit noise: increasing M beyond this level would be nonsense. This M_{max} limit depends on the actual case (actual APD and circuit).

It is the maximum factor F_{max} tolerable in the case of interest that actually determines the M_{max} level. In critical cases (typically InGaAs APDs, which have F≈M) a fairly high value F_{max} turns out to be tolerable, even up to $F_{max} \approx 10$.

Thanks to the low k factor, Silicon devices have the lowest excess noise among APDs and achieve the highest gain levels.

Si-APD devices specially designed for low k have:

$$F \le 2,5 \text{ up to } M \approx 100$$
$$F \le 5 \quad \text{up to } M \approx 500.$$

Ordinary Si-APD devices have fairly lower performance, i.e. typically:

$$F \le 4 \quad \text{up to } M \approx 100.$$

APDs are currently developed with different materials, for example to obtain better detection efficiency in the near infrared range. Germanium and III-V semiconductors are typical examples. In Germanium, holes ionize more efficiently than electrons, i.e. $k > 1$. Instead of k, the reciprocal $1/k$ must be considered, and it is advisable to have the primary current carried by holes. $1/k$ in Ge-APDs increases with the electric field in a way similar to k in Si-APDs, but the values of $1/k$ in Ge are higher than k in Si. Fairly high useful gain has been demonstrated with Ge-APD devices, typically:

$$F \leq 4 \ \text{ up to } M \approx 50.$$

On the other side, in III-V semiconductors (GaAs, InP, InAlAs, etc.) the ionization efficiencies of electrons and holes are equal ($k=1$) or at least comparable ($k \approx 1$). The positive feedback thus is very strong, and F increases as M.

For InP-InGaAs and other III-V devices the useful gain range is quite limited, typically:

$$F \leq 10 \text{ up to } M \approx 10$$

Nevertheless, InGaAs-APDs are in general preferred to Ge-APDs for detecting IR optical signals because they have lower dark-current (lower detector noise) and higher quantum detection efficiency, with cut-off extended to longer wavelengths (typically $\lambda \leq 1{,}7$ μm).

CHAPTER 23

Single – Photon Avalanche Diodes

Is it possible to reach single photon sensitivity with an integrable device? APD is not the most appropriate detector: indeed, we saw that its gain is not high enough in order to reach this goal. To bypass such limitation, a new kind of detector has been introduced: the Single Photon Avalanche Diode (SPAD). Its main feature is that it can work biased above the breakdown voltage of the p-n junction, in the so called Geiger-mode.

24.1 Introduction

During the last decades, special semiconductor detectors, called Single-Photon Avalanche Diodes (SPADs), were developed; some examples are shown in figure 24.1. These devices, also called "Geiger-mode avalanche photodiodes" or "triggered-avalanche detectors", have been worked up and elaborated in order to detect single photons.
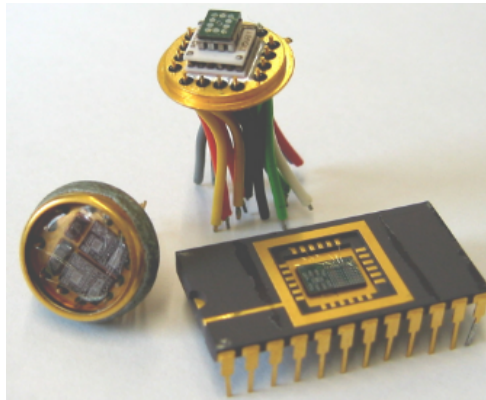
Figure 24.1: Examples of SPAD detectors and its integrated circuits

Basically, SPADs consist of a p-n junction reverse biased at a fixed overvoltage with respect to the breakdown voltage. When a photon impinges on the surface, it can be absorbed and it can generate an electron-hole pair. The two carriers can be accelerated by the electric field, and if any of them reaches a sufficient level of energy, it can trigger an avalanche as a consequence of the impact ionization phenomenon. Thanks to the positive feedback, a macroscopic current can be generated, and detected, in response to the absorption of a single photon.

SPADs can be used in several fields: basic quantum mechanics, cryptography, astronomy, single molecule detection, luminescence microscopy, biology, chemistry, material science, fluorescent lifetime decays, optical fiber testing. Currently available devices are characterized by good properties such as low dark count rate, high photon detection efficiency (PDE) and excellent temporal resolution. All these features are addressed in the following paragraphs.

But first, an advanced analysis of the limited sensitivity and limited Single-Photon Counting (SPC) capabilities of APDs in linear amplifying mode must be considered.

24.2 APD limitations: sensitivity and SPC

Before discussing the features of SPADs, let us discuss why there was the necessity to develop a new detector to achieve single-photon sensitivity. Which were the main limitations of APD in Single-Photon Counting? Let's discuss them in detail.

In order to achieve single-photon sensitivity, a high gain is required. In fact, the output of an APD is given by the number of primary (photogenerated) carriers multiplied by the gain. If we use a comparator with a threshold to discriminate the detection events, it is quite clear that the output of the APD output pulse due to a single-photon must be at least higher than the threshold, which is lower-limited by noise. Unfortunately, the gain of APD is a statistical process; therefore, it introduces some additional noise. We have seen in the previous chapter that, in order to limit the F factor, i.e. the parameter that takes into account random fluctuations of the gain, and so to avoid an increment of the noise rms value, the gain must be kept low.

In general, to achieve true single-photon sensitivity, two main requirements must be met:

- Efficient rejection of noise, i.e. low probability of *false detections* due to the noise
- Efficient detection of photon pulses, i.e. high probability of detecting the single-photon pulses, which can have variable amplitude with large statistical fluctuations

With noise amplitude having gaussian distribution (most frequent case) with variance σ_n (rms value), the noise rejection threshold level must be at least $N_{nr} \geq 2{,}5\sigma_n$, in order to keep below <1% the probability of false detection. This is graphically recalled in figure 24.2.
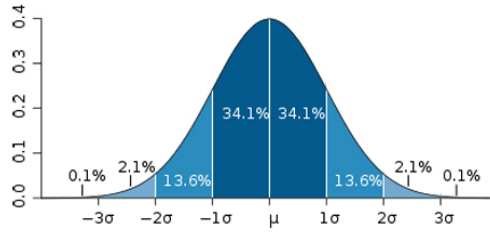
Figure 24.2: Gaussian distribution percentage: values higher than $2,5\sigma_n$, occur with a probability lower than 1%.

We have seen in previous chapters the typical acquisition scheme for a high-impedance sensor, like photodetectors. So, considering that measurement chain, recall that the rms noise (in number of electrons) with an optimum filter can be computed as follows:

$$\sigma_n = \frac{\sqrt{2C_L\sqrt{S_V}\sqrt{S_I}}}{e}$$

Where e^- is the electron charge. Reasonable values for the capacitance and the electronics noise are:

$C_L \approx 0{,}1$ to 2 pF $\rightarrow$ load capacitance

$\sqrt{S_v} \approx 2$ to $5 \; \frac{nV}{\sqrt{Hz}} \rightarrow$ series noise

$\sqrt{S_I} \approx 0{,}01$ to $0{,}1 \; \frac{pA}{\sqrt{Hz}} \rightarrow$ parallel noise

With high-quality preamp we typically get $\sigma_n \approx 40$ to 120 e^-. Taking some margin and considering $2,5\sigma_n$, we need at least $M \approx 100$ to 300, which is not easily feasible with APDs. Furthermore, M just higher than the minimum is not sufficient for having single-photon pulses higher than the threshold: we will see that a *much higher M is necessary.*

Even without variations, the minimum APD gain that is needed can be higher than what we just computed, in case the setup exploits an input resistor R_L; indeed, the resistor can introduce a dominant contribution to the noise. To have a better insight into this issue, let's consider a typical amplification scheme, as the one reported in figure 24.3. The pulse signal coming from the sensor is read-out with a load resistance R_L. C_L takes into account all capacitive contributions, both due to the sensor and the acquisition circuit. Considering some typical values as $R_L=1k\Omega$ and $C_L= 2$ pF as an example, we get a pole with $T_A= R_L C_L$. Noise is due to the preamplifier and resistor, but considering some typical values for the preamplifier, the noise is usually dominated by the chosen resistor value ($1k\Omega$).
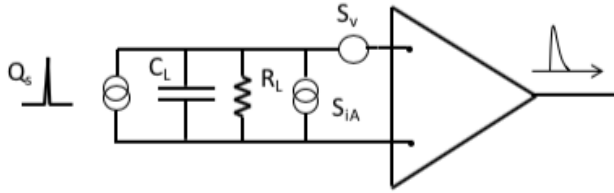
Figure 24.3: Circuit configuration for noise and signal analysis

Considering an amplifier with the same bandwidth of the $R_L C_L$ network, in this configuration there are two poles at the same frequency. Let's now proceed with the signal and noise analysis.

The transfer function due to the presence of the two poles is:

$$H_b = \frac{1}{(1+jwT_A)} \frac{1}{(1+jwT_{pa})}$$

Considering $T_A = T_{pa}$, the two poles are at the same frequency and so in the Laplace domain:

$$H_b = \frac{T}{(1+sT)^2}$$

Computing the Fourier Transform of H_b, the expression of the filter in time can be obtained:

$$h(t) = \frac{t}{T} e^{-\frac{t}{T}}$$

With δ-like single-photon detector pulse of charge Q_s, the voltage output signal is equal to

$$V_s = \frac{Q_L}{C_L} \frac{t}{T_L} e^{-\frac{t}{T_L}}$$

having a maximum at $t = T_L$:

$$V_s = \frac{Q_L}{C_L} \frac{1}{e}$$

The noise is dominated by the resistor in this case

$$\overline{v_n^2} \approx S_{iR} R_L^2 \frac{1}{8T_L} = 4KTR_L \frac{1}{8T_L}$$

Therefore the S/N results:

$$\frac{S}{N} = \frac{V_S}{\sqrt{\overline{v_n^2}}} = \frac{Q_S}{C_L} \frac{1}{e} \frac{\sqrt{8T_L}}{R_L\sqrt{S_{iR}}} = \frac{Q_S}{C_L} \frac{1}{e} \frac{\sqrt{8T_L}}{\sqrt{4KTS_{iR}}}$$

The rms noise referred to the detector output in terms of charge is

$$\sqrt{\overline{q_n^2}} = \frac{e}{\sqrt{8}} \sqrt{S_{iR}}\sqrt{T_L}$$

And in electron number:

$$\sigma_n = \frac{\sqrt{\overline{q_n^2}}}{q_{el}} \approx 1055 \text{ e}^-$$

with this acquisition scheme, the necessary noise-rejection threshold level is

$$N_{nr} \approx 2.5\sigma_n \approx 2600 \text{ e}^-$$

This would require a gain that is extremely high for an APD.

The exploitation of the optimum filter allows us to maximize the SNR; nevertheless, there are some other aspect that must be taken into account. Recall that in the acquisition chain considered at the beginning of this discussion, the system parameters set the bandwidth of the optimum filter (low pass), with a time constant T_{nc} (due to the preamp noise and the input capacitance) that typically ranges from a few nanoseconds to some hundreds of nanoseconds. As a result, output pulses are fairly long and this brings drawbacks.

In fact, the goal is not to detect just one photon but to detect more than one. A *count loss* can occur because of the overlap of the tail of one pulse and the following peak of another pulse. As sketched in figure 24.4, when the time interval between two photons is shorter than the output pulse width, pulse pile-up occurs (i.e. the two pulses overlap), the comparator is triggered only once and only one count can be recorded.
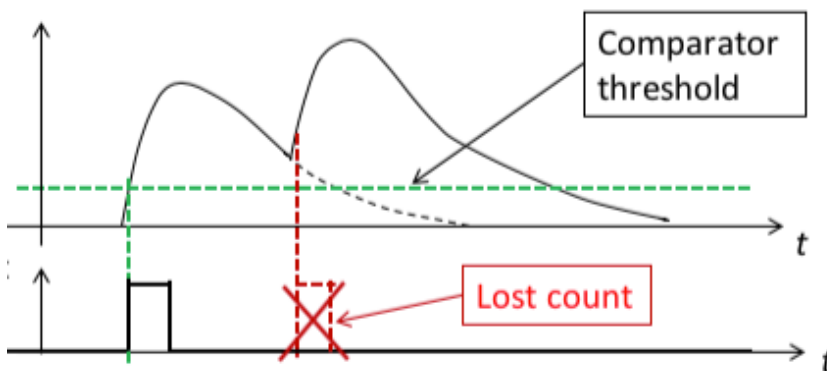
Figure 24.4: Top graph: single photon output of an approximated optimum filter. Bottom graph: comparator output fed to the counter
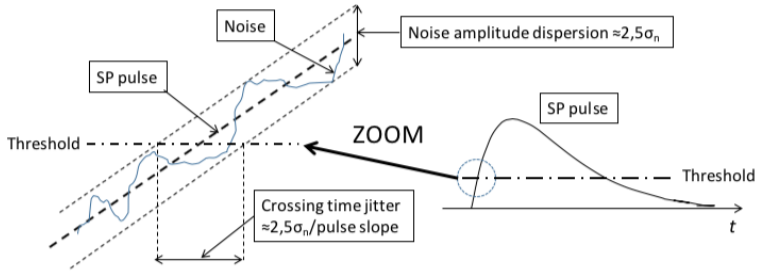
Figure 24.5: Noise amplitude dispersion on a single pulse signal

Photons occur randomly, so pile-up can occur; this situation causes a missing photon event and therefore a distorsion of the recorded signal. It is quite intuitive that the percentage of lost counts increases as the pulse-width is increased.

To reduce count losses, we need to reduce the duration of the pulse tail; this could be achieved exploiting a filter with a higher bandwidth with respect to the optimum filter. Nevertheless, this solution would cause the collection of a higher amount of noise. We see that there is a tradeoff between speed, which can translate into signal loss in this case, and noise. Moreover, in several applications also the arrival time of the signal is a parameter of interest, leading to another drawback. Considering the same readout scheme as before, i.e. the signal resulting from a single photon is detected by means of a comparator, the problem is noise and speed. Indeed, if we have a quite slow signal, the superimposed noise causes a crossing of the threshold with different delays causing the so-called *time jitter*. This scenario is shown in figure 24.5. As a consequence, observing with high precision ($\approx$ ps) the instant of arrival of the photon is not straightforward. The time jitter is inversely proportional to the derivative of the signal, which means that increasing the BW could reduce the jitter (figure 24.5). This new parameter takes part into the previously considered trade off. In order to reduce count-losses and time jitter, the APD pulses must be processed with a high filter bandwidth, typically higher than the optimum filter. This implies higher noise and, as consequence, a higher threshold value and a higher gain required to the APD. Moreover, M just higher than the noise rejection threshold level N_{nr} would be enough if the APD gain were constant for all SP pulses, but this is not the case.

Since the gain M has strong statistical fluctuations, a high excess noise factor F>>1, which is directly related to the relative variance of M, must be considered.

$$F = 1 + v_M^2 = 1 + \sigma_M^2 / \left(\overline{M}\right)^2$$

Thus

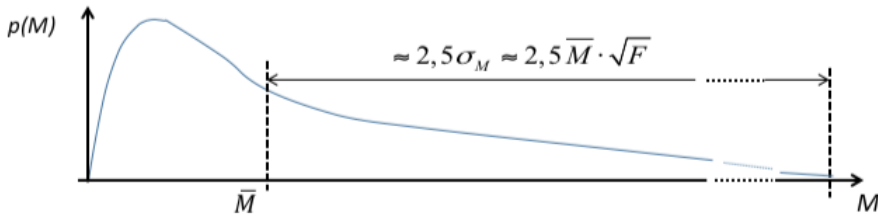$$\sigma_M = \overline{M}\sqrt{F-1} \approx \overline{M} \cdot \sqrt{F}$$

Figure 24.6: Statistical distribution of the gain M

In figure. 24.6 it can be seen that the statistical distribution of the gain M is not Gaussian (as it is not symmetric with respect to the mean value $\bar{M}$), but it is skewed in favour of low-gain values. So the probability to have a gain higher than the mean value is lower than the probability to have a gain lower than $\bar{M}$. Therefore, with a mean gain $\bar{M}$ just above the noise rejection threshold, a major percentage of the SP pulses is rejected. This downgrades the photon detection efficiency, i.e. the basic performance of the detector.

In order to limit the reduction of detection efficiency due to the threshold, the mean gain $\bar{M}$ should be higher than the noise rejection threshold N_{nr} by a factor G>>1.

In the most favorable case (special Si-APD with optimum filtering), the value of $\bar{M}$ necessary for attaining the noise rejection threshold N_{nr} is near to the maximum available APD gain, but there is still some margin. In other cases (regular Si-APDs with wideband electronics) there is no margin at all.

In conclusion, photon counting with linear amplifying APDs is possible only with special Si-APDs and with photon detection efficiency strongly reduced with respect to that obtained with the same APDs by measuring the analog current signal.

24.2 SPAD: features and generalities

The term Single-Photon Avalanche Diode (SPAD) defines a class of photodetectors able to detect low intensity signals (down to the single photon) and to acquire the time of the photon arrival with high temporal resolution. Its most important features are analyzed in the following paragraphs.
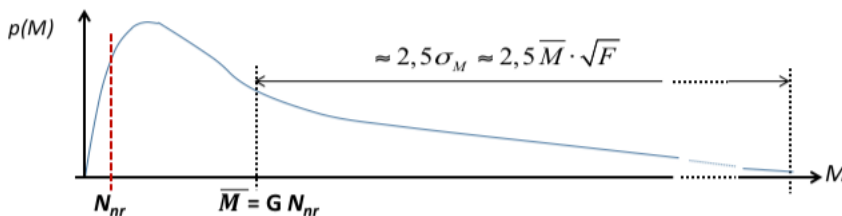

Figure 24.7: Statistical distribution of the gain M, with the value of $\bar{M}$ that should be chosen

24.2.1 I-V characteristic

The I–V characteristic of a p-n junction engineered to work above the breakdown voltage is sketched in figure 24.8. As you can see, it is not a steady state curve: in fact, in the breakdown operating condition it has a bistable behavior. When reverse voltage above the breakdown is applied and no free carriers are present in the diode depletion region, the detector operates at $I_a=0$ (blue line). As soon as a carrier is injected into the depletion region, impact ionization may cause an avalanche, and the diode will follow the steady-state curve (red line).

It is interesting to discuss how it is possible to trace the I-V characteristic above reported. SPAD devices, as all photon detector, are affected by dark count; this means that it is not possible to trace its bistable characteristic by simply applying a constant voltage. If we tried to use that approach, only the static characteristic (red line) can be detected because the avalanche is for sure activated due to the dark count rate. In order to detect also the curve in which $I_a=0$ (blue line) a fast and repetitive voltage scanning with a duration of about 10ms must be applied. In this scenario, we have that in some periods of the scanning no electrons are generated, thus, the avalance is not activated and is it possible to trace the branch $I_a=0$ (blue line); in some other periods an avalanche is triggered and the steady-state condition (red line) can be traced as well.

Note that the current I_a has been considered positive in the cathode to anode direction. For this reason the reverse bias characteristic is placed in the first quadrant.

24.2.2 Geiger-mode

The fundamental difference between SPADs and APDs is that SPADs are specifically designed to operate with a reverse bias voltage well above the breakdown voltage. This kind of operation is also called Geiger-mode. In order to be able to operate in this regime, a diode should have uniform properties over the sensitive area: in particular, it must be free from defects causing local field concentration and lower breakdown voltage (the so-called microplasmas, due to metal precipitates, higher dopant concentration, etc.
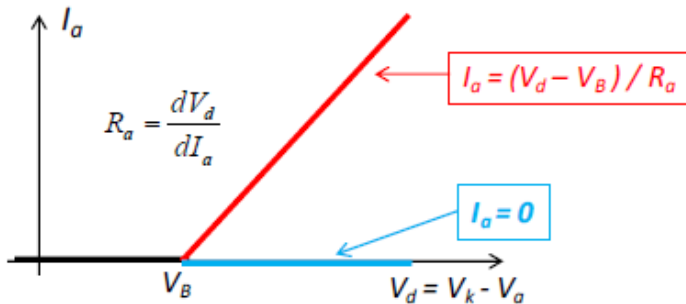

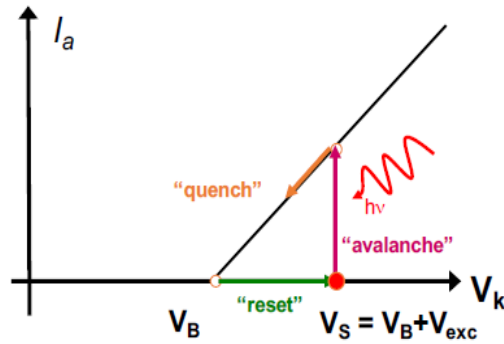**Figure 24.8:** Bistable behavior of I-V characteristic

Figure 24.9: Geiger mode operation

Bias supply voltage V_s exceeds the breakdown voltage V_B of the junction by an amount called *excess voltage* V_{exc}, which has a major influence on the detector performance. When a photon arrives, the avalanche is triggered and a macroscopic current flow into the device. Then the avalanche must be quenched by pulling down V_d to (or below) V_B in a passive or active way (see sections. 24.3, 24.4). The quenching delay has to be minimized in order to limit the aferterpulsing effect (see section 24.6). Finally, the diode voltage V_d must be restored to V_s to make the sensor ready to detect another signal (Fig.24.9).

Pulses are produced in SPADs also by the spontaneous thermal generation of single carriers in the diode junction and constitute a dark count rate (DCR) similar to that observed in PMTs. Low DCR is a basic requirement for an avalanche diode to be employed as SPAD.

Various parameters characterizing the detector performance strongly depend on the excess bias: probability of avalanche triggering, hence the photon detection efficiency; amplitude of the avalanche current pulse; dark count rate; delay and time-jitter of the electrical pulse with respect to the true arrival time of the photon.

The breakdown voltage depends on the structure of the device and on doping levels. Thin devices (typically 1 μm) have lower breakdown voltage (15 to 40V), and so lower power dissipation, while thick devices (typically from 20 to 30 μm) have higher breakdown voltage (300 to 400V) and power dissipation. The fabrication of thick devices is usually more complicated, but they can provide better quantum efficiency.

Breakdown voltage V_B also strongly depends on junction temperature. At constant supply voltage V_S, the increase of V_B causes a decrease of excess bias voltage V_{exc} if a fixed bias voltage is applied. At a high counting rate, the mean power dissipation causes a significant temperature increase, particularly in SPAD's with high V_B. The effects on device performance are significant, therefore it is very important to stabilize the junction temperature in working conditions.

24.2.3 Equivalent circuit of the diode above breakdown

In figure 24.10 the equivalent circuit of a SPAD is reported. When the diode is biased below the breakdown voltage V_B, the switch S is always open: in fact, an avalanche can't be triggered in these conditions. When the diode is biased above the breakdown voltage V_B, the switch S could be open or closed: when the switch is closed, the avalanche current I_a can flow, so closing S is like triggering the avalanche in the diode and this is equivalent to a carrier injection or generation in the high-field region. On the contrary, an open switch corresponds to no current flowing into the detector: this can happen before an avalanche is triggered or when the avalanche is quenched pulling the detector voltage down to V_B (or below).

C_d is the junction capacitance while R_a is the resistance of the diode. The R_a value is given by the ratio between the derivatives of V_d and I_a, and it depends on the semiconductor device structure. It is typically lower than 500Ω for devices with a wide area and thick depletion layer, and it ranges from a few hundred ohms to various kilo ohms for devices with a small area and a thin junction.
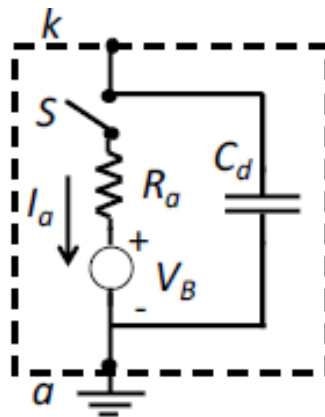

Figure 24.10: Equivalent circuit

24.3 Passive Quenching Circuit (PQC)

The avalanche mechanism is self-sustaining, meaning that an external circuit is required to stop the current flowing and restore the initial detector bias conditions that make it ready to detect another photon. Such a circuit is usually called *quenching circuit*. Its tasks include i) sensing of the leading edge of the avalanche current; ii) generation of a standard output pulse synchronous with the avalanche build-up; iii) quenching of the avalanche by lowering the bias down to or below the breakdown voltage; iv) restoring the photodiode to the operative level where it is ready to detect another photon. Considering the aforementioned SPAD equivalent circuit, a passive quenching configuration (PQC) is here discussed. It basically consists in a high-

resistive load R_L (in the order of 1MΩ) connected in series to the SPAD, as shown in figure 24.11.
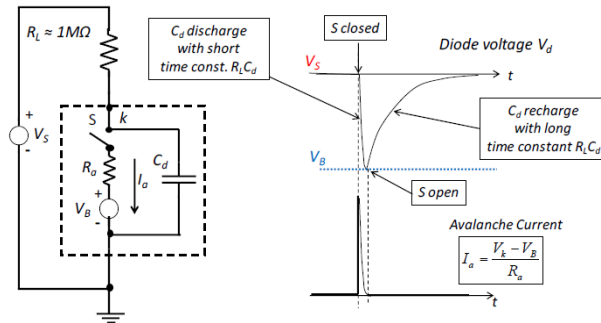

Figure 24.11: Passive quenching circuit (PCQ) and its operation

In this way, the avalanche current quenches itself by simply generating a voltage drop on the load R_L: if the voltage drop is high enough, the diode voltage goes down to V_B and the avalanche is no more self-sustaining since there is a high probability that no carrier flows in the depletion region for a while. To reach this situation the final current should be in the order of 50-100μA.

Let's analyze the circuit behavior looking at the graph of figure 24.11. As a starting condition, the switch S is opened, so no current flows in the circuit and the voltage drop across C_d is equal to V_s ($V_d=V_k-V_a=V_s$), where V_k is the cathode voltage and V_a is the anode voltage. When a photon is detected, I_a starts to flow causing a voltage drop on the high load impedance R_L: the diode voltage V_d falls down to the breakdown voltage and the avalanche is quenched. At this point the voltage on C_d is restored to the initial condition (V_s) with an exponential time constant given by R_LC_d. Since the switch S is opened when the avalanche is quenched, the small current carried by the load resistor R_L flows into the diode capacitance, slowly recharging it. A photon arriving during the first part of the recovery transient is almost certainly lost, since the avalanche triggering probability is very low. The subsequent arriving photons have a progressively higher probability of triggering an avalanche. For this reason, the gradual recovery from V_B to the diode bias voltage after each quenching can cause count loss in photon counting, reducing in this way the detection efficiency. The recovery transient does not always reach V_s. If a photon arrives in this time window, the avalanche current starts to flow and the voltage on C_d drops down again (figure 24.12). For this reason, at the beginning of the avalanche, the voltage on C_d is not always the same and the current-peaks amplitude can vary (figure 24.14).
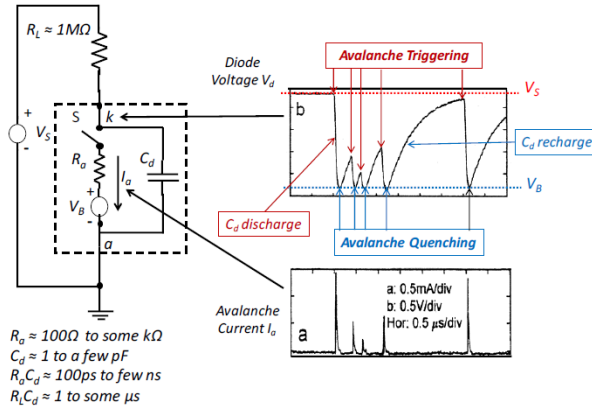
$R_a \approx 100\Omega$ to some kΩ
$C_d \approx 1$ to a few pF
$R_aC_d \approx 100ps$ to few ns
$R_LC_d \approx 1$ to some μs

Figure 24.12: Voltage and current behavior in PQC

The more is the delay between peaks, the more is the current amplitude because V_d is bigger. In order to read an output pulse from a PQC it is possible to insert a low value resistance R_s ($\approx 50\Omega$) between the diode and the ground of the circuit (figure 24.13). The obtained pulse is an attenuated replica of the diode voltage waveform and it is therefore called voltage-mode output.
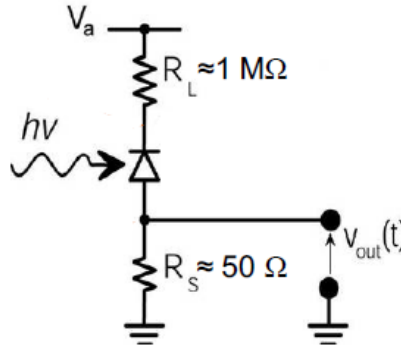

Figure 24.13: Circuit configuration to detect output pulse

In conclusion, SPADs used with a PQC can be useful for applications with very small probability of occurrence of an event during recovery transients, such as applications with low dark-count rate, low count-rate of background photons and low count-rate of the signal photons.

24.4 Active Quenching Circuit (AQC)

Passive quenching is simple, but as we have seen, it has some drawbacks. In particular it suffers from:

- Long and not well defined deadtime, that is the exponential charge of the diode capacitance in order to recover towards the potential V_a.

- As a consequence of the previous problem, we will have normally a low maximum counting rate, below 100kc/s.

- Finally, if a photon arrives during the voltage recovery, the arrival time measured can suffer of an increased delay and time-jitter with respect to the correct diode operation at V_a.

So, in order to avoid drawbacks causes by a slow recovery from avalanche pulses and exploit at the best the inherent performance of SPADs, a new approach was devised and implemented.

The basic idea was simply to sense the rise of the avalanche pulse and react back on the SPAD, forcing, with a controlled bias-voltage source, the quenching and reset transitions in short times.

Fig.24.14 illustrates the principle of the *active quenching* method. The rise of the avalanche pulse is sensed by a fast comparator whose output switches the bias voltage source V_a to breakdown voltage V_B or below. After an accurately controlled hold-off time, the bias voltage is switched back to operating level V_a. A standard pulse synchronous to the avalanche rise is derived from the comparator output to be employed for photon counting and timing.
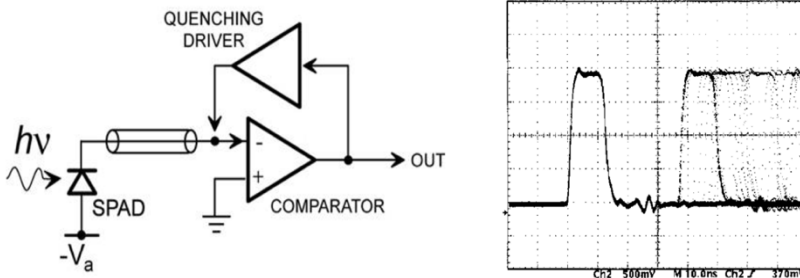
Figure 24.14: Active Quenching Circuit (AQC) and its output

The basic advantages offered by the AQC approach are:

- Fast transitions from quenched state to operating level and vice versa.

- Short and well-defined durations of the avalanche current and of the dead time.

- Standard output pulses.

- High counting rate, above 1Mc/s, since we have no more long recovery time to V_a

The introduction of the Active Quenching Circuit (AQC) opened the way to practical application of SPADs.

24.5 Comparison SPAD/APD and SPAD/PMT

Let's analyze the main differences between SPADs and APDs. The main features of the Avalanche Photodiode can be resumed as follows:

- It's biased close to but below the breakdown voltage
- It works in a linear-mode as an amplifier with a limited gain well below 1000.
- Its output is an analog one

On the contrary, the Single-Photon Avalanche Diode has the following characteristics:

- It's biased in the so-called Geiger-mode, well above the breakdown voltage (from 10 to 50V in thin-junction SPADs; from 200 to 500V in thick-junction SPADs)
- It's a bistable device since it can work with no current when no photon is impinging on the sensor, or with macroscopic current when the avalanche is activated by the generation of carriers.
- Its output is digital, it can assume just an high or low state.
- In Geiger-mode, so with a bias voltage well above the breakdown, the gain is meaningless.
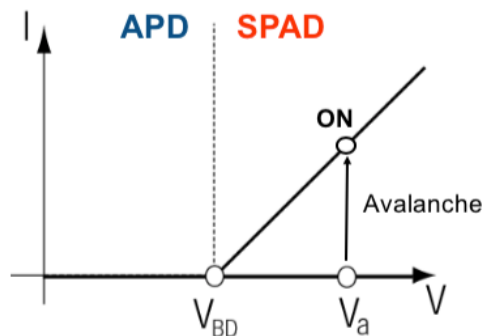
Figure 24.15: Different working region of APD and SPAD

SPAD shows some important advantage also compared to PTM in terms of:

- Miniaturization of the device, so smaller dimensions, lower voltages, lower power consumption, ruggedness, bulkiness, etc.

- Higher photon detection efficiency, since they do not rely on electron emission in vacuum from a photocathode as do PMTs, but on the internal photoelectric effect. It can be seen in figure 24.16, where the efficiency in the overall visible spectrum (390 nm – 700 nm) of SPAD is clearly higher, particularly in the red and near-infrared range
- Better photon timing: regarding the PMT, different factors are involved in the generation of transit time jitter, such as the distance from the electrode and the dynode and the initial random velocity of the electrons, so geometry and the potentials of the electrodes should be carefully designed. Classical values of FWHM for PMT ranges from a few hundreds of ps to a few ns. The FWHM for a fast planar SPAD is very low, in the order of some ps (figure 24.17).
- Comparable or lower noise, only if the area is not taken into consideration, otherwise PMT has a better dark count rate.
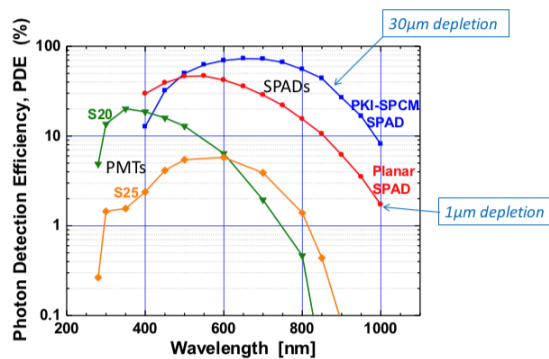

Figure 24.16: Photon detection efficiency vs wavelength in different technologies for PMT and SPAD
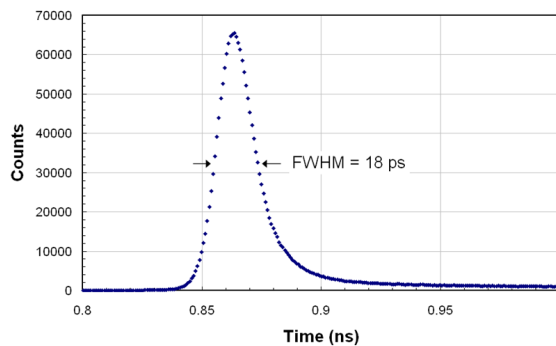

Figure 24.17 Timing Jitter of Fast Planar SPAD

In conclusion, if the area and the compactness is not important for a specific application, the optimal result is obtained by the PMT. This is due to the fact that its sensitive area can be enlarged until consistent values are reached, without increasing too much the noise. However the new trend is to develop integrable and compact devices so the SPAD has now become the best choice in many applications.

24.6 SPAD noise: dark count

SPAD is affected by an internal noise that must be considered for a correct analysis: the dark count. Dark count is the average rate of registered counts in a dark environment, without any incident light. This determines the minimum count rate at which the signal is dominantly caused by real photons.

Dark counts are due to carries thermally generated within the SPAD junction, so the dark count rate increases with temperature (figure 24.18, left). Thermal generation effects produce current pulses even in the absence of illumination, and the Poissonian fluctuation of these dark counts represents the internal noise source of the detector. The SPAD dark count rate increases also with excess bias voltage (figure 24.18, right) V_{exc} because of two effects, namely, field-assisted enhancement of the emission rate from generation centers and an increase of the avalanche triggering probability.
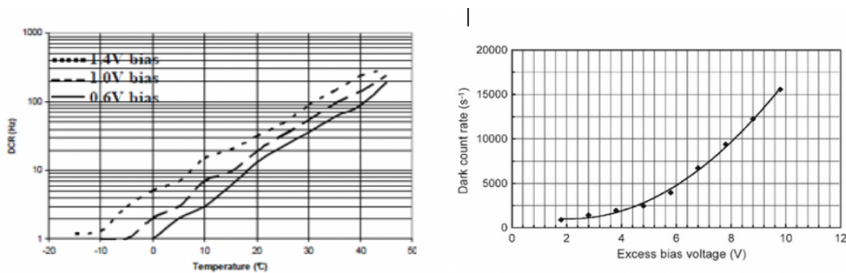

Figure 24.18: Dark count vs temperature (left) and Dark Count vs Excess bias voltage (right)

Both the quality of the material and the technological process used in the device fabrication have strong impact on the density of deep energy levels and therefore on the generation rate. Metal contamination may occur during silicon handling, high temperature heat treatments or ion implantations. Unintentional contaminants, Fe, Cu, Ti, Ni are usually found in silicon in concentration of about 10^{11} to 10^{12} cm^{-3}.

Poole-Frenkel and trap-assisted tunneling effects that occur at high electric fields (>10^5 V/cm), can greatly enhance the emission rate of deep energy levels (field-enhancement generation). Tunneling probability of carriers is, at first order, exponentially dependent on the energy barrier that they have to pass through.

Tunnel-assisted generation is not reduced by lowering the temperature and therefore sets a limit to the reduction of the dark count rate obtained by cooling the detector.
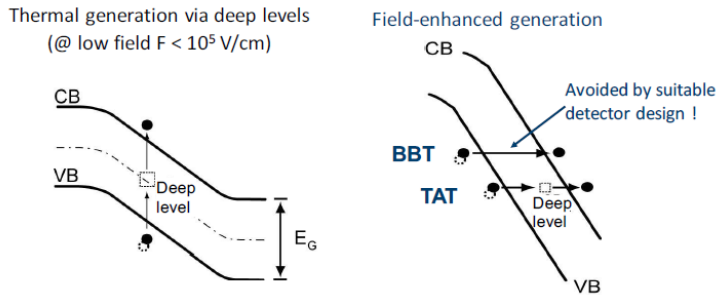
Figure 24.19: Example of thermal and tunneling carrier generation

Detector noise is also enhanced by a secondary effect called afterpulsing. During the avalanche some carriers are captured by deep levels in the junction depletion layer and subsequently released with a statistically fluctuating delay, whose mean value depends on the deep levels actually involved. Released carriers can retrigger the avalanche, generating afterpulses correlated with a previous avalanche pulse. The number of carriers captured during an avalanche pulse increases with the total number of carriers crossing the junction, that is, with the total charge of the avalanche pulse. Therefore, afterpulsing increases with the delay of avalanche quenching and with the current intensity, which is proportional to excess bias voltage V_{exc}.
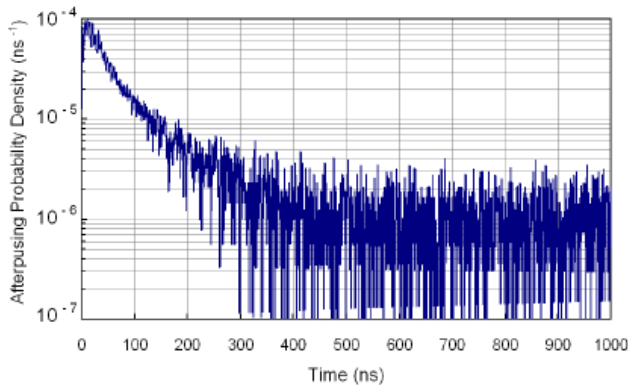

Figure 24.20: Probability density of afterpulse generation after an avalanche in a 100μm-diameter silicon SPAD operating at room temperature

24.7 Prototype SPAD structure

The first SPAD configuration was introduced at the Shockley laboratory in the early 1960's. The approach was to fabricate many n^+p junctions with very small diameter (a few microns) sorrounded by a deeply diffused guard ring and then select the few devices that didn't contain extended defects. Diffusing a shallow n^+ layer in a p bulk substrate two key features are reached: a low voltage operation ($\approx$30V), with the

benefit of a lower power consumption during the avalanche, and the fabrication of devices on a silicon wafer with a planar technology. The planar process is a manufacturing process used in the semiconductor industry to build individual components of a transistor, and in turn, connect those transistors together. In this way it's possible to integrate SPADs with other detectors and circuits.
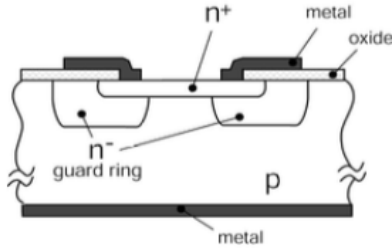
Figure 24.21: Planar technology implementation of SPAD (no epitaxy)

This planar configuration shown some weaknesses over time. In fact the deep guard-ring diffusion compromized the detection efficiency. It wasn't uniform in the active zone, and this introduced a strong limit in the diameter of the active n^+p junction, limiting also the scalability of the structure.

To overcome the drawbacks of early planar structures, the planar epitaxial devices were introduced in 1988 (figure 24.22). This structure fabrication starts from an n-type substrate on top of which a p^+ / p^- double-epitaxial layer is grown. The resulting p-n junction limits the neutral region from which carriers are collected, this effect reduces also the diffusion tail. The former guard ring is now substituted by a virtual one; this concept is based on the locally enhanced, in the central region of the shalow n^+p junction, of the electric field throught an higher p-doping. In this way the active region of the device is defined. Finally, a n-type diffusion region completely surrounds the detector. As result of the process the SPAD is enclosed in a p-well delimited by the isolation (n-type) and by the substrate allowing it to be electrically insulated from other devices.
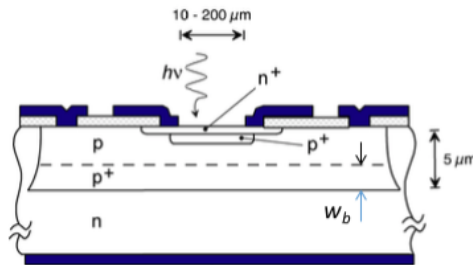
Figure 24.22: p-p+-n Double-Epitaxial SPAD structure

The continuous improvement of the planar epitaxial technology makes it possible an increasing of the active area diameters (up to 500μm) and the achievement of an optimum compromise between breakdown voltage (still ≈30V), total afterpulsing probability (≈1% at -15°C) and time jitter (better than 40ps FWHM, full width half maximum).
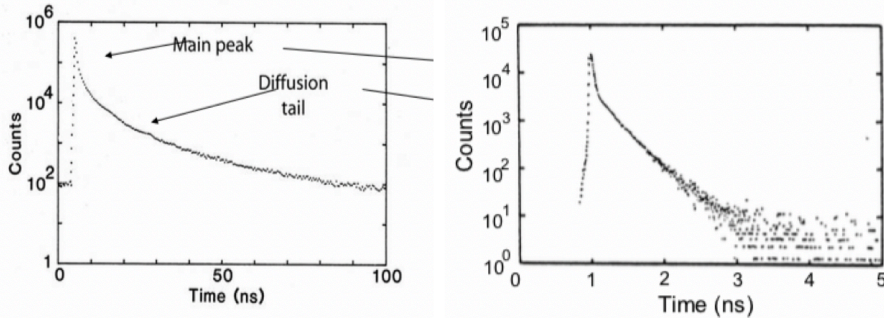

Figure 24.23: Tail of a planar technology devices (left), tail of a Double-Epitaxial SPAD structure (right)

Figure 24.23 highlights how the diffusion tail is strongly reduced. In fact, on the left side the old technology's response is analyzed and the diffusion tail lasts ≈100ns. Instead on the right side a great improvement is reached with the epitaxial technology, in which the diffusion tail lasts ≈3-5ns.

The lifetime of the tail depends on the square value of the neutral region thickness:

$$\tau = \frac{w_b^2}{\pi^2 D_n}$$

However, reducing the neutral region brings to a partially reduction of the detection efficiency, because the carriers in this region can diffuse until the depletion region. The device has to be customized taking this problem into consideration in order to reach the desired goal.

24.8 SPAD array

Two approaches are followed nowadays in SPAD detector technology: standard CMOS technology or custom one. The first one refers to the use of standard design for digital electronic, based on ultra-scaled complementary and symmetrical pairs of p-type and n-type MOSFET. The second refers to the use of a fully custom technology where any lithographic step can be modified to achieve better performance.

Without any technology customization, the goal is to obtain an array of SPADs having a low dark-count rate, high photon-detection probability, low afterpulsing probability, and acceptable timing jitter and breakdown voltage. Standard CMOS approach uses small pixel diameter (<50μm), introduces large numbers of pixel (>100 pixel) and takes advantage of smart-pixel architecture that has in-pixel electronic circuitry (figure 24.24).
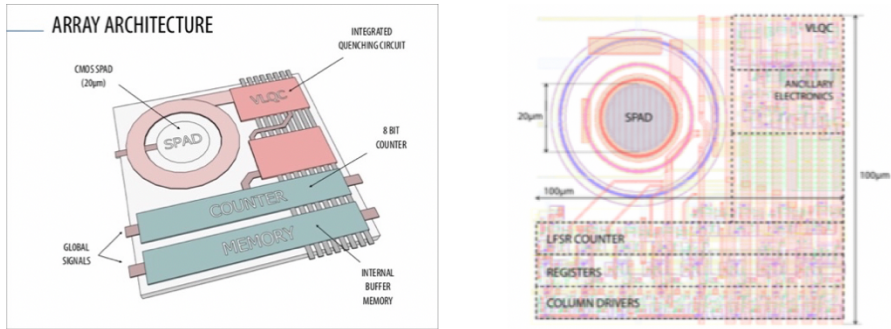
Figure 24.24: Example of pixel with SPAD and sorrounding electronics

In this configuration a fully parallel operation is possible in terms of pixel Single Photon Imager (High frame rate single photon imaging), or can also act as a "Single pixel" large area detector, that is a cooperation between all pixels in order to obtain a better performance (low dead time, high count rate and photon number resolution). A critical factor is the cost-down without sacrificing the system performance. Furthermore, at device level, CMOS SPAD's performance could not be optimized with adjustable layer arrangement and doping profile, which is in general not free, even not possible, for the existing CMOS technology.

For the applications where high photon-detection efficiency is requested, SPAD arrays are fabricated using a custom process, instead. In fact, this approach allows the design of a high-quality pixel array with wide pixel diameter (>100µm). In this case, the integration with CMOS circuitry is more difficult, since only a limited number of steps are implemented, so normally a moderate number of pixel is introduced (<100 pixels).

24.9 Silicon Photomultiplier

The Silicon Photomultiplier is a detector based on array of SPADs, where each pixel has its own load resistance ($R_L \approx 100k\Omega$) and capacitance ($C_L \approx 100$ fF). All pixels have a common ground terminal, connected to a low resistance external load, ($R_s = 50\Omega$) (figure 24.25, left). All the pixels currents flow in this terminal and they are added among them. In this way the output amplitude is proportional to the excited photodetectors and to the number of incident photons (figure 24.25, right). The detector pixels are individually triggered by incident photons, quenched by the discharge of pixel capacitance and reset by the recharge of C_L with time constant $R_L C_L$.

The Silicon Photomultiplier provides the low-light detection capabilities of the PMT while offering all the benefits of a microelectronic device. The SiPM features low-

voltage operation, insensitivity to magnetic fields, mechanical robustness and excellent uniformity of response. The SiPM is capable of discriminating the precise number of photons detected as distinct, discrete levels at the output node as a function of the number of triggered cells.

However, SiPMs have also drawbacks with respect to PMTs. For example, the fill factor (percentage of the sensor surface area that is sensitive to light, active area/ total area) it is not 100% and this brings to a reduction of the photon detection efficiency because the active area. Furthermore, SiPMs have much higher dark current density over the active area.
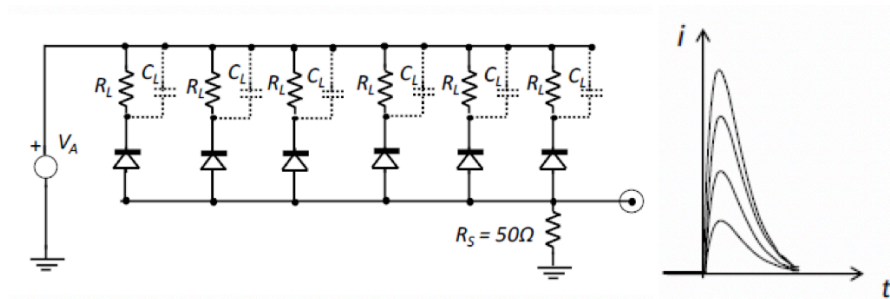
Figure 24.25: Equivalent circuit configuration Silicon PhotonMultiplier (left), currents at different number of photons detected and excited photodetectors (right)

CHAPTER 24

Temperature Sensors

Temperature is one of the most important and commonly measured physical quantities. Consequently, temperature sensors cover the largest segment of the sensor market by volume. Moreover, many of the physical phenomena being sensed and measured (e.g., humidity, pressure, flow, stress and gas concentration) have some temperature dependence and therefore, need to be compensated for temperature variations. Major applications of temperature sensors are thus focused at sensing temperature for thermal compensation. In this chapter, four types of temperature sensors are examined: Thermocouple (TC), Resistive Temperature Detector (RTD) and Thermistors.

25.1 Thermocouples (TC)

Thermocouples are probably the most diffused temperature sensors. They are cheaper than other temperature-measuring devices and offer the widest temperature range, even though they are characterized by low accuracy and require calibration. The thermocouple consists of two wires of different metals welded together at the ends. In the following discussion, two separated metals at equal temperature T_R are considered. The two metals A and B are characterized by their work functions Φ_A and Φ_B (the minimum thermodynamic work needed to remove an electron from a solid to a point in the vacuum outside the solid surface).
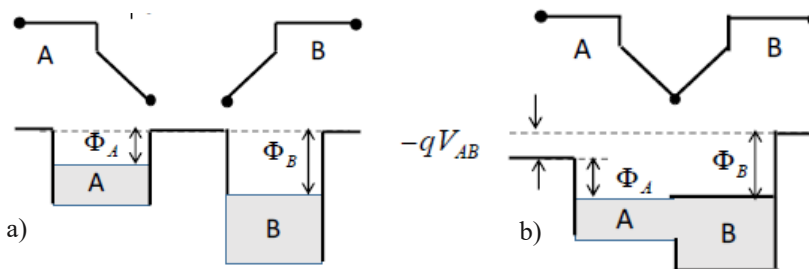
Figure 25.1: Two metals at the same temperature T_R, a) separated; b) in contact

When the two metals are put in contact, electrons in higher energy states in metal A move to free states at lower energy in metal B. Thus, B gets negative charge and A positive charge; a voltage drop is established between A and B and the difference

of Fermi levels is reduced. The equilibrium is reached when the Fermi levels in A and B are equal. The potential difference V_{AB} at the A-B contact is given by

$$V_{AB} = \frac{\Phi_A - \Phi_B}{-q}$$

Considering a chain of metal wires in contact (for example A – B – A), all at the same temperature T_R, equilibrium is reached when all metals have equal Fermi level, as before. It is worth noting that the potential difference between the opposite ends of a metal sequence does not depend on the intermediate metals, but only on the end metals. In fact, if the boundary metals are equal, the end-to-end potential difference is nil. The situation is different if one of the junctions is characterized by a different temperature T_S. This effect is exploited in the working principle of the thermocouple.

25.1.1 Seebeck Effect

Thermoelectric effects consist in the conversion of a temperature gradient, that a conductor experiences, to a voltage and vice versa. This conversion can be performed by a thermocouple. Thermoelectric effects include different phenomena, as the Seebeck, the Peltier and the Thomson effect.

The Seebeck effect is the conversion of a temperature difference directly into electricity at the junction of different types of wires, i.e. conductors; on the other hand, the Peltier Effect consists in the generation of a temperature difference by making a current flowing between two junctions; finally, the Thomson effect consists in the absorption or generation of heat when electric current passes through a circuit composed of a single material that has a temperature difference along its length.

In a thermocouple, the flow of electric current in the wires is negligible, hence it exploits only the Seebeck effect, which is present also with zero current flowing, i.e. in open circuit. As an example of the Seebeck effect, let's consider a wire with higher temperature T_S at one end and lower T_R at the other end. At higher temperature, the random thermal motion of free electrons is enhanced, and electrons diffuse away from the hot region. Since the lattice ion space charge is permanent, a net space charge builds up (positive in the hot region, negative in the cold region), which generates a potential variation along the wires. A potential difference between the metal ends is established, which depends on the temperature difference $\Delta T = T_S - T_R$ and on the properties concerning the free carrier motion, which are different in different metals.

The magnitude of the voltage produced between the two junctions depends on the material and on the temperature difference ΔT through the linear relationship defining the *Seebeck coefficient S*, specific for the material.

For each material, S is defined as the potential difference along the metal wire per unit temperature difference

$$S = \frac{dV}{dT}$$

The value (and the sign) of S depends on the metal properties concerning the free carrier dynamics; it is in a range from about 1 µV/K to a few 10 µV/K.

The voltage difference between two points at different temperatures T_R and T_S in a metal is given by:

$$\Delta V = V_S - V_R = \int_{T_R}^{T_S} S dT$$

For the metals currently employed in thermocouples, the Seebeck coefficient S is different for different metals and it has just small variations with the temperature.

Therefore, the Seebeck voltage ΔV_{SR} is approximately proportional to $\Delta T = T_S - T_R$, with different coefficient S for different metals.

25.1.2 Thermocouple Principle

In the chain of metals previously considered isothermal at temperature T_R, let us now bring one of the A-B junctions at a different temperature T_S, leaving the rest at T_R; as a result, a potential difference $\Delta V_{21} = V_2 - V_1 = V_T$ is established at the end terminals.
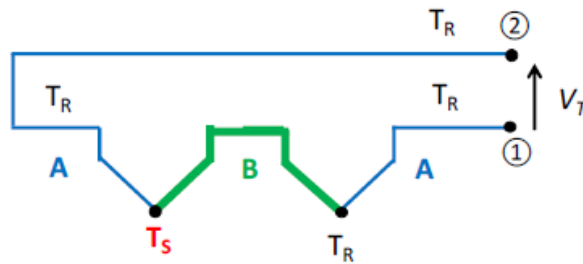

Figure 25.2: Potential difference at the junction ends

V_T is given by the sum of the potential differences due to the Seebeck effect in the A and B wires contiguous to the junction at temperature T_S.

$$V_T = V_2 - V_1 = (V_S - V_R)_B + (V_R - V_S)_A = \int_{T_R}^{T_S} S_B dT + \int_{T_S}^{T_R} S_A dT$$

$$V_T = \int_{T_R}^{T_S} (S_B - S_A) dT$$

If S_A and S_B are approximately constant in the temperature range of T_R and T_S, the dependence of voltage from temperature can be considered linear and the integral can be simplified in the following way:

$$V_T = V_2 - V_1 \approx (S_B - S_A)(T_S - T_R) = S_{BA}(T_S - T_R)$$

where S_{BA} is the difference between the two Seebeck coefficients.

Under the hypothesis of operating in a linear range and if T_R is a constant and known value, the potential difference measured between the junctions can be converted in a temperature difference ΔT and the T_S value is obtained:

$$\Delta T = T_S - T_R = V_T/S_{BA}$$

$$T_S = \Delta T + T_R = V_T/S_{BA} + T_R$$

Therefore, the junction behaves as a temperature sensing device. Thermocouples are employed over a wide temperature range, from 0°C to 1000°C, and down to negative temperatures, but the behavior of voltage versus temperature can become significantly nonlinear. For this reason, TC manufacturers provide calibration plots: they show the behavior of V_T versus temperature, measured with reference to T_{R0}=0°C (at the water triple point, liquid-ice-vapour in thermal equilibrium in a dewar vessel). About negative temperature, the nonlinearity in the TC behavior is even more marked, since the Seebeck coefficient decreases with temperature. Thus, thermocouples are not efficient for cryogenic measurements.

In figure 25.3 a calibration plot is reported, with curves related to the most diffused typologies of TCs.
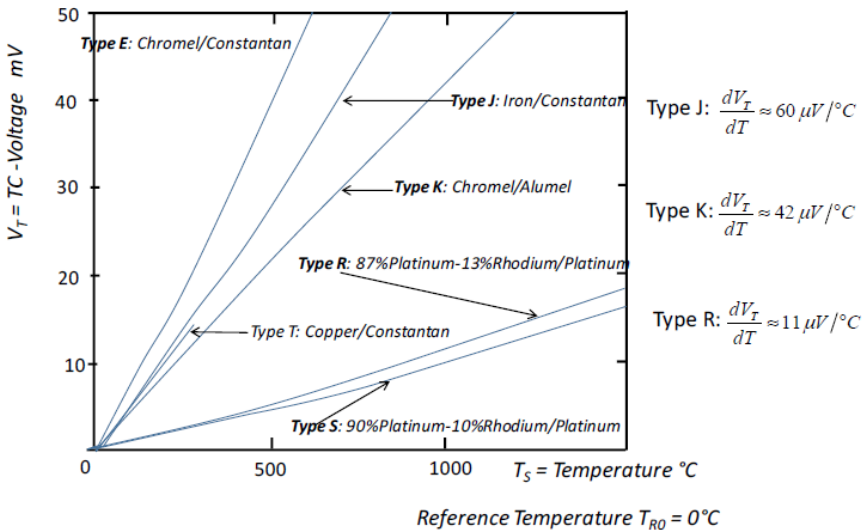

Figure 25.3: Calibration plot

25.1.3 TC operation in nonlinear range

Using calibration tables, a thermocouple can be well exploited also in a range of nonlinear voltage-temperature behavior. The temperature difference $\Delta T = T_S - T_{R0}$ is

directly obtained from the measurement of the potential difference V_T by employing the nonlinear calibration plot.
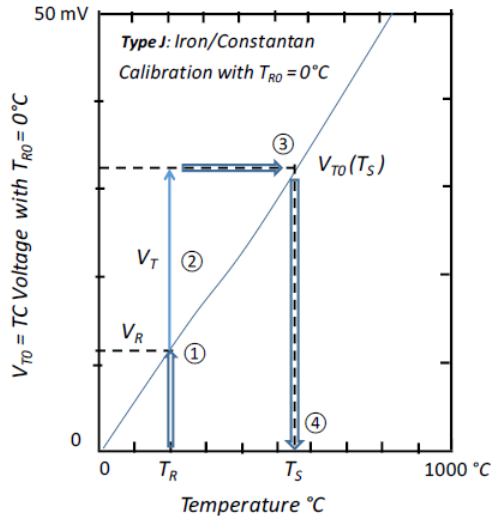

Figure 25.4: Calibration plot for a type J thermocouple

If the reference temperature T_R is different from the T_{R0} in the calibration, the measured voltage V_T must be corrected before using the calibration table.

1) In the TC Calibration, $V_{T0}(T_S)$ is measured with reference to $T_{R0}=0°C$, while in the TC Operation $V_T(T_S)$ is measured with respect to a different temperature T_R. Therefore, to obtain T_S, we should read from the plot the value $V_R=V_{T0}(T_R)$.
2) The value V_T measured in operation is identified on the plot.
3) $V_R + V_T = V_{T0}(T_S)$
4) The correct value of T_S can now be extracted.

It can be noticed that accuracy and precision of the temperature measurement are limited by the knowledge of the reference temperature T_R. Operating with the reference junction at standard temperature T_{R0} is a technically sound approach, but the set-up which must be used to establish and maintain accurately T_{R0} is quite complex and is often unsuitable to practical applications. An option is to use a thermostat at $T_R > T_{R0} = 0°C$; the value of T_R is measured with a precise auxiliary thermosensor and controlled in a feedback loop with a heater. Since the control loop must operate in a narrow range of temperature, a Resistive Temperature Detector (RTD, see paragraph 25.2) is suitable.

25.1.4 Reference Junction and Input Compensation

The terminals of the thermocouple have to be connected to an electronic circuit to read the differential voltage V_T: a differential preamplifier is required. This connection introduces an additional junction of different materials, which brings an unwanted contribution to the voltage, impairing the measurement.
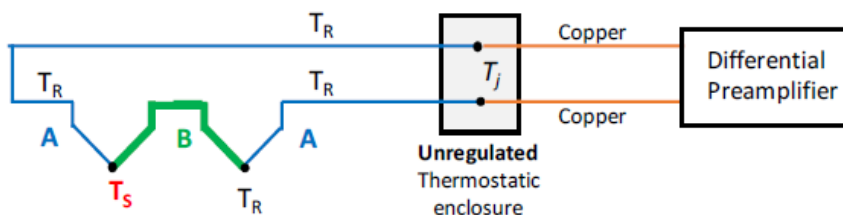

Figure 25.5: Compensation with unregulated thermostatic enclosure

This contribution can be cancelled if the two junctions A-metal/Copper are held at the same temperature T_j, even if the value of T_j is unknown. The Seebeck voltages of the two couples A-metal/Copper have the same value but opposite sense in the circuit, so they cancel out. This solution, achieved by putting the input contacts in a thermostatic enclosure, provides good results even without regulation.
Another possibility is the Reference Input Configuration.
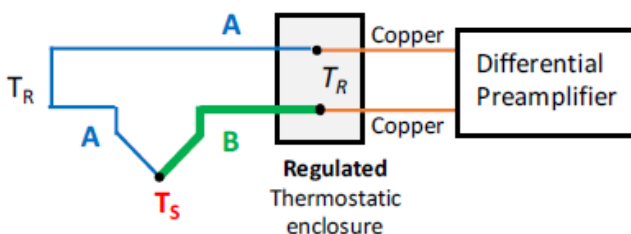
Figure 25.6: Compensation with reference input configuration

A regulated thermostatic enclosure is used to keep the two junctions at the same known temperature. The enclosure can be set to a reference T_R or it can be left free to follow the ambient temperature, but it is continuously measured, so that each measurement takes into account the actual value of T_R. Using this configuration, the second junction A/B acting as a reference can be avoided, because the input junction itself acts as a reference. It can be easily verified that when the thermometric junction A/B is at temperature T_R, the end potential difference is zero and the chain is isothermal.

25.1.5 Operation with TC Remote from Electronics

The wires that connect a thermocouple to the following electronic circuitry can be several meter long if the sensor has to be placed far from the electronics: some examples are the distributed temperature controls in industrial workshops or the measure of high temperature in furnaces. As we have seen before, a regulation of the reference temperature T_R is needed, and this control must be implemented far enough from the object of the measurement. Thus, the long connections cannot be made in copper, but the same metals adopted in the TC have to be used.
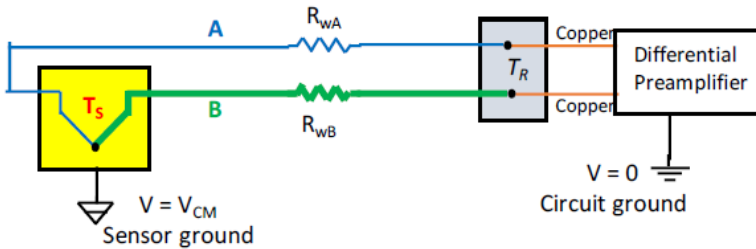

Figure 25.7: TC remote from electronics

Furthermore, the two metals of a TC have different and fairly high resistivity, while the long connection wires have small section (to reduce heat transmission as well as cost and space). As a result, the connection wires are typically characterized by fairly high resistances, in the order of hundreds Ohm, and remarkably different values. This difference in the resistances causes problems in thermocouples, as the ground potential differences due to the operation remote from the associated circuit.

25.1.6 Electronics associated to the sensor

As said before, a differential preamplifier is needed to read the voltage at the ends of a thermocouple. In this paragraph, the value of parameters required for the circuitry are examined. The preamplifier must have CMRR sufficiently high to keep within acceptable limit $\Delta V_{C,max}$ the error caused by V_{CM} in the measurement of V_T. In fact, the ground potential difference V_{CM} from circuit to remote sensor is often remarkable, in the order of tens of Volts.
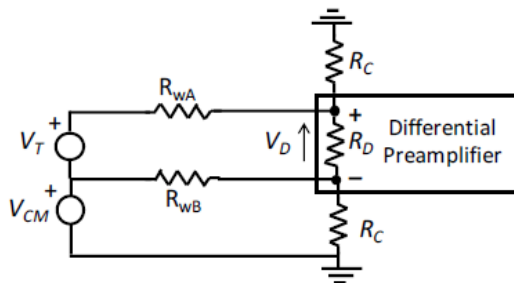

Figure 25.8: Differential configuration

$$CMRR \geq V_{CM}/\Delta V_{C,max}$$

The $\Delta V_{C,max}$ value is set by the limit ΔT_{max}, the maximum error in temperature that can be tolerated in the measurement:

$$\Delta V_{C,max} = S \cdot \Delta T_{max}$$

S is the Seebeck coefficient of the employed TC. In most cases the requirements for the Common Mode Rejection Ratio are fairly stringent, around 100-120 dB.
Another requirement regards the differential resistance at the amplifier input terminals: the non-negligible resistance of the connection wires causes a signal loss, that must be maintained lower than a minimum acceptable value.

$$\Delta V_T = V_T - V_D = V_T \frac{R_{WA} + R_{WB}}{R_D//2R_C + R_{WA}+R_{WB}} \approx V_T \frac{R_{WA}+R_{WB}}{R_D}$$

$$\frac{\Delta V_T}{V_T} = \frac{R_{WA} + R_{WB}}{R_D} < \frac{\Delta T_{max}}{T - T_R}$$

$$R_D > (R_{WA} + R_{WB})\frac{T-T_R}{\Delta T_{max}}$$

Usually the value of R_D required is about a few MΩ; if $R_D > some\ M\Omega$, it can be considered ideal, therefore approximated to an infinite resistance.
Since the resistances of the two wires are different, they cause different voltage drops on the wires and generate a fake differential signal. This signal must be lower than the maximum allowed error $\Delta V_{C,max}$. This consideration translates into a minimum acceptable value for the common-mode input resistance R_C.

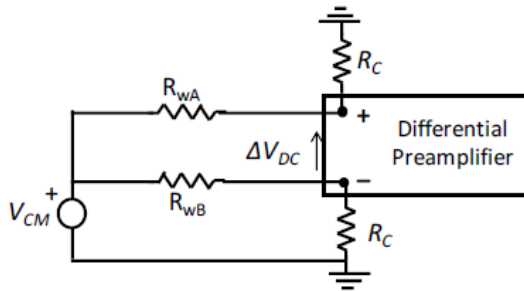**Figure 25.9:** Equivalent circuit for computation of common mode input signal

$$I_{WA} = \frac{V_{CM}}{R_C + R_{WA}} \approx \frac{V_{CM}}{R_C}$$

$$I_{WB} = \frac{V_{CM}}{R_C + R_{WB}} \approx \frac{V_{CM}}{R_C}$$

$$I_{WA} \approx I_{WB} \approx I_W = \frac{V_{CM}}{R_C}$$

$$\Delta V_{DC} = I_{WA}R_{WA} - I_{WB}R_{WB} = V_{CM}\frac{R_{WA} - R_{WB}}{R_C}$$

$$R_C > (R_{WA} - R_{WB})\frac{V_{CM}}{\Delta V_{DC,max}}$$

In many cases a very high value of R_C is required, in the order of hundreds of MΩ. An Instrumentation Amplifier may be the appropriate solution.

25.1.7 Noise

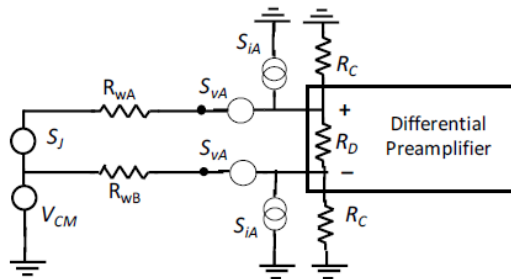In figure 25.10 all the noise equivalent generators are put in evidence.

Figure 25.10: Equivalent circuit for noise computation

The sensor is a metal resistor, so it does not add a relevant contribution to flicker noise. The sensor resistance R_{wA} + R_{wB}, usually a few hundreds of Ohms, are characterized by an equivalent noise voltage generator in the range

$$\sqrt{S_j} \approx 1 \ \ to \ \ 10nV/\sqrt{Hz}$$

This contribution has to be considered as well as the amplifier noise. Anyway, in many cases the dominant contribution is given by the amplifier voltage noise, while S_{iA} is usually negligible. The amplifier is characterized by 1/f noise components as well. This contribution has to be carefully treated, because temperature is in general a very slow signal, with a low-frequency bandwidth.

To reduce the effect of flicker noise on the measurement of TC signals, modulation is a good choice. Modulator circuits cannot be used since TC signals may not be higher than the modulator flicker noise. Switches are instead suitable.

Switches with the lowest noise are the electromechanical metal-contact ones; a drawback is represented by the limited range of operating frequencies (the switching frequency can be chosen up to a few hundreds of Hz).
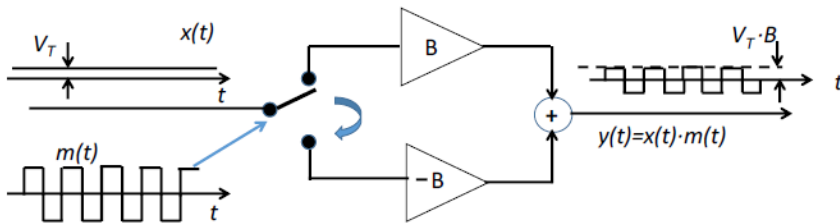
Figure 25.11: Signal modulation with switches

Alternatively, electronic switches like MOSFETs and diodes can operate up to high frequencies (even MHz). Unfortunately, they are characterized by higher noise than metal switches; anyway, MOSFETs in switching operation have lower noise than in linear operation (amplifiers) and can be used for TC signal modulation.

After the modulation, a high-performance filtering solution is represented by a Lock-In Amplifier.

25.1.8 TC Dynamic Response

The response of a thermocouple to a step in the temperature to be measured is determined by the thermal transient of the TC junction. The temperature-to-voltage transfer function of the TC is a low-pass filter.

A fair approximation is to consider the TC in linear operation as a single-pole low-pass filter. Similarly, to what happens for thermal photodetectors, the thermal transient is characterized by two parameters: the thermal resistance R_T and heat capacitance C_T.

R_T is the resistance between the object measured and the TC junction; C_T is the capacitance of the TC junction. Their values strongly depend on the fabrication technologies of TCs and vary from case to case.

The time constant $\tau = R_T C_T$ of the single pole response of the TC has anyway a value remarkably longer than usual in electronic filters and can widely vary in the different practical cases, from millisecond range to seconds.

25.1.9 Applications

Table 25.1 reports the most diffused topologies of thermocouples. Each type is characterized by a value of sensitivity, a temperature range and a color code. Actually, different internationally recognized standards have been developed. In table 25.1 the United States ASTM and ANSI standard is reported.

Thermocouples are widely used in science and industry. Applications include temperature measurement for kilns, gas turbine exhaust, diesel engines, and other industrial processes. Thermocouples are also used in houses, offices and businesses as the temperature sensors in thermostats, and also as flame sensors in safety devices for gas-powered appliances.

It can be noticed from table 25.1 that thermocouples of type R, S and B all use alloys of platinum and rhodium/platinum. These are among the most stable thermocouples, but they have lower sensitivity than other types, approximately 10 µV/°C. They are usually employed only for high-temperature measurements due to their high cost and low sensitivity.

Type K is the most common general-purpose thermocouple with a sensitivity of approximately 42 µV/°C. It is inexpensive, and a wide variety of probes are available.

Type J is instead characterized by a more restricted range of temperature than type K but it presents a higher resistivity.

Thermocouples of type T are often used as a differential measurement, since only copper wire touches the probe.

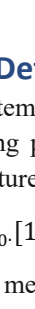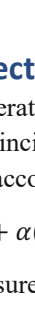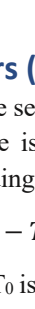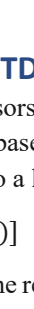| Type	Metals	Sensitivity (dV$_T$/dT)	Temperature Range	Colour Code
J	Iron/ Constantan	60 µV/°C	0 to 750°C	J (black) + / -
K	Nickel Chromium / Nickel Aluminium	42 µV/°C	-200 to 1250°C	K (yellow) + / -
R	87%Platinum- 13%Rhodium/Platinum	11 µV/°C	0 to 1450°C	R/S (green) + / -
S	90%Platinum- 10%Rhodium/Platinum	10 µV/°C	0 to 1450°C	R/S (green) + / -
T	Copper/Constantan	48,2 µV/°C	0 to 350°C	T (blue) + / -
B	70%Platinum- 30%Rhodium/Platinum	10 µV/°C	800 to 1700°C	B (grey) + / -

Table 25.1: Most common Termocouples

25.2 Resistive Temperature Detectors (RTD)

The RTD is one of the most common temperature sensors used in many applications as well as thermocouples. Its working principle is based on the variation of the resistivity of a metal with the temperature according to a linear equation:

$$\rho(T) = \rho_0 \cdot [1 + \alpha(T - T_0)]$$

 where T is the temperature we want to measure, T_0 is the reference temperature, $\rho(T)$ is the resistivity at T and ρ_0 at T_0.

The second Ohm's law says that the resistance of a conductor is directly proportional to its length L and inversely proportional to its section S:

$$R = \rho \frac{L}{S}$$

Combining together the two expressions, the relationship between resistance and temperature can be obtained, at a first order approximation:

$$R(T) = R_0 \cdot [1 + \alpha(T - T_0)]$$

where α is the temperature coefficient of resistance for the conductor material. With a good approximation, the resistance has a linear behavior on wide T range for various metals.

Calling $\Delta T = (T - T_0)$ and $\Delta R_S = (R_S - R_0)$ is possible to define $\Delta R_S = \alpha \Delta T R_0$, therefore the variation of resistance is connected to the variation of the temperature through the α coefficient, that is around $\approx 4 \cdot 10^{-3}$ for metals currently employed in RTDs. The value of α for some metals typically used in RTDs is reported in the following table.

Metal	α
Platinum Pt	$3,9 \cdot 10^{-3}$
Copper Cu	$4,3 \cdot 10^{-3}$
Tungsten W	$4,6 \cdot 10^{-3}$
Nickel Ni	$6,8 \cdot 10^{-3}$

RTD sensors usually are made by pure materials such as copper, nickel, tungsten or platinum. Usually this kind of sensors are used for small range of temperature, from less than 0°C to some hundred degrees; for higher temperatures, instead, thermocouples are employed. It is useful to represent the curve of the relative variation of resistance as function of temperature. Platinum is chemically inert and resistant to contaminations; in fact, the α coefficient, whose value can change with impurities, is quite constant. As a result, the characteristic of Platinum is linear with good approximation from -200°C to about 500°C and with small deviation from linearity up to 800°C.

All these properties make platinum the best material for RTD sensors. Common RTD sensing elements made of platinum are PT100 and PT1000; the value of their resistance is respectively 100Ω and 1000Ω at reference temperature (0°C). Although the platinum is a quite expensive material, the amount used in the sensor is really small, consequently the cost of the overall sensor is not high.

Pt is the material of choice in many cases and is used in official metrology to define the International Practical Temperature Scale (from 13,81 K to 903,89 K).
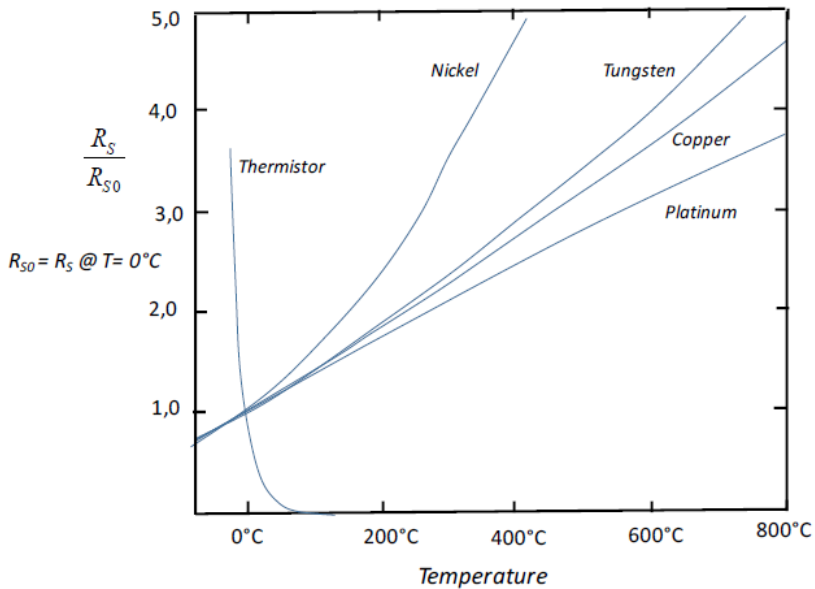
Figure 25.12: Plot of relative variation of resistance as function of temperature

The RTD fabrication technology is not so simple. Since the α coefficient is very small, the resistance employed should have a big value, like some tens or some hundreds of ohms, in order to be able to measure a significant variation of relative resistance. One of the big problems of RTDs is not the electronics or the sensor itself, but the connection of the sensor to the object to be measured without introducing errors during the reading phase. The package must be compact and ensure good thermal contact between the resistor and the object measured; good electrical isolation is also needed: insulators such as PVC or silicone rubber are used at temperatures above $T \approx 250°C$ and grass fiber or ceramic for higher temperatures.

Many RTD sensors consist of a long and thin wire wrapped in spiral around a ceramic or glass core. The mechanical structure must avoid strain of the metal wire due to thermal expansion or contraction, otherwise the piezoresistive effect would cause unwanted resistance variations and consequent errors in ΔT.

RTDs do not generate an electrical signal, so a power supply is necessary to get current and voltage in the sensor; the higher is the bias voltage, the higher is the signal, but some problems can arise: above all, Joule self-heating is the main limitation for these kinds of sensors. This effect leads to an increase of power consumption and could modify the temperature of the object under test. The only solution is to limit the maximum power $P_{S,max}$ below the threshold that the system

can tolerate, depending on the final application and on the thermal conductivity both of the sensor and of the object.

$$P_S = \frac{V_S^2}{R_S} \qquad\qquad\qquad P_S \le P_{S,max}$$

Hence, $V_S \le \sqrt{R_S \cdot P_{S,max}}$ is the value of the maximum voltage allowed in the RTD. Typical value of the sensor resistance is $R_S \approx 100\Omega$ and power $P_{S,max} \approx 100\mu W$, thus the maximum value for the bias voltage is around 100mV, that definitely leads to small voltage variations.

25.3 Circuit for measurement

In modern electronics, a simple and practical approach to measure the variation of temperature uses a current generator and a voltmeter:
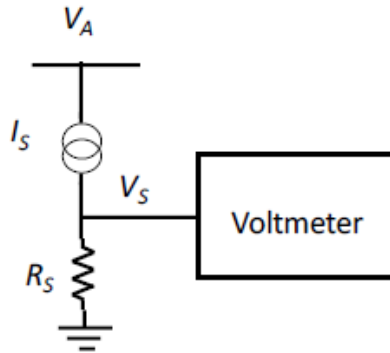
Figure 25.13: RTD operation at constant current

R_S is biased with a constant current generator I_S, then the voltage V_S across R_S is measured and at any temperature T, it is exactly proportional to the sensor resistance. The variation of voltage ΔV_S consists just in the difference between the measured value of V_S and the value of the voltage across R_S at the reference temperature (0°C), V_{S0}; ΔV_S gives an accurate measure of ΔR_S, that is a precisely known function of $\Delta T = T - T_0$

$$V_{S0} = R_0 I_S \qquad\qquad \Delta R_S = f(\Delta T) \approx \alpha R_0 \Delta T$$

$$\Delta V_S = V_S - V_{S0} = I_S \cdot \Delta R_S = V_{S0}\frac{\Delta R_S}{R_0} \approx V_{S0}\alpha\Delta T$$

Since α is a very small value and ΔV_S is directly proportional to the coefficient, also the overall variation of voltage is small and reading it over a bias value of some volts is not easy nor convenient. A possible solution is to use a differential measurement.
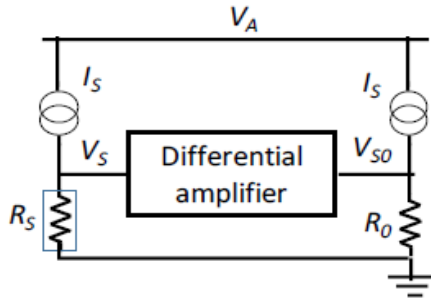
Figure 25.14: Differential configuration to avoid self-heating

The configuration shown in figure 25.14 includes the reference voltage V_{S0} on the right part of the circuit and makes it possible to take directly differential measurements of ΔV_S. On the left, we have the signal that has to be measured. However, in various cases the RTD is placed on a measured object not near to the circuit and a couple of wires brings the signal from the sensor to the amplifier.
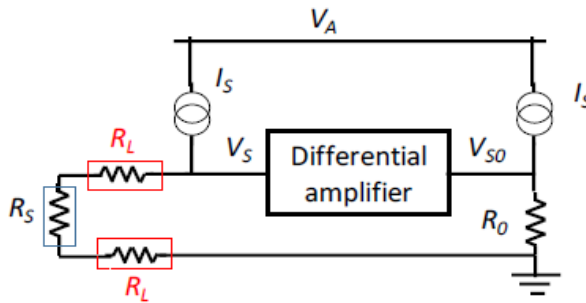
Figure 25.15: Two-wire-connection. Highlight on wire resistances

The wires are both made of the same material and have the same resistance R_L in which flows I_S from the current generator, causing an unwanted voltage drop. Hence, an error occurs due to the misreading of the correct value of ΔV_S. The circuit above represents the so called "Two-wire-connection" and since the resistances on the left side can all be considered in series, the voltage drop added to V_S is equal to $2 I_S R_L$. Errors in ΔV_S due to wire resistance R_L are avoided using a "Three-wire-connection".
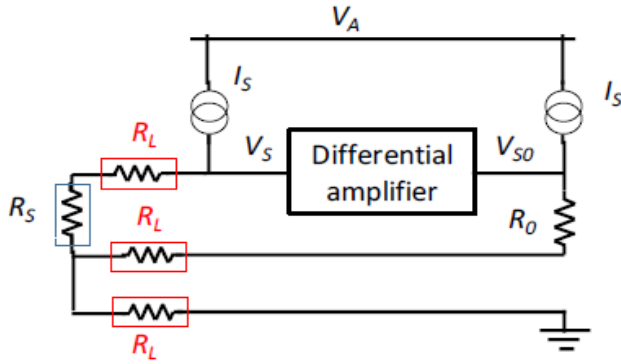
Figure 25.16: Three-wire-connection

The third wire resistance is inserted in the common return path to ground. With this configuration both left and right side of the circuit include in series a wire resistance R_L and, under the hypothesis of having R_S equal to R_0, the voltage drop is the same for both sides. An alternative configuration, useful when current generators are not available, is the Wheatstone bridge: it requires only four resistances and it is cheaper, easier to use and it occupies less space than a current generator, therefore the costs of the system can be reduced.
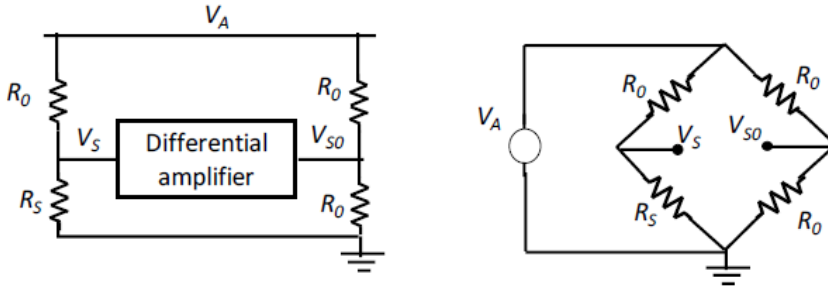
Figure 25.17: Wheatstone bridge

The only drawback is that it is not linear. It is based on the principle of a voltage divider, implemented by the R_S of the RTD in series with a reference resistor R_0; the variations of the divider output voltage correspond to the variation of the sensor resistance. In a first analysis, the four resistances can be considered equal; R_S is the sensor resistance and the other three R_0 have the same value.

$$V_{S0} = V_A \frac{R_0}{R_0+R_0} = \frac{V_A}{2} \qquad V_S = V_A \frac{R_S}{R_0+R_S}$$

Since the sensor resistance is connected to the object, a ΔR_S variation takes place due to a variation of temperature, so $R_S = R_0 + \Delta R_S$.f the variation of the relative resistance is less than 5% ($\Delta R_S < 0,05R_0$), the voltage variation ΔV_S can be considered linear and its value is the derivative of V$_S$ as function of the resistance R$_S$:

$$\Delta V_S = \Delta R_S \left(\frac{dV_S}{dR_S}\right)_{R_S=R_0} = \frac{V_A}{4}\frac{\Delta R_S}{R_0} = \frac{V_A}{4}\alpha\Delta T$$

The result is a function both of the bias voltage V$_A$ and of the α coefficient; nevertheless, the variation is very small it does not represent an issue because a modulation technique can be used to improve the voltage value, i.e. Lock-in Amplifier arrangement. A possible doubt could arise about the optimum value of the three resistance in the bridge.
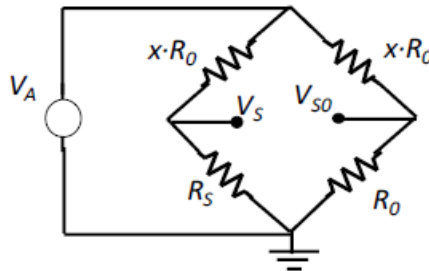
Figure 25.18: Wheatstone bridge with different values of R$_0$

For example, in this configuration the resistances on the upper part of the circuit have a value x times R$_0$; computing the same calculations as before the result is different, since all the voltages, including the variation ΔV_S, become function of the x value:

$$V_{S0} = V_A \frac{R_0}{R_0 + xR_0} = \frac{V_A}{1+x}$$

$$V_S = V_A \frac{R_S}{xR_0 + R_S}$$

$$\Delta V_S = \left(\frac{dV_S}{dR_S}\right)_{R_S=R_0} \Rightarrow \Delta R_S = V_A \frac{x}{(1+x)^2}\frac{\Delta R_S}{R_0}$$

The Wheatstone bridge can be employed with any ratio x of the voltage divider. However, with x=1 the highest output ΔV_S is obtained:

$$\max\left[\frac{x}{(1+x)^2}\right] = \frac{1}{4} \quad for \quad x=1$$

If the x value changes also the sensor sensitivity will change: an increase of the x value will lead to worsen ΔV_S, so a worse SNR.

It is possible to extend the application of the Wheatstone bridge also to cases with greater variations of ΔR_S, even if they have a non-linear but known dependence of ΔV_S on ΔR_S. Considering again all the resistances equal, the new value of the voltage variation is:

$$\Delta V_S = V_S - V_{S0} = V_A \frac{R_0 + \Delta R_S}{2R_0 + \Delta R_S} - \frac{V_A}{2} =$$

$$= \frac{V_A}{2} \cdot \frac{\dfrac{\Delta R_S}{2R_0}}{1 + \dfrac{\Delta R_S}{2R_0}}$$

Even if it is possible to use resistances instead of current generator, a problem still remains: in case of remote RTD operations, the long wire which connects the object to the sensor has to be considered.
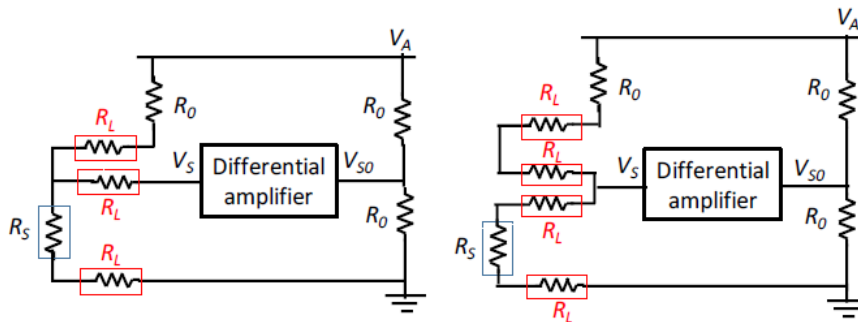
Figure 25.19: Remote RTD Operation

Two different configurations can be used. The first one, shown in figure 25.20 on the left, is the "Three-wire-connection": on the left side of the circuit there are $R_0 + R_L$ and $R_S + R_L$; the R_L in the middle has no drop since it is connected to an ideal differential amplifier, so a negligible current flows in that branch. The resistance is just added in order to be sure that all the Thomson, Seebeck and Peltier effects are equally distributed among the three wires. On the right side of the circuit, the current flows in the series of the two resistances R_0 and the overall voltage is $V_A/2$.

The second circuit (figure 25.19 on the right) represents the "Four-wire-connection" and is just a more symmetrical version than the previous one. Although with four wires the probability of introducing parasitism and interference increases, the symmetric configuration is able to perform more precise and detailed measurements.
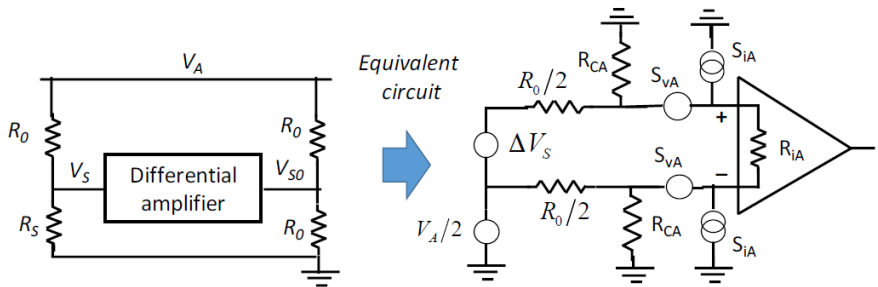
Figure 25.20: RTD preamplifer and equivalent circuit for noise computation

Since the source resistance is low, typically $R_0=100\Omega$, the input differential resistance R_{iA} and the input-to-ground resistance R_{CA} can have moderately high values. The contribution of the input current noise generators is reduced; therefore, the input voltage noise generators are dominant.

Since the differential signal ΔV_S is accompanied by a high common mode signal $V_A/2$, the CMRR should have an adequately high value at the frequency of the supply V_A. An AC supply voltage V_A can be selected with a frequency of several kHz in order to reduce the 1/f noise contribution.

26.3 Thermistors

Thermistors are temperature sensor based on the same principle of RTD, i.e. using the variation of the relative resistance variation of a certain material to measure the variation in temperature. Nevertheless, they present a different behavior: as the temperature increases, the resistance decreases, because thermistors are not made of the same materials as RTDs. In particular, semiconductor ceramic elements and oxides of Cr, Mn, Fe, Co and Ni are the most used ones. The relative resistance variation as function of temperature is much higher than in RTDs; indeed, while in RTDs it goes from ≈ 1 at standard reference temperature (0°C) to ≈ 4 in about 800°C, with thermistors ΔR_S goes from ≈ 4 to less than 1 in a temperature range smaller than about 5°C. The resistance-temperature relationship can be described by the equation:

$$R = \exp(B/T)$$

where T is the absolute temperature in Kelvin degrees, B is constant; B is the so-called characteristic temperature of the thermistor and usually ranges from 2000K to 4000K. Considering also the reference resistance value R_0 at a known temperature T_0, the final expression of the resistance becomes:

$$R = R_0 \exp\left[B\left(\frac{1}{T} - \frac{1}{T_0}\right)\right]$$

The drawbacks of these kinds of sensors are the strongly non-linear characteristic, the lower accuracy and lower reproducibility, which limit the application of thermistors in automatic control systems and in integrated circuits.

The big advantage, instead, is the possibility to have large variations of the resistance in a very small temperature range.

Strain Gauges

Strain gauges are sensors used to measure the deformations induced in a body when an external force is applied. A brief physical explanation of the strain gauge principle is given in this chapter. The main features and limitations of a metal thin-foil strain gauge are described. Some measurement and compensation techniques are presented. Finally, the performance of metal strain gauges is compared to that of semiconductor ones.

26.1 Introduction

A strain gauge (or strain gage) is a resistor, made of metal or semiconductor material, used to acquire information about strain. In fact, when an external force is applied to a resistor, its resistance value can change due to both the mechanical deformation and to a phenomenon called piezo resistivity. The change is deterministic: equations can be derived to link the resistance variation to the applied deformation. The resistance change can be simply linked to an electrical variable, which can be measured to obtain information about the strain. However, the measurement procedure is not trivial, and several factors should be taken into account to obtain a sufficient level of sensitivity and avoid reading errors. Aspects like temperature of operation, adhesive used to bond the sensor to the surface, length of the connection wires, and so on, will all influence the measurement in different ways and must be properly compensated. A single optimum strain gauge does not exist: instead, several options are available, and for each measurement set-up the best suited one should be chosen as a trade-off among the characteristics of the sensors. In order to make the best choice in such trade-off, the main characteristic of a strain gauge and its limits should be understood.

26.2 Strain Gauge principle

For a thin slab of conductor material like the one schematized in figure 26.1, the resistance is:

$$R = \rho \cdot \frac{L}{A}$$

$$\rho = resistivity \ [Ohm \cdot m]$$
$$A = W \cdot H = section \ [m^2]$$
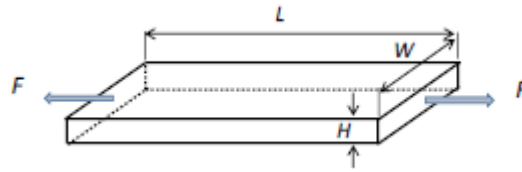
Figure 26.1: schematic view of a thin metal bar

The variation of resistance experienced when a strain is applied can be linked to these three parameters.

26.2.1 Measuring Strain

Whenever an external force F is applied to a body, reactive forces are induced within the body. These forces are distributed throughout the body: **Stress (N)** is the force density per unit area:

$$N = \frac{F}{A}$$

$F = force\ [N]$
$A = section\ area\ [m^2]$
$N = stress\ \left[\dfrac{N}{m^2}\right]\ usually\ expressed\ in\ [MPa]$

A stress can be *tensile* (if it tends to pull the atoms of the body farther apart from each other), *compressive* (if it pushes the atoms of the body closer to each other), or a *shear stress* (if it tends to produce relative slip of atomic planes on one another). As an effect of the applied stress, the dimensions of a body change. *Strain (ε)* is the relative change in dimension of the body, with respect to the original size.

$$\varepsilon = \frac{\Delta L}{L}$$

$\Delta L = extension\ (or\ compression)\ due\ to\ F\ [m]$
$L = original\ dimension\ [m]$
$\varepsilon = strain\ [\%]$

Deformation depends on the orientation of the applied force. It can be reversible, so that the body returns to the original size when the force is removed (elastic range) or irreversible, if the deformation is persistent even after the force is not acting on the body anymore (plastic range).

26.2.2 Elastic module & Poisson module

Let us consider the case of applying a tensile stress to a thin metal bar of length L at rest; the section of the bar is $A = W \cdot H$. A pulling force (F) is applied to the two sides of the bar, like in figure 26.1, resulting in a tensile stress.

For small ε, the regime is elastic: Strain (ε) is linearly proportional to stress (N) via a constant called *young modulus* (E). For larger deformations (plastic regime), the dependence of N on ε is not linear anymore.

The young modulus of a material is inferred as the slope of the curve in the Stress/Strain diagram obtained with a tensile test: it is a proper characteristic of each material, as well as the $ε_L$ that sets the limit to the elastic range.

Moreover, whenever a tensile stress is applied, due to Poisson effect, the section dimensions (H and W in the example) will reduce as length increases. The section contraction is proportional to ε via the *Poisson module (v)* – adimensional – which is also a property of the material considered.

$$\frac{\Delta W}{W} = \frac{\Delta H}{H} = -v \cdot \varepsilon$$

It follows that the deformation of the section area $A = W{\cdot}H$ will be:

$$\frac{dA}{A} = \frac{dWH + dHW}{WH} = \frac{dW}{W} + \frac{dH}{H} = -2 \cdot v \cdot \varepsilon$$

So, passing to the discrete world:

$$\frac{\Delta A}{A} \approx \frac{\Delta W}{W} + \frac{\Delta H}{H} = -2 \cdot v \cdot \varepsilon$$

26.2.3 Piezoresistive effect

The *piezoresistive effect* is a change in the resistivity of a material when mechanical strain is applied. In contrast to the piezoelectric effect, the piezoresistive effect causes a change only in electrical resistance, not in electric potential therefore, to exploit piezo resistivity in a sensor, an external voltage source is needed.

In metals, the effect is moderate, and it only contributes in a marginal way to the overall change of resistance when an external force is applied. The link between resistivity and stress is linear, so it is possible to define a piezo resistivity coefficient $β$ such that:

$$\rho = \rho_0(1 + \beta N)$$

$$\frac{d\rho}{\rho_0} = \beta N = \beta E \varepsilon$$

Where ρ_0 is the nominal value of resistivity, when no external strain is applied.

For semiconductors, instead, piezo resistivity is much stronger and resistivity has a non-linear and fairly complex dependence on strain.

26.2.4 Gauge factor

Considering a metal strain gauge, the previous relationships can be exploited:

$$R_0 = \rho_0 \cdot \frac{L_0}{A_0}$$

$$\frac{dR}{R_0} = \frac{d\rho}{\rho_0} + \frac{dL}{L_0} - \frac{dA}{A_0} = \beta E \varepsilon + \varepsilon + 2\nu\epsilon = \varepsilon(1 + 2\nu + \beta E)$$

Therefore, if resistance variation ΔR is small, the following equation is valid:

$$\frac{\Delta R}{R} = \varepsilon (1 + 2\nu + \beta E)$$

From which we can define the Gauge factor (G) as follows:

$$G = \frac{\frac{\Delta R}{R}}{\varepsilon} = (1 + 2\nu + \beta E)$$

The gauge factor represents the gain of the sensor system and it varies depending on the shape of the strain gauge and the material adopted. For metal strain gauges, typical values of G are between 1 and 3, and it is constant for a wide range of operation environments. For strain gauges fabricated with semiconductors, G is typically higher (G values of some hundreds are achievable), but the value is usually less stable than the case of metal. A higher G is desirable as it would allow to measure strains with higher resolution, but it is not the only factor influencing the sensitivity of a strain-gauge measuring system, as will be clarified in the following sections.

26.3 Design and Technology

The most intuitive and straightforward way to implement a strain gage is just a piece of wire glued to the surface of the body where we want to measure the strain. To have a larger resistance (thus lower power consumption) one could simply place the piece of wire curved in alternate direction, forming S shapes, like in figure 26.2 (Wound-wire SG).
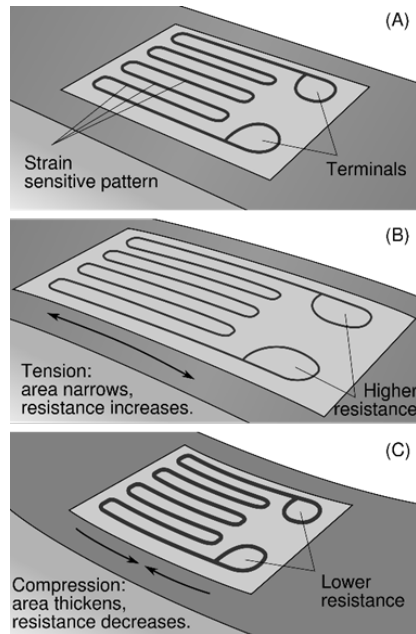
Figure 26.2: schematic view of a wound-wire strain gauge

Longer portions of the wire should be aligned with the principal axis (the axis along which the strain must be measured), so the shorter portions would be on the transvers axis. If a tensile stress is applied, R increases on the main axis (proportional to strain via the Gauge factor: $\frac{\Delta R}{R} = G\varepsilon$); however, the portion of wire along the transversal axis will show a change of resistance with the opposite sign, as it will experience an opposite deformation due to Poisson effect.

The ratio between the change of resistance along the transversal axis and the change along the principal axis is called *cross-sensitivity*: this parameter should be as low as possible. It is roughly proportional to the ratio of the total length of the wire portions on the two axes.

Most of modern strain gauges are manufactured with Lithographic technology. Strain gauges fabricated with this technology are called thin foil strain gauges: the manufacturing process consists of laying a thin metal foil of the sensing material onto a carrier (usually a plastic or glass material), and then cutting away portions of the foil, to obtain the grid pattern required. Thin-foil technique is preferred to simple wire implementation, as it shows several advantages:

- Larger resistance is achievable, since the foil can be made much thinner than a wire.
- Surface area is larger if the cross section is a rectangle, instead of a circle. This grants better heat management (heat diffusion from surface) and easier application on the carrier

- Lower cross sensitivity can be achieved, because transversal portions of the meander can be wider, to have a lower resistance and so a lower contribution to transversal gauge factor
- Easier lead connection, as the connection pad can be realized of the desired size
- Precise manufacturing process with lithographic technology grants high reproducibility. This is a key point when matched strain gauges are required to implement compensation techniques.
- Smaller overall dimensions are achievable, so localized stress can be measured.
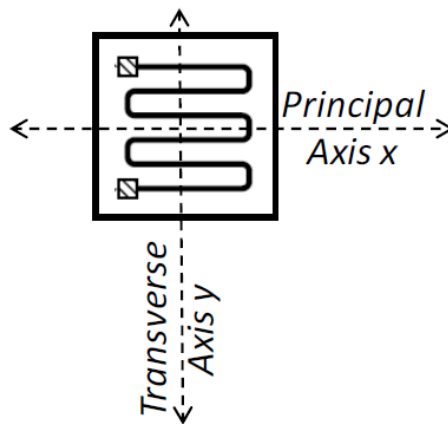
Figure 26.3: schematic view of main and transverse axis in a wound wire strain gauge

26.3.1 Strain gauge parameters

When selecting a strain gauge for a particular application, several different constraints should be met. Such constraints can be expressed as requirements that the strain gauge should fulfil regarding:

- *Accuracy* of the measured value, for instance: absence of zero drift
- *Stability* of the Gauge factor throughout the measurement, in spite of effects such as creep or temperature drift
- *Temperature*: immunity to temperature drift in a given temperature range
- *Elongation* that the sensor should withstand in that application
- *Test duration*: the strain gauge should maintain its properties unaltered for the whole duration of the test
- *Cyclic endurance*: fatigue-related failure should not show before the end of the test

- *Ease of installation*: for instance, if the site where the sensor should be installed is very narrow or has a curvature with small radius, the strain gauge should have the needed flexibility and size
- *Environment*: the materials used should be compatible with the environment where they are used. For instance, waterproof strain gauges are needed for some applications.

The parameters of choice for a sensor are the following:
  o Strain-sensitive Alloy: the material with which the measuring grid is fabricated
  o Backing material (carrier): usually an insulating material, it is a scaffold for the sensor.
  o Grid resistance: a higher resistance grants lower power consumption and thus a better heat management, but requires a larger size
  o Gauge pattern: the choice strongly depends on the orientation of the strains in the body
  o Self-temperature compensation number: a parameter of compatibility between a strain gauge made in a temperature-compensating alloy and the body, related to the thermal expansion coefficients of the materials involved
  o Gauge length: depends on the application and it influences resistance and resolution

The selection of a gauge for a certain purpose is not trivial for two main reasons. First, because even if these parameters may vary at some extent, it is not possible to choose them in any combination at will. In fact, manufacturer usually realize strain gauges as complete systems, comprised of a choice of foil/carrier combination and determined features such as pads material or shape or (if present) lead wire length. Moreover, the choice of a parameter that satisfies a certain requirement is often detrimental for other aspects. For example, to measure the strain on a small filet, where the surface area is limited, requires a very small device. Such strain gauge would unavoidably have a reduced stability and fatigue life, as well as a small resistance. This example shows that compromises are always needed when choosing a device and understanding the effect of such compromises is necessary for making the best choices in each situation.

Alloys:
The selection of the alloy is usually a balanced trade-off considering gauge factor stability against temperature and fatigue, optimal value of resistivity, temperature range and sensitivity. Some alloys are mentioned as an example.
Constantan is the common name of a copper–nickel alloy. Of all modern strain gauge alloys, constantan is the oldest, and still the most widely used. The reason is that

constantan features a combination of properties which make for an optimum trade-off for several applications: gauge factor is adequately high and insensitive to strain level or temperature; resistivity is high enough (*49.0 x 10^{-8} Ω·m*) to achieve suitable values of resistance even in small grids. The main feature (that also gives the name to this alloy) is the fact that resistivity is quite constant over a wide range of temperatures. Constantan sensor can typically measure up to 20000 microstrains. However, even if fatigue life of constantan is quite good, it exhibits permanent change of resistance under high cyclic strains, which causes a zero shift in the strain gauge. An *Isoelastic Alloy* (also called D alloy by some manufacturers) is preferable when purely dynamic measurements have to be made, that is, if it's not necessary to maintain a stable reference zero. One of D-alloy main features is a greater fatigue life, together with high gauge factor. However, its thermal output is also very high, thus it cannot be used unless a system level temperature compensation technique can be adopted. Other alloys, like *K Alloy* from VPG Micromeasurements, are designed to offer a quite flat temperature response and great fatigue life in quite extreme temperature context (–269° to +260°C), but the installation technique of such strain gauges is more difficult than others.

Fatigue life
Just like any other metal, the sensing grid of a thin foil strain gauge is subject to fatigue damage when strained cyclically at a sufficiently high strain level. Fatigue damage is a critical issue in strain gauges, influenced by several different factors that can impair the measurement in different ways. For instance, the unstrained resistance of the sensor may change due to fatigue, causing a signal at the output even when no signal is applied ("zero shift"). This causes reading errors when a static strain should be measured, but the effect may be negligible for dynamic measurements. However, with cyclic simulation, microfractures may appear in the sensing grid, that would affect dynamic measurements. Consider as an example a sinusoidal strain: the cracks would appear only when the tensile strain is applied, but they would close back in unstrained condition: their effect is therefore visible only for dynamic measurements, as the resulting sine wave would appear distorted. Fatigue damage may also cause variation in gauge factor, that may appear increased for tensile stresses and decreased for compressive forces.
The strain gauge bonding pads are where the risk of fatigue-related damage is higher, because of the abrupt shift in section of the grid. A good practice to limit this effect is to mechanically decouple the pads from the sensing grid, or at list to position the sensor so that they are subject to reduced strains. The presence of an encapsulating cover can reduce fatigue related damages. The number of cycles that a sensor can withstand without failing is specified by the manufacturer and is called fatigue life. A possible definition for fatigue life is the number of cycles after which the sensor shows a zero shift of $100\mu\varepsilon$: typical values for metal strain gauges are around 10^7 cycles.

Carriers

The support serves several functions at once: it provides a mean for handling the foil pattern during installation, as well as electrical insulation between the sensing foil and the test object. Also, it presents a readily bondable surface for adhering the gauge to the test specimen. The backing material should also be unaffected by temperature changes, while being thin and strong enough to transmit the force from the test object to the sensing wire. As already mentioned above, carrier and sensing material are not two independent parameters. Instead, strain gauges are designed as systems and only some combinations of alloy and carriers are available. Manufacturers usually offer different strain gauge "series" with different combinations, and for each series the main characteristics and suggested applications are listed. Typical carrier materials are polyimide – a polymeric material with high heat resistance (temperature range from -195° to +175°C) and capable of elongations up to 20% – or other glass-fiber-reinforced materials, which may withstand even more extreme temperatures (up to -269°C +290°C), but for smaller strains (around 2%).

Glues

The choice of the adhesive to fix the strain gauge to the test object is not trivial and cannot be overlooked: the glue must accurately transmit the strains from the body to the sensor. One of the critical parameters is the application temperature range of work of the sensor, as some glues are cold-curing while other are hot-curing, and a temperature shift may alter the adhesive properties of the chosen glue. Moreover, the adhesive should be insulating and have high resistance to creep. Manufacturer usually include a section about glues in their catalogues, and in the description of each strain gauge series the best suiting glue is suggested. The surface of the test object where the strain gauge is mounted must often undergo a preparation procedure before bonding. Such procedures are suggested by strain gauge manufacturers too.

Connections

Integrated solder pads may be present at one end of the device, allowing direct soldering on the strain gauge. This makes the connection process very easy, but in this way the solder pads are not mechanically decoupled from the sensing grid and a strain would induce a change of resistance in the pads too, causing possible measurement errors. Mechanical decoupling can be partially achieved with bigger solder pads shaped in a way that would make them less sensitive to strain. Alternatively, some SG are provided with wires directly attached to them, and offer full mechanical decoupling between grid and connector, but a long cable would cause a change in the nominal resistance of the device, so different compensation techniques may be necessary at system level, to avoid static errors.

Patterns

The gauge pattern refers cumulatively to the shape of the grid, the number and orientation of the grids in a multiple-grid gauge, and the solder tab configuration. For a single-grid gauge, connectors size and placement should be compatible with the application (consider again the example of measuring the strain on a thin metal filet). The main axis should be oriented along the direction of the strain; thus, such solution can be employed only when the direction of the strain is known a priori. A trade-off ensues also when considering the size: a narrower grid would minimize the averaging error due to the presence of strain gradients in directions perpendicular to the main one; on the other hand, a wider grid would allow better heat dissipation and higher resistance (thus, lower power dissipation).

When the direction of the strain to be measured is not known and cannot be determined a priori, instead, a multielement rosette is required. Many different configurations are available, as sketched in figure 26.4.
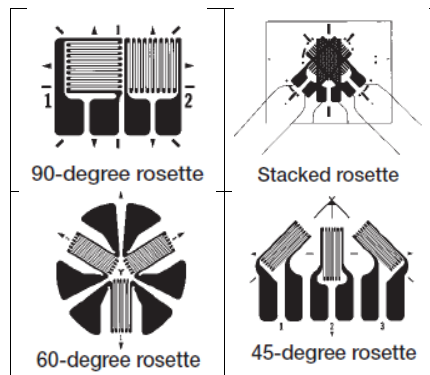
Figure 26.4: some possible configuration of multi-grid strain gauges, from VPG

A two-element rosette is typically used when the direction of the main strain can be at least supposed from the test object geometry. In other cases, a three-element rosette can be used to determine the direction. A single-plane rosette (60-degree or 45-degree) is superior to a stacked rosette configuration in terms of heat transfer (since there is more surface available for power dissipation) and more accurate when perpendicular strain gradient is also present (for example: bending), since all grids are as close as possible to the surface. On the other hand, a stacked rosette is advantageous when there are large strain gradients in the plane of the test subject: since the area is reduced, the averaging error will be also reduced.

Summary

Strain gauge manufacturers usually provide guides and tables where all available combinations are listed, with the relative parameters of temperature range, strain

range and fatigue life in different operating conditions. As an example, a table taken from the website of VPG Micromeasurements is reported in figure 26.5.

| Type of Test or Application	Operating Temperature Range	Test Duration in Hours	Accuracy Required **	Cyclic Endurance Req'd		Typical Selection	
				Maximum Strain, $\mu\varepsilon$	Number Of Cycles	Gage Series	M-Bond Adhesive
General Static or Static-Dynamic Stress Analysis*	−50° to +150°F (−45° to +65°C)	$<10^4$	Moderate	±1300	$<10^6$	C2A, L2A, W2A, CEA, EA	200 or AE-10
		$<10^4$	Moderate	±1300	$<10^6$	C2A, L2A, W2A, CEA, EA	AE-10 or AE-15
		$<10^4$	Very High	±1600	$<10^6$	WA, SA	AE-15 or 610
		$<10^4$	High	±2000	$<10^6$	WK, SK	AE-15 or 610
	−50° to +400°F (−45° to +205°C)	$<10^3$	Moderate	±1600	$<10^6$	WA, SA	600 or 610
		$<10^3$	High	±2000	$<10^6$	WK, SK	600 or 610
	−452° to +450°F (−269° to +230°C)	$<10^3$	Moderate	±2000	$>10^6$	WK, SK	610
	<600°F (<315°C)	$<10^2$	Moderate	±1800	$<10^6$	WK	610
	<700°F (<370°C)	<10	Moderate	±1500	$<10^6$	WK	610
High-Elongation (Post-Yield)	−50° to +150°F (−45° to +65°C)	<10	Moderate	±50 000	1	CEA, EA	AE-10
		$>10^3$	Moderate	±100 000	1	EP	AE-15
		$>10^3$	Moderate	±200 000	1	EP	A-12
	0° to +500°F (−20° to +260°C)	$<10^2$	Moderate	±15 000	1	SA, SK, WA, WK	610
	−452° to +500°F (−269° to +260°C)	$<10^3$	Moderate	±10 000	1	SK, WK	600 or 610
Dynamic (Cyclic) Stress Analysis	−100° to +150°F (−75° to +65°C)	$<10^4$	Moderate	±2000	10^7	ED	200 or AE-10
		$<10^4$	Moderate	±2400	10^7	WD	AE-10 or AE-15
	−320° to +500°F (−195° to +260°C)	$<10^4$	Moderate	±2000	10^7	WD	600 or 610
		$<10^4$	Moderate	±2300	10^6	WD	600 or 610
Transducer Gaging	−50° to +150°F (−45° to +65°C)	$<10^4$	1 to 5%	±1300	$<10^6$	CEA, EA	AE-10 or AE-15
		$<10^6$	1 to 5%	±1300	$<10^6$	CEA	AE-15
	−50° to +200°F (−45° to +95°C)	$<10^4$	Better than 0.2%	±1500	10^6	N2A	600, 610, or 43B
	−50° to +300°F (−45° to +150°C)	$<10^4$	0.2 to 0.5%	±1600	10^6	WA, SA	610
	−320° to +350°F (−195° to +175°C)	$<10^4$	Better than 0.5%	±1800	10^6	WK. SK	610

* This category includes most testing situations where some degree of stability under static test conditions is required. For absolute stability with constantan gages over long periods of usage and temperatures above +150°F (+65°C), it may be necessary to employ half- or full-bridge configurations. Protective coatings may also influence stability in cases other than transducer applications where the element is hermetically sealed.

** It is inappropriate to quantify "accuracy" as used in this table without consideration of various aspects of the actual test program and the instrumentation used. In general, "moderate" for stress analysis purposes is in the 2 to 5% range, "high" in the 1 to 3% range, and "very high" 1% or better.

Figure 26.5: table from a VPG Micromeasurements technical note, listing the most suited strain gauge series to adopt in different situations, together with the suggested adhesive and the strain gauge main parameters

26.4 Electronics & Temperature compensation

A strain applied to a strain gauge causes a variation of the resistance value which, at least for a limited range of values, follows a linear relationship: $\Delta R = R_0 G \varepsilon$
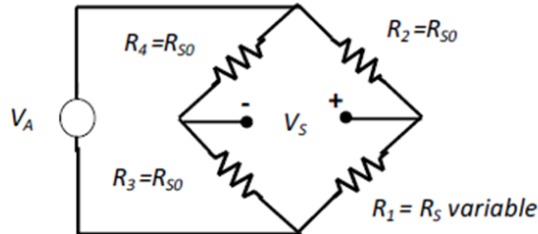
Figure 26.6: the Wheatstone Bridge configuration

In the previous chapter we studied this configuration; in this case we obtain:

$$\Delta V_S = \Delta R_S \left(\frac{dV_+}{dR_S}\right)_{R_2 = R_0} = \frac{1}{4} V_A \frac{\Delta R_S}{R_0} = \frac{1}{4} V_A G \varepsilon$$

$$\frac{V_S}{\varepsilon} = \frac{1}{4} V_A G \quad \left[\frac{V}{\mu\varepsilon}\right] = Sensitivity$$

This relationship shows that sensitivity can be increased with a better strain gauge (with higher gauge factor) o choosing a higher level of power supply.

If a measurement requires a high level of precision or if the relationship $\frac{\Delta R_S}{R_0} \ll 1$ is not valid, non-linearities have to be considered. However, strain gauge manufacturers like Vishay provide mathematical tools to correct for errors like this. From the computation, it is confirmed that, for small strain values, the error is zero, but it becomes non negligible for larger values of ε.

Sensitivity can be doubled if two sensors are employed, instead of just one. The use of two strain gauges with astute positioning also allows the separate measurement of different components of strain even if they are present simultaneously.

As an example, consider the case of a thin metal bar subject to bending (**blue arrows**) and compression (**red arrows**) forces at the same time (figure 26.8).

To maximize sensitivity, the two strain gauges should be positioned on the two opposite sides of the metal bar. In this case:

- The effect of compression would be the same for bot strain gauges: a compression will reduce the nominal resistance value (while an extension would increase it).

- Bending would have opposite effect on the two SG: the upper part of the bar would be extended, so R_{SU} would experience and increase in resistance; the opposite is true for the lower part, thus the resistance of R_{SL} would decrease instead.
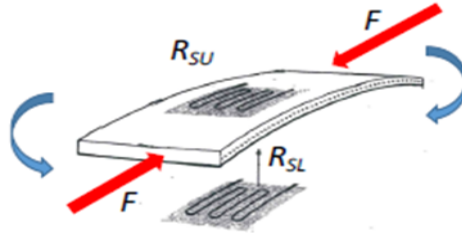

Figure 26.7

First, consider the case when only *bending* must be measured. As mentioned above, a compression strain causes a resistance variation in the same direction in both sensors: it is a common mode signal. If $R_1 = R_{SU}$ and $R_3 = R_{SL}$, the common mode variation induced will cause the two voltage partitions to vary by the same amount, that means: $V_S = V_+ - V_- = 0$.

On the other hand, the effect of bending is the opposite for the two sensors and will cause V_+ to increase and V_- to decrease. If we consider a small resistance variation, $\frac{\Delta R_S}{R_0} \ll 1$, the linear approximation is valid, and the output voltage can be computed exploiting the principle of superposition of effects: the contribution to the output of the two strain gauges can be computed independently and then summed up.

Considering first only $R_1 = R_{SU}$, the output variation due to a resistance variation ΔR_S has already been computed and it is:

$$V_{S,R_1} = \frac{1}{4} V_A \frac{\Delta R_{SU}}{R_0}$$

The contribution to the output voltage of $R_3 = R_{SL}$ can be easily computed and it is:

$$V_{S,R_3} = V_A \left(\frac{1}{2} - \frac{R_0 + \Delta R_{SL}}{2R_0 + \Delta R_{SL}} \right) \approx -\frac{1}{4} V_A \frac{\Delta R_{SL}}{R_0}$$

For compression/extension, $\Delta R_{SL} = \Delta R_{SU}$; therefore, when summing the two contributions, it results, as expected:

$$V_S = V_{S,R_1} + V_{S,R_3} = 0$$

If a bending strain is considered instead, the upper part of the metal bar will extend due to the bending, while the lower part will shrink. The resistance variations of the two strain gauges can be considered equal in absolute value but they will have opposite sign:

$$|\Delta R_{SU}| = |\Delta R_{SL}| = \Delta R_S$$
$$\Delta R_{SU} = -\Delta R_{SL}$$

This leads to:

$$V_S = V_{S,R_1} + V_{S,R_4} = \frac{1}{4} V_A \frac{\Delta R_S}{R_0} + \frac{1}{4} V_A \frac{\Delta R_S}{R_0} = \frac{1}{2} V_A \frac{\Delta R_S}{R_0}$$

$$V_S = \frac{1}{2} V_A G \varepsilon$$

That means sensitivity is doubled for the detection of bending strains, with respect to the case when only one SG is employed. Another possibility to detect bending without detecting compression/elongation would be placing the two strain gauges in the same branch of the Wheatstone bridge (for instance $R_{SU} = R_1, R_{SL} = R_2$). The computation steps would be similar to the ones reported for the previous case.

For measuring *compression/elongation* instead, R_{SL} should be placed in the bridge as R_4, while R_{SU} is still used as R_1. In this configuration, the effect of the bending strain on the output will be nil. This can be shown with a reasoning similar to the previous one, considering the two strain gauges one at a time. The resistance variation of the two strain gauge will be equal in magnitude, but with opposite sign:

$$|\Delta R_{SU}| = |\Delta R_{SL}| = \Delta R_S$$
$$\Delta R_{SU} = -\Delta R_{SL}$$

The effect of R_{SU} is the same as above, being it in the same position:

$$V_{S,R_1} = \frac{1}{4} V_A \frac{\Delta R_{SU}}{R_0}$$

Let us compute the effect of a variation of $R_4 = R_{SU}$:

$$V_{S,R_4} = V_A \left(\frac{1}{2} - \frac{R_0}{2R_0 + \Delta R_{SL}} \right)$$

$$V_{S,R_4} = V_A \left(\frac{2R_0 + \Delta R_{SL} - 2R_0}{4R_0 - 2\Delta R_{SL}} \right) \approx \frac{1}{4} V_A \frac{\Delta R_{SL}}{R_0}$$

It follows, substituting $\Delta R_{SU} = \Delta R_s$ and $\Delta R_{SL} = -\Delta R_s$:

$$V_S = V_{S,R_1} + V_{S,R_4} = \frac{1}{4}V_A\frac{\Delta R_S}{R_0} - \frac{1}{4}V_A\frac{\Delta R_S}{R_0} = 0$$

While for compression/extension, since $\Delta R_{SU} = \Delta R_{SL} = \Delta R_s$, the result would be again:

$$V_S = V_{S,R_1} + V_{S,R_4} = \frac{1}{4}V_A\frac{\Delta R_S}{R_0} + \frac{1}{4}V_A\frac{\Delta R_S}{R_0} = \frac{1}{2}V_A\frac{\Delta R_S}{R_0}$$

That is: sensitivity is doubled. Following the same reasoning, a full bridge configuration may be adopted, where all four of the resistors are active sensors. This would further increase sensitivity, doubling the result obtained:

$$V_S = V_A\frac{\Delta R_S}{R_0} = V_A G\varepsilon$$

26.4.1 Voltage supply

It is evident from the previous examples that the output is related linearly to the power supply value V_A: one might consider to just increase this value as much as possible to increase sensitivity. Although this would work in principle, non idealities must be taken into account when considering this option. In fact, a constraint of max power dissipation on each sensor must be observed, in order to limit grid self-heating that would cause measurement errors. There is no standard value for the optimum power supply to be adopted, as it depends on the sensors used (max power is specified by the manufacturer) and the conditions of the measurement. For instance, if the measuring grid has a large surface and the test object is made of metal and connected to a heat sink, the requirements on maximum heat tolerable are relaxed. If strains must be measured on a plastic object, instead, it is critical to keep the heating very low, as plastic is an insulator and a change in temperature might also influence the properties of the glue, impairing the measurement in more than a way.

For a Wheatstone bridge with power supply V_A, the mean power dissipation on a strain gauge would be:

$$P_{diss} = \frac{V_A^2}{4R_0}$$

Reversing this equation gives a first level estimation of the maximum power supply that may be adopted:

$$V_A < \sqrt{4R_0 P_{diss,max}}$$

The chosen value should be strictly smaller than the maximum, as the signal would introduce further power dissipation on the sensor. Typical values range from some hundreds of mV to a couple of Volts.

In some cases, an AC voltage supply may be used, rather than a DC one. This may be useful when the signal has a low frequency spectrum and the conditioning electronics introduces a non-negligible noise at low frequency. If a sinusoidal voltage supply is used, the resulting signal V_S will be modulated at the frequency of the power supply, and then a LIA may be used to extract the signal, escaping from 1/f noise. Even if the applied strain is already modulated (for instance, a strain caused from the vibration of a motor rotating at 15000rpm would give a signal modulated at 250Hz), voltage supply modulation can be superimposed to it and the measurement can be still performed using a LIA filtering setup. Another advantage brought by using a sinusoidal reference is that the peak value V_A can be set higher by a factor $\sqrt{2}$ than the one computed for the static reference to respect the constraint on maximum power. In fact, the power delivered to a resistor R_0 by a sinusoidal power supply of amplitude $V_A/2$ (considering the partition between the SG and the resistor, only half of the power supply falls on the sensor) is:

$$P_{diss,AC} = \left(\frac{V_A}{2\sqrt{2}}\right)^2 \frac{1}{R_0} = \frac{V_A^2}{4*2}\frac{1}{R_0}$$

$$V_{A,max,AC} = \sqrt{2*4P_{max}R_0} = \sqrt{2}V_{A,max,DC}$$

A higher peak voltage supply is advantageous, as it yields higher sensitivity.

26.4.2 Temperature compensation

Since the output of the bridge is proportional to a variation of the sensor resistance, any drift of the resistance from the nominal value will result in an output signal, regardless of the nature of the variation itself. The resistivity of a metal is temperature-dependent: $\rho = \rho_0 + \alpha\Delta T\rho_0$. This means that even if no strain is applied, a change in temperature ΔT will cause a change in resistance ΔR_S proportional to ΔT via a temperature coefficient α which depends on the metal employed for the sensing grid. For metals used in SG manufacturing, a typical value of α is around $\alpha = 4*10^{-3}/K$. A resistance variation due to a temperature shift will be interpreted at the output as a strain.

$$\left(\frac{\Delta R_S}{R_0}\right)_N = G\varepsilon$$

$$\left(\frac{\Delta R_S}{R_0}\right)_T = \alpha\Delta T$$

A temperature variation appears as a strain according to the following equation:

$$\varepsilon = \frac{\alpha \Delta T}{G}$$

Assuming $\alpha = 4 * 10^{-3}/K$ and $G = 2$, the equation leads to a factor of 2000 μstrain for each kelvin degree temperature variation.

The origin of this problem lays in the difference between the temperature coefficients of the SG and of the regular resistors adopted to complete the bridge: if they were equal, a temperature variation would cause the same resistance variation in all resistors, and thus a nil output. Following this reasoning, one way to compensate for this issue is to just use another strain gauge as a resistor, but oriented in a way so that it is insensitive to strain.

Strain Gauge manufacturers usually provide two-grid strain gauges where the two sensing grids are placed with the sensitive axes one perpendicular to the other, so that only one would be sensitive to the strain (Active Gauge), while the other acts as a Dummy cell and it is only sensitive to temperature variations. Adopting a configuration like the one in figures 26.8, the output will be nil when no strain is applied.

Dummy Gauge   Active Gauge
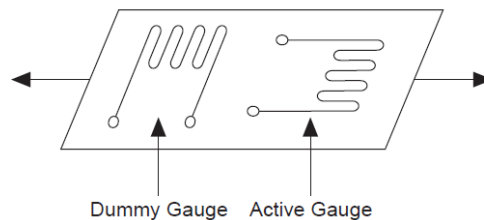
Figure 26.8: schematic view of Active and Dummy gauge configuration

If two strain gauges are needed (for instance, to measure a composite strain like bending), the other two resistors of the bridge can be used as dummy sensors to compensate for temperature effects.

A temperature variation will also cause thermal expansion, both in the test object and the strain gauge applied to it, that may lead to measuring errors if it is not the variable of interest. This issue can be circumvented adopting a strain gauge made with a sensing material for which the thermal expansion coefficient is the same as that of the test object (or at least very close to it).

26.4.3 Lead wires resistance

If the strain gauge circuit is far away from the electronics, the resistance of the connecting wires may be an additional source of error. Considering the case of a ¼ bridge, if just two wires are used to connect it to the rest of the circuit, a voltage offset will be present at the input of the differential amplifier, due to the difference in the partition branches (figure 26.9, the voltage supply is here called V_{EX} instead of V_A).
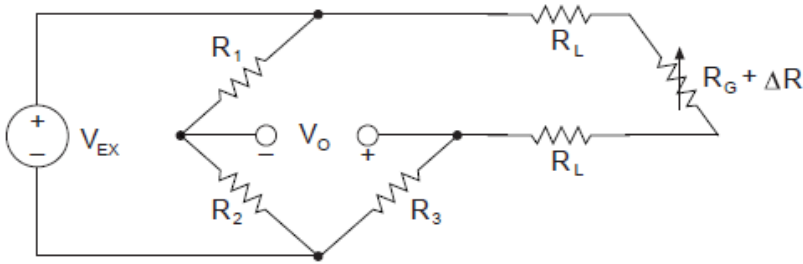
Figure 26.13: Two-wire connection of a quarter bridge circuit

Even if all the resistors in the bridge have the same nominal value R_0, the two partitions are different:

$$V_- = \frac{1}{2}V_A \text{ , while } V_+ = V_A \frac{R_0}{2R_0 + R_{L1} + R_{L2}} < \frac{1}{2}V_A$$

Although the offset may be easily accounted for during signal processing, further errors may arise from the fact that the temperature coefficient of the lead wires (usually made of copper) is greater than that of the gauges by 1-2 orders of magnitude. Therefore, a slight change of temperature can lead to a measuring error of several microstrains. Using a three-wire connection can eliminate this effect and reduce the offset to a negligible value at the same time.
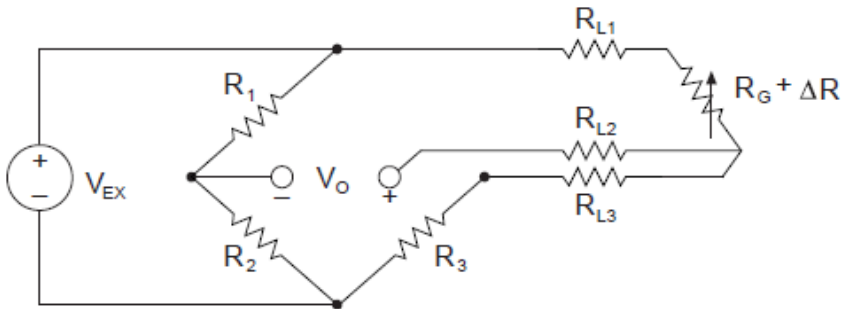

Figure 26.14: Three-wire connection

In fact, the positive pin is connected to a large impedance differential amplifier via a wire with resistance R_{L2} (see figure 26.14). However, the current flowing in R_{L2} is negligible, so we can consider that there is no voltage drop over this resistance. This means that the voltage at the positive pin is $V_+ \approx \frac{1}{2}V_A$, and any fluctuation of resistances R_{L1}, R_{L3} cancel out. If the current entering the differential amplifier is not negligible, a fourth wire may be added to make the structure perfectly symmetrical and eliminate any residual offset. A large wire resistance may also cause reading

errors due to voltage drops along the wire itself: this issue can be accounted for and compensated during the signal processing phase once again using mathematical tools provided by the manufacturer.

26.5 Semiconductor Strain Gauges

As already mentioned in the first section, piezo resistivity is a prominent effect in semiconductors like Si and Ge, therefore such materials are commonly employed for the manufacturing of strain gauges. The high piezo resistivity allows the fabrication of sensors with a gauge factor 5 to 10 times higher than that achievable with metal strain gauges. Moreover, semiconductor sensors can be much smaller (less than 1mm) than metal ones, and with higher resistance, as the resistivity of silicon is some order of magnitude higher than that of constantan. They can be directly integrated in a circuit, for measuring very small local strains (down to $10^{-2}\mu\varepsilon$) with low power consumption. Performance is also improved for what concern hysteresis: semiconductors are less subject to creep, so fatigue life is increased. However, since piezo resistivity in semiconductors is a complex phenomenon, the resistance variation has a non-linear dependence on strain, i.e.: the gauge factor markedly varies with the level of the applied strain and with temperature, with deviations of 10%-20% from the nominal value. For this reason, the dynamic range of measurable strain is limited, and manufacturers usually provide correction coefficients to apply to the values obtained to compensate for non-linearity. Gauge-factor magnitude and sign can be controlled by doping level and sign, since it is related to the mobility of carriers in the crystal: p-type doped Si has a positive G, whereas n-type shows negative G.

This allows to implement temperature compensated Wheatstone bridge structures, where sensitivity is increased with the collaboration of two or more sensors that react in opposite way to the same strain.

Characteristic	Metal Strain Gauge	Silicon Strain Gauge
Size for 1kOhms	$0.04 inch^2$	$0.0004 inch^2$
Gauge Factor	$1 - 3$	$50 - 300$
Resistance	120 – 5kOhms	120 – 10kOhms
Resistance Temperature coefficient	1% per 100F	10 – 20% per 100F
Gauge Factor Temperature coefficient	-3% per 100F	1 – 20% per 100F
Dynamic range	$500 - 10.000\ \mu\varepsilon$	$500 - 1.000\ \mu\varepsilon$
Typical Resolution	$0.1\ \mu\varepsilon$	$0.01\ \mu\varepsilon$

Table 26.1: summary of semiconductor and metal strain gauge main characteristics

In summary, semiconductor strain gauges may be an optimal choice when very small strains over a limited range should be measured, as they offer great sensitivity, therefore high resolution is achieved. However, the measurements require additional care with respect to the case of metal strain gauges, as the many non-linearities may be a great source of error and more correction steps are necessary.

A comparison of the main figure of merit of metal SG and semiconductor SG is reported in table 26.1 below (data taken from Kulite Semiconductor Products inc.).

CHAPTER 26

This book is intended not just to lead students to know and properly describe the electronic techniques and instrumentation developed for recovering sensor signals from noise, but rather to gain a good insight in the problems faced and in the approaches developed for overcoming them. This implies to critically evaluate the solutions, avoiding the attitude where sensors and electronics are designed and employed just according to the established rules and standards. It is instead necessary to clarify the reasons of the choices and decisions in the light of the physics of phenomena involved, of the basic principles of signal and noise processing and of the actual performance of the available devices. It is necessary to clearly distinguish the intrinsic limitations set by physical laws from the current limitations set by the state of the art, which can be overcome by the technological progress. In essence, to gain insight means to progress at the pace of the technology evolution and be able to contribute to it.