SAMPLING

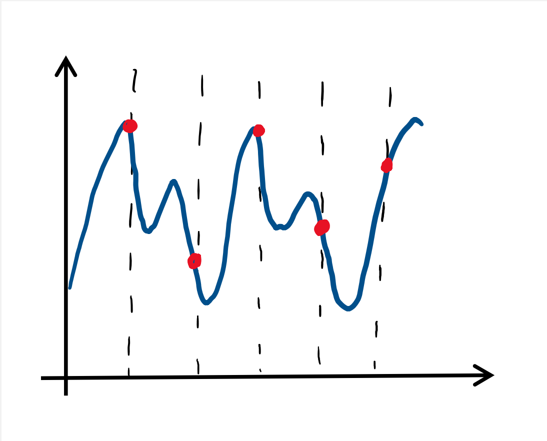
Introduction

The focus of the course will be not the instrumentation sites, but the understanding of complex signals, we'll see algorithms and techniques to extract information from signals. First of all, we must study how the acquisition of data works.

Acquisition of data (or sampling): introduction

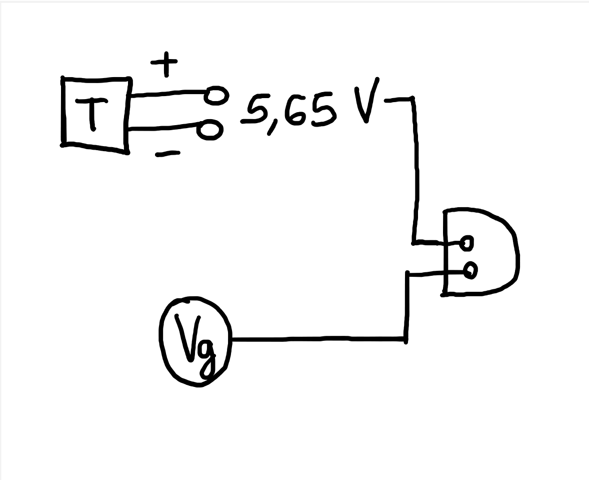
In general, we want to transform the information that we are measuring (like the temperature, for example) into a voltage signal because it's easier to carry the signal far apart from where the measurement takes place and so it's easier to collect all the data. Usually, we would like to obtain a linear relationship between the physical phenomenon and the output signal.

Let's imagine having a plot that represents all the data that we have collected as the one depicted in the picture. From this plot we can extract some information like, for example, the maximum value or the minimum value, but we are not able to obtain other parameters like the average due to the fact that we have a continues measurement that is not a function. In case of a function we would be able to compute the average since it is defined as the ratio between the integral of the function itself and the intervals:



$$avarage = \frac{\int f(t)}{T}$$

So, we have understood that one of the main problems is the fact that we are dealing with a continuous measurement and not a function. Another important element to consider is trying to understand how we should choose the time interval: this choice is up to us and it depends on what we think is meaningful. Depending on how many time intervals we choose we could increase or decrease the resolution[1] of our measurement that can be defined as the minimum variation in the input that causes a variation of the output, in other words it is the minimum variation that causes a variation of the output. Anyway, after choosing the time interval we can extract samples data[2]. With the term "sampling[3]" we mean understanding what we are seeing and reducing the information by extracting what we think is relevant. When we sample, we are losing some information. In general, we are losing whatever is in the between of the time interval that we have chosen. Since we are dealing with complex phenomena and with fast phenomena, we are going to see a device that allow the process in an automatic way; this device is the ADC converter. Let's now try to understand how it works.



Acquisition of data (or sampling): how it works

Let's, now, imagine having a transducer (T) that has two wires which are supposed to carry electricity in terms of voltage. How can we translate that voltage into a number? We can bring the voltage coming from the transducer into a component that compares that value with a standard one; this component is called SAR (Successive Approximation Register). The standard voltage used to make the comparison can be obtained

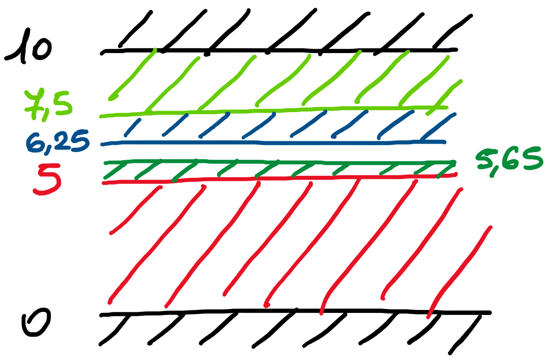
---

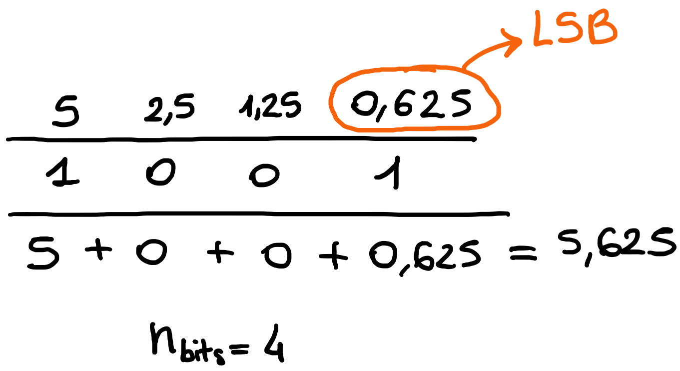
[1] Sensibilità

[2] Valori di campionamento

[3] Campionamento

thanks to another devise, called voltage generator (V_g). Since the voltage generator is able to generate a voltage between two number, we usually first generate a voltage in between these two values, then we compare this value that has been generated with the value that comes from the transducer: by doing a comparison between these two values we can understand if the voltage coming from the transducer is higher or lower than the one coming from the voltage generator. By repeating these steps, we end up finding our value; generally, the number of comparisons is equal to the number of bits. Of course, we can increase our resolution by improving the number of the intervals, or, in other words, by increasing the number of bits. Form these considerations we can easily understand that the resolution of our measurements is strictly connected to the number of bits of the converter. We'll see in just a minute the definition of resolution. Anyway, thanks to this technique of comparison we can also obtain a code in 0 (if the value that comes from the transducer is lower than the generated one) and 1 (if the value that comes from the transducer is higher than the generated one) that is easy to send to a computer.





Of course, the number that we obtain by applying this technique is usually different form the real one, this error, that can be defined as the difference between the real voltage and the output result of our measurement is called quantization error and it is at maximum equal to $\pm LSB/2$ (we'll see that LSB stands for Least Significant Bit, for us it is the same of resolution):

$$quantization\ error = real\ voltage - output\ result$$

This method of obtaining a voltage output is called, as already mentioned, ADC which stands for Analogue Digital Converter. We can foresee how much the resolution is by taking the range of the converter and divided it by 2 to the power to the n choice of sampling (2^n). The number of choices is the bits and so we could also write 2^bits:

$$resolution = RES = \frac{range}{2^n} = \frac{range}{2^{bits}}$$

<u>Examples:</u>

- $0 \div 10$: range = 10
- $-10 \div +10$: range = 20

Usually a very good range is 24 bits, but also 16 bits can be considered a good range, too. The last value of the one obtained by using this method is called LSB which stands for Least Significant Bit, for us it is the same of resolution. So, we call resolution:

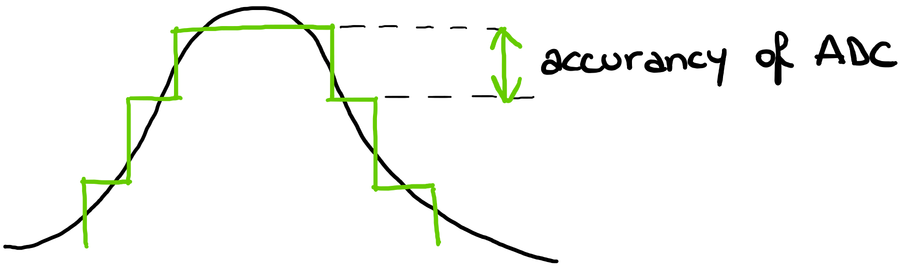
- the number of bits (informatic resolution)
- the least significant bit (informatic resolution)
- the definition we have previously provided (electrical resolution)

In general, when we'll talk about the electrical resolution, we'll refer to the definition in volt whereas when we'll say informatic resolution we'll mean the number of beats or the LSB.

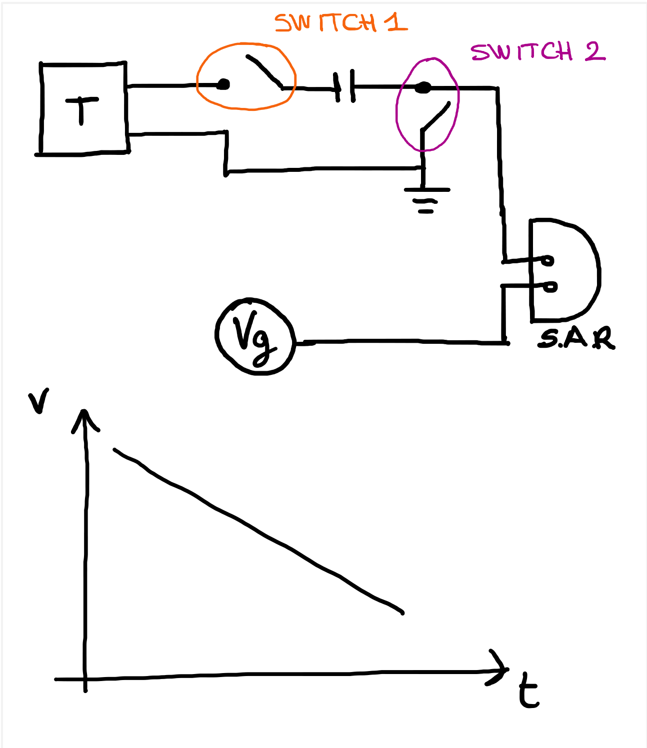
Acquisition of data (or sampling): technological issues

So, until now we have understood what sampling is; as we already seen it can be defined as the process of extracting significant data from a population. Then we saw how to transform an electrical signal to a number, and we defined this process as ADC. Actually, there are a couple of technological issues that we must cover:

1. If we should design a converter where would we invest our money on? What is most crucial component for an ACD? In term of accuracy the voltage generator is the most critical part. Of course, also the number of bits is important, but the voltage generator is the most crucial one. We can define the accuracy of the ADC as the difference between two steps as depicted in the picture:



2. The time that is required to convert the signal into a number is the second important issue that we must cover. In fact, if the conversion process is too long, in that time the signal could have been changed. So, we need to freeze the voltage and then operate a conversion. Let's now try to understand how we can freeze the voltage. In this case we can freeze the voltage by installing a capacitor which is a device used to store electric charge that consists of one pair of conductors separated by an insulator. The new system will be the one depicted in the picture; as we can see we also have two switches and the negative wire is connected to the ground. Let's see how it works; when I decide to convert the signal, I close the switch on the positive wire and keep open the witch on the negative wire; thanks to this the capacitor accumulates the charges, then I switch off the switch on the positive wire so I can convert the signal. Due to the fact that there is nothing in nature that is able to keep the charges constant, the voltage that has been stored into the capacitor starts decreasing but in a slow way; of course the speed of decreasing should be slower than the time I need to do all my comparison and conversion and so it must be lower than the LSB. In other words, I have to freeze the signal and to do this we need to charge a capacitor, at this point the voltage is not coming anymore from the transducer but it comes from the capacitor, the capacitor cannot do anything else than losing its charges. By a practical point of view, we cannot use a huge or infinite capacitor since I would spend too much time to charge and to discharge it, and this is not convenient. When the conversion is over, I have to discharge the capacitor and to do this I close the switch on the negative wire so that the capacitor is connected to the ground.



Another important point to keep in mind is that most of the time the users are not interested in seeing the measurements expressed in bits or voltage, because they only care of what they are measuring, like the temperature, for example.

Example: let's imagine having a thermometer with:

- Sensitivity: $S = 0{,}1\ V/°C$
- Range: $\pm 10\ V$
- Resolution: $12\ bits$

Let's compute:

- Electrical resolution: $Res_v = \frac{10-(-10)}{2^{12}} = 4{,}88\ mV$
- Resolution in number of levels: $Res_n = 2^{12} = 4096$
- Resolution in engineering units (it is a link between the sensitivity and the electrical resolution):

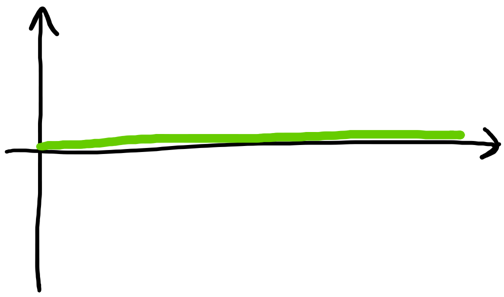
$$Res_{E.U} = \frac{Res_v}{S} = 4{,}88\ mV \cdot 0{,}001 \frac{V}{°C} = \frac{4{,}88\ mV}{100\ \frac{mV}{°C}}$$

Finally let's highlight that with 16 bits the electrical resolution is in the order of the micro volt whereas with 24 bits the electrical resolution is in the order of nano-volts. In this case we should use amplifiers because otherwise the value measured is not "valid".

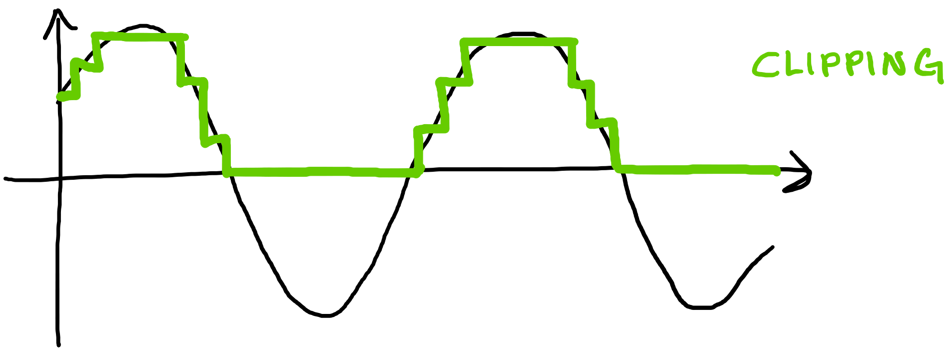
Acquisition of data (or sampling): problems

Let's now discuss a little bit about the different errors that we can end up with during the measurements if we choose the wrong range into the converter and so if we choose wrongly the values on the vertical axis (in the next paragraph we'll discuss the errors that can occur when we act wrongly on the horizontal axis, the time):

- Problem 1 = if we are measuring an amplitude lower than the LSB we'll end up having a 0 since the converter is not able to catch that signal due to the fact that the range on which we are working on is not big enough



- Problem 2: clipping = it is a form of distortion that limits a signal once it exceeds a threshold. Clipping may occur when a signal is recorded by a sensor that has constraints on the range of data it can measure. For example; if we are dealing with a harmonic signal that goes form -10 to $+10$ but we are using a range of $(0, +10)$, we'll obtain something like the figure shown since the negative signal is not converted because it out of the range of the converter itself



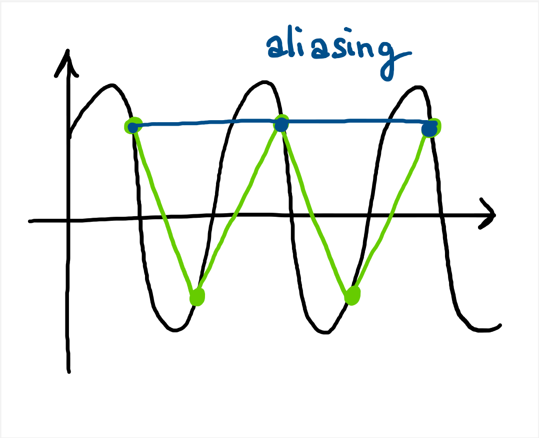
- Problem 3 = it can occur when we use the wrong range in terms of unit of measurements in fact if the range is expressed in V whereas the signal is in mV we could not be able to convert correctly the signal itself.

Time of sampling: introduction

Until now we have studied the conversion only by the amplitude point of view. Let's now discuss about time. When do we do the sampling? When do we start sampling and how often we need to do that? The time of sampling is crucial because we can introduce strong errors in our sampling. Usually errors are related to fast measurements, but they can occur also in slow measurements.
We can define the time of sampling[4] T as the product between the number of samplings N and the time interval dT or the ratio between the number of samplings N and the frequency of sampling $f_{sampling}$:

$$T = N \cdot dt = \frac{N}{f_{sampling}}$$

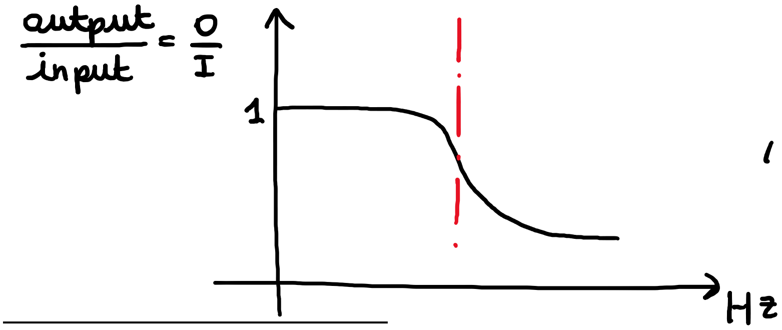
Time of sampling: aliasing

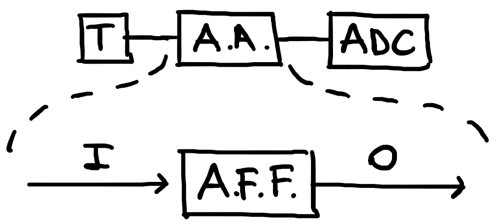


Let's now study the following plot and let's assume that the points indicated are at a distance of $1\ ms$; in this case choosing $1\ ms$ as the sampling time is not smart because we would extract the same value each sampling; a nice idea would be do measure twice per cycle. We can define the sampling frequency $f_{sampling}$ as the average number of samples obtained in one second (samples per second). The sampling frequency is important because if we choose it in a wrong way, we could reconstruct a signal that is different for the real one; this problem is called aliasing. An example of aliasing is what we can see in movies when we see a car that is moving forward whereas the wheels seems to move backwords. There is a theory that we are going to follow according to which the sampling frequency must be higher than twice the signal frequency; this theory is called Shannon theory:

$$f_{samplig} > 2f_{signal}$$

The main problem of this theory is that we could not know the frequency of the signal and so it's like a paradox because if we know something, we do not measure it, usually we measure it when we do not know the signal. So, how can we solve this paradox or problem? The only way is unfortunately to limit the damage status to our measurements; since I cannot know the signal until I have measure it, but if I do not respect this condition I do an error, the only thing I can do is removing from the signal all the things that do not depend on this relationship. So, to avoid this problem, we usually add an antialiasing device which is an apparatus that removes all the component of the signal which have a frequency higher than twice my sampling frequency. Since we use it in this specific way, we call it filter. If we study the ratio between the output and input as the function of frequency, the filter has a function that can be plotted as the following one:



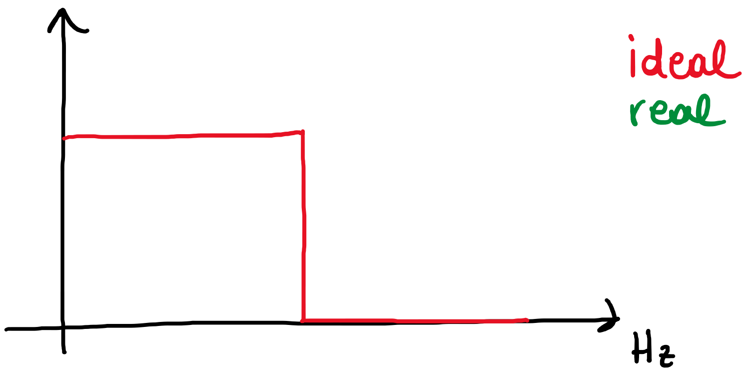


[4] Periodo di campionamento

Of course, the higher the frequency of the signal is, the lower is the ratio output/input (the idea is that I choose the sampling frequency, the red line; if the frequency of the signal are lower than this value, I keep them, otherwise if they are higher, they tend to be zero and so I neglect them). This type of filter is called low-pass filter; it can be defined as a filter that passes signals with a frequency lower than a selected cut off frequency and attenuates signals with frequencies higher than the cut off frequency. In other words, we are killing all the components that do not respect the Shannon theory. Usually I can write:

$$f_{aliasing} = f_{sampling} - f_{signal}$$

Example: let's assume having a sampling frequency at $1000\ Hz$. In this case the limit, according to Shannon theory, would be $500\ Hz$, this frequency is called Nyquist frequency or cut off frequency. I can plot it to clarify everything. In the yellow zone of the plot I have frequencies that pass that should not passed and frequencies that should have been passed but they didn't pass.



Usually the cut off happens at 60% of the ratio output/input which is more or less $\sqrt{2}/2$. But we must be sure to not have aliasing and so we take the cut off frequency and multiply it per 2,56. This value comes from the point after which the attenuation of the signal is 100dB and so very negligible. The concept is that the filter that we choose is characterised by a specific cut off frequency, we know that we are able to sampling correctly only signal characterised by the Nyquist frequency (which is equal to half of the sampling frequency), but if this frequency is higher than the cut off frequency we for sure not be able to sampling anything and so we'll choose a sampling frequency equal to the double of the cut off frequency. In addition to this to be extremely sure to sample correctly we'll actually choose a sampling frequency equal to $2{,}56 \cdot f_{cut\ off}$.
[La frequenza ci cut off è propria del filtro che si sceglie: sapendo che il filtro taglia tutto ciò che sta sopra ad essa e sapendo che si riesce a campionare correttamente dei segnali che hanno una frequenza minore della metà di quella di campionamento (cioè $f_{Nyquist} = f_{sampling}/2$), se tutto ciò che sta sopra a $f_{cut\ off}$ si perde, allora si sceglie una $f_{sampling} = 2 \cdot f_{cut\ off}$. Poi per stare sicuri, siccome il filtro non è netto, ma ha un gradino addolcito, si prende $f_{sampling} = 2{,}56 \cdot f_{cut\ off}$.]
The only question that we haven't answered yet is: at which frequency do I keep the signal? I must simply use the common sense.

Exercise: imagine having three signals: a, b and c with 8 bits and with three ranges that can be $\pm 10V$; $\pm 5V$; $\pm 25V$.

a. $3\ V\ @\ 1000\ Hz$
b. $6\ V\ @\ 2500\ Hz$
c. $0{,}01\ V\ @\ 100\ Hz$

Let's answer to these questions:

1) Which range should we choose? The range that we should choose is $\pm 10V$ since is the smallest range in which the three signals that are given are defined. Of course, we could have chosen also $\pm 25V$,but by choosing this range we would have decrease the resolution of the measurement. I cannot for sure choose the range $\pm 5V$ since one signal would not be included $(6\ V)$
2) Which is the sampling frequency? In order to answer this question, we must analyse the higher signals since it would be one easiest to be subjected to aliasing. So, we start from $f_{signal} = 2500\ Hz$. From this we could decide one value as our cut off frequency $f_{cut\ off} = 2750\ Hz$, for example (every value higher than our f_{signal} will be fine, of course we must use the common sense). After computing the cut off frequency we can determine the sampling frequency by multiplying it per 2,56: $f_{sampling} = f_{cut\ off} \cdot 2{,}56 = 7040\ Hz$.