L.N.L. - I.N.F.N. (REP) 71 / 92

Signal and Noise
Lecture notes

S. Vitale

University of California at Berkeley

A cura del servizio
Documentazione dei
L.N.L.

"Signal and Noise"

Lecture Notes

Stefano Vitale
University of California at Berkeley
Physics 250 Fall Semester 1992

Introduction

These notes discuss some basic concepts in signal processing. Through them I will refer to the conceptual picture of a physical experiment shown in fig. 1

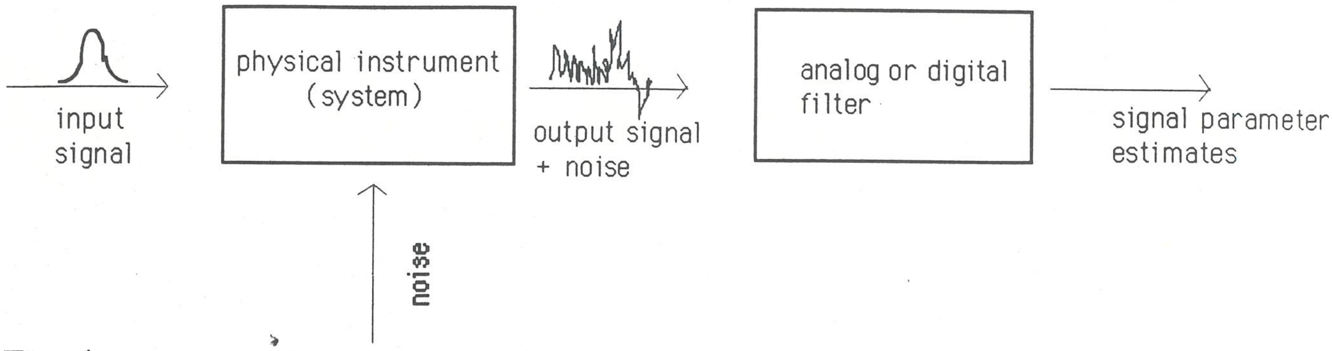

Fig. 1

The physical information one whish to extract is carried by some parameter (amplitude, power, duration etc) of the input signals. The signals are physical quantities that depend on the time or on some other physical parameter (frequency, coordinates etc.).

The physical instrumentation is thought as a black box or a "system", or as a series of systems, the purpose of which is to convert the input signals in other signals, the outputs, that can be measured and eventually converted to a set of digital data. Thus the physical apparatus act in a sense as a matematical operator that convert a real function to some other real function. The operator can be "measured" by calibrating the instrumentation and, for an ideal apparatus free of any disturbance, the input signal and its parameter would be reconstructed with infinite precision from the knowledge of the output.

Unfortunately the physical instrumentation is plagued by the noise. However skilled is the experimentalist in designing the apparatus, there are fundamental noise sources that cannot be eliminated. The disturbances that affect the physical instrumentation can be treated as random signals which add or mix to the real input and output signals and prevent the one-to-one reconstruction of the inputs from the knowledge of the output.

Nevertheless the noise has in most cases some feature that allows a partial discrimination of the signal. For instance a signal that consists of a fast pulse can be easily discriminated from a noise that consist in a slow drift of the instrumentation output. To take advantage of these differences between signal and noise they have obviously to be known and an essential part of the calibration of a physical apparatus is then the measurement of the relevant parameters of its noise sources.

The extraction of signal from the noise and the estimate of its parameters is finally obtained by some other "system", an analogic or numeric filter. The purpose of the signal processing theory is to find out what is the filter design that allows, for given properties of signal and noise, to estimate the signal parameters with the smallest possible uncertainty.

These notes will proceed through the steps of the picture above: signals, systems, stochastic processes, fundamental noise processes, measurement of noise parameters, filtering and signal detection.

1 The Signal

I will call a signal any physical (measurable) quantity s(t) which is a function of another physical parameter indicated here with t. I will assume, unless differently specified, that the uncertainty with which t is measured is neglegible, that is that the error on s due to that uncertainty, $\delta s \approx \left|\frac{ds}{dt}\right| \delta t$, is neglegible in comparison to the overall uncertainty on s.

For sake of clarity I will mainly focus on the case where t is the time (time signals), however this assumption is not necessary and t could be any other parameter like, for instance, a space coordinate or a frequency.

I will not deal with signals depending on more than one parameter. However the main results that will be obtained for the single parameter case can often be extended, with no very much effort, to many parameters signals.

Fourier Transform

Time signals availables at the output of physical instrumentation have limited duration and take finite values. Thus they obey the Dirichelet condition:

$$\int_{-\infty}^{\infty} |s(t)|\,dt < \infty \qquad 1$$

and can be Fourier transformed:

$$s(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} s(\omega)e^{i\omega t}\,d\omega \qquad 2$$

with

$$s(\omega) = \int_{-\infty}^{\infty} s(t)e^{-i\omega t}\,dt \qquad 3$$

The Fourier transform converts the time signal s(t) to the angular frequency signal s(ω) which plays a crucial role in signal processing.

s(ω) is a complex number and consists of two real signals. As s(t) is a real function however, then s(ω)=s*(-ω). The transform maps then one real function given on the whole time axis to the two real functions Re{s(ω)} and Im{s(ω)} that one needs to know only on one of the two angular frequency semiaxis.

Elementary properties of Fourier transforms that will be used in the next are shown in table 1.

Table1

Elementary properties of Fourier transforms and selected exemples

	Function	Transform
1	$s(t)$	$s(\omega)$
2	$a\cdot s(t)+b\cdot q(t)$	$a\cdot s(\omega)+b\cdot q(\omega)$
3a	$s(t)\cdot q(t)$	$\frac{1}{2\pi}\int_{-\infty}^{\infty} s(\omega')q(\omega-\omega')d\omega'$
3b	$\int_{-\infty}^{\infty} s(t')q(t-t')dt'$	$s(\omega)q(\omega)$
4	$\frac{ds}{dt}$	$i\omega s(\omega)$
5	$s(t)=s(-t)$	$Im\{s(\omega)\}=0$
6	$s(t)=-s(-t)$	$Re\{s(\omega)\}=0$
7	$\int_{-\infty}^{\infty} s(t)q(t)dt=\frac{1}{2\pi}\int_{-\infty}^{\infty} s(\omega)q^*(\omega)d\omega$	
8	$\int_{-\infty}^{\infty} \lvert s(t)\rvert^2 dt=\frac{1}{2\pi}\int_{-\infty}^{\infty} \lvert s(\omega)\rvert^2 d\omega$	
9	$e^{-\frac{t}{\tau}}$ $(t\geq 0)$, 0 $(t<0)$	$\frac{1}{\frac{1}{\tau}+i\omega}$
10	$\frac{1}{\omega_1}e^{-\frac{t}{2\tau}}\sin(\omega_1 t)$	$\frac{1}{\omega_0^2-\omega^2+\frac{i\omega}{\tau}}$ $(\omega_0^2=\omega_1^2+\frac{1}{4\tau^2})$

11	$1\ (-\frac{T}{2} \le t \le \frac{T}{2})$; 0 elsewhere	$T\frac{\sin\left(\frac{\omega T}{2}\right)}{\frac{\omega T}{2}}$
12	$\frac{e^{-\frac{t}{2\tau^2}}}{\sqrt{2\pi\tau^2}}$	$e^{-\frac{\omega\tau^2}{2}}$
13	$\delta(t)$	1

Uncertainty Principle

The Fourier transform preserve the signal energy, an important signal parameter defined by

$$E= \int_{-\infty}^{\infty} |s(t)|^2 dt. \qquad 4$$

In fact, from line 8 of table 1 one can see that, besides the factor 2π, the energies of s(t) and of s(ω) are the same.

Despite this, the distribution of the energy for the two signals is quite different. In fact the uncertainty principle that we discuss here shows that the more the energy of the time signal is concentrated the more the frequency signal is broadned.

To discuss this priciple let first give a definition of a width of a signal. To do that one first converts a signal to a normalized energy density f(t):

$$f(t)=\frac{|s(t)|^2}{E} \qquad 5$$

with this density one can define a mean signal time $\bar{t}= \int_{-\infty}^{\infty} tf(t)dt$ and a time width $\delta t=\sqrt{\int_{-\infty}^{\infty}(t'-\bar{t})^2 f(t')dt'}$. Notice that by an appropriate shift of the time origin one can always assume $\bar{t}=0$ without changing δt.

As the modulus of s(ω) is an even function of ω, the mean signal angular frequency is always zero and the signal angular frequency width is

$$\delta\omega=\sqrt{\int_{-\infty}^{\infty}\omega'^2 f(\omega')d\omega'}.$$

With these definition of δt and of δω one can then evaluate a lower limit to the product δtδω from

$$(\delta t\delta\omega)^2=\frac{\int_{-\infty}^{\infty}t'^2|s(t')|^2dt'\int_{-\infty}^{\infty}\omega'^2|s(\omega')|^2d\omega'}{2\pi E^2}=$$

$$=\frac{\int_{-\infty}^{\infty}t'^2|s(t')|^2dt'\int_{-\infty}^{\infty}\left|\frac{ds}{dt'}\right|^2dt'}{E^2}\geq\frac{\left|\int_{-\infty}^{\infty}t's(t')\frac{ds}{dt'}dt'\right|^2}{E^2} \qquad 6$$

Here I have used line 4 of table 1. E in eq. 6 is the energy of the time signal.

The integral in the right hand side of eq. 6 can be performed by part

$$\int_{-\infty}^{\infty}t's(t')\frac{ds}{dt'}dt'=\frac{1}{2}[(t's^2(t'))_{t'=\infty}-(t's^2(t'))_{t'=-\infty}]-\frac{1}{2}\int_{-\infty}^{\infty}s^2(t')dt'=$$

$$=-\frac{1}{2}E \qquad 7$$

where I have used the observation that $t's^2(t')$ has to vanish at $t'=\pm\infty$

Substituting in eq. 6 one gets

$$(\delta t\delta\omega)^2\geq\frac{1}{4} \qquad 8$$

which is in fact the uncertainty principle. In the following table 2 I list energies,time width and frequency width for some of the function listed in table 1

Table 2

	Function	Transform	E	δt	δω
1	$e^{-\frac{t}{\tau}}$ (t≥0), 0 (t<0)	$\frac{1}{\frac{1}{\tau} + i\omega}$	$\frac{\tau}{2}$	$\frac{\tau}{2}$	∞
2	$\frac{1}{\omega_1} e^{-\frac{t}{2\tau}} \sin(\omega_1 t)$	$\frac{1}{\omega_o^2-\omega^2 + \frac{i\omega}{\tau}}$ $(\omega_o^2 = \omega_1^2+\frac{1}{4\tau^2})$	$\frac{\tau}{\omega_o^2}$	τ $(\omega_1\tau>>1)$ 3τ $(\omega_1\tau<<1)$	ω_o
3	$1\ (-\frac{T}{2} \le t \le \frac{T}{2})$; 0 elsewhere	$T\frac{\sin\left(\frac{\omega T}{2}\right)}{\frac{\omega T}{2}}$	T	$\frac{T}{\sqrt{12}}$	∞
4	$\frac{e^{-\frac{t}{2\tau^2}}}{\sqrt{2\pi\tau^2}}$	$e^{-\frac{\omega\tau^2}{2}}$	1	$\frac{\tau}{\sqrt{2}}$	$\frac{1}{\tau\sqrt{2}}$

Notice that only for the gaussian function in the fourth row the limit set by the uncertainty principle is reached while for the other ones the product $\delta t \delta\omega$ is much larger than 1/2 or even infinite. This is in part due to the fact that the width of the signal along the angular frequency axis is related to the integral of the square modulus of its derivative (eq. 11). Thus any discontinuity in s(t), contributing by a large amount to the integral, strongly broadens the spectrum. It can be seen in fact that the transform of a unit step $\theta(t)$, that can be obtained as the limit of the exponential relaxation in line 1 of table 2 for $\tau \to \infty$, is $\frac{1}{i\omega}$ and thus the associated energy density decays only as $\frac{1}{\omega^2}$ when $|\omega| \to \infty$ so that the integral entering in the definition of $\delta\omega$ becomes infinite.

Narrowband Signals

A consequence of the uncertainty principle is that a physical signal, which has always a finite duration, cannot have infinitely narrow lines in its Fourier transform. A narrow line in $s(\omega)$ around some angular frequency ω_o, translates, in the time domain, to an approximately sinusoidal signal of angular frequency ω_o. thus, to clarify this issue, let define a narrow band signal has a signal that can be represented by:

$$s(t)=a(t)\sin(\omega_1 t)+b(t)\cos(\omega_1 t) \qquad 9$$

where a(t) and b(t) are slowly varying function of time. With the term slowly varying I mean that the Fourier transform of a(t) and b(t) approach zero when $\omega > \omega_c$ with ω_c an angular frequency such that $\omega_c << \omega_1$. Usually a(t) and b(t) are called the quadrature components of the signal.

According to line 3 of table 1 the Fourier transform of the signal is the convolution of the transforms of a(t) and b(t) with the transforms of $\sin(\omega_1 t)$ and $\cos(\omega_1 t)$. These last ones are not defined in a strict sense but can be obtained using the Dirac delta. In fact

$$\sin(\omega_1 t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\pi}{i}[\delta(\omega-\omega_1)-\delta(\omega+\omega_1)]e^{i\omega t}d\omega \qquad 10a$$

$$\cos(\omega_1 t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\pi[\delta(\omega-\omega_1)+\delta(\omega+\omega_1)]e^{i\omega t}d\omega \qquad 10b$$

Performing the convolution one gets that

$$s(\omega)=\frac{a(\omega-\omega_1)-a(\omega+\omega_1)}{2i}+\frac{b(\omega-\omega_1)+b(\omega+\omega_1)}{2} \qquad 11$$

As $a(\omega)$ and $b(\omega)$ vanish for $\omega>\omega_c$, $s(\omega)$ will then be different from zero only in the two angular frequency intervals $\omega=\pm\omega_1\pm\omega_c$ and will consist of two lines of width ω_c around $\omega=\omega_1$. If s(t) have a finite duration T, then both a(t) and b(t) have finite duration and their time width is $\delta t\approx T$. Thus $\omega_c>\frac{1}{2T}$. If, for instance, s(t) is a gaussian packet resulting from the product of the function n.4 in table 1 and a pure sinusoidal term, then $s(\omega)$ will consists of two lines of width $\frac{1}{\sqrt{2}\delta t}$ around ω_1.

Exercises

1) Find the Fourier transform of a triangular signal $s(t)=1+\frac{2t}{T}$ for $-\frac{T}{2}\le t\le 0$, $s(t)=1-\frac{2}{T}t$ for $0\le t\le\frac{T}{2}$. Find also the signal energy, the time width and the energy width.

2) A sinusoidal signal $s(t)=\sin\omega t$ is switched on at time 0 and switched off at time T. What is its Fourier transform?. Give an estimate the time and frequency width.

3) A pulsed laser beam consists of a train of 100 gaussian bursts separated by a time T=1 ms. The width of the single burst is τ=1 μs. Find the Fourier transform of the beam intensity.

Sampling and Band Limited Signals.

In the preceding sections I considered signals which are continous functions of the parameter t. In numerical processing of signals however one deals only with a numerable set of data s_n that results from the sampling of a continous signal at fixed time values. Usually the sampling occurs at a fixed frequency ν_s so that $s_n=s(nT)$ with $T=\frac{1}{\nu_s}$. I will call this set of data a discrete signal.

In this section I discuss some properties of discrete signals and I give an estimation of the loss of information related to the sampling of a continous signal.

The Sampling Theorem

Let consider a signal s(t) whose Fourier transform is $s(\omega)$ and the sequence of its samples at fixed time interval $s_n=s(nT)$. Let form the signal

$$s'(t)=\sum_{n=-\infty}^{\infty}\frac{s_n\cdot\sin\left[\frac{\pi}{T}(t-nT)\right]}{\frac{\pi}{T}(t-nT)} \qquad 12$$

One can verify that s(nT)=s'(nT) so that the two signals have the same samples at t=nT. Now I want to show that for a certain class of signals the two coincide for any value of t.

The Fourier transform of s'(t) is

$$s'(\omega)=\begin{cases} T\sum_{n=-\infty}^{\infty} s_n \cdot e^{-i\omega nT} & |\omega|\leq\frac{\pi}{T} \\ 0 & |\omega|>\frac{\pi}{T} \end{cases} \qquad 13$$

Recalling that $s_n=\frac{1}{2\pi}\int_{-\infty}^{\infty} s(\omega)e^{i\omega nT}d\omega$ one gets that, for $|\omega|\leq\frac{\pi}{T}$, eq. 13 becomes

$$s'(\omega)=\frac{T}{2\pi}\int_{-\infty}^{\infty} d\omega' s(\omega')\sum_{n=-\infty}^{\infty} \cdot e^{i(\omega'-\omega)nT} \qquad 14$$

Now it can be shown[1] that

$$\sum_{n=-\infty}^{\infty} \cdot e^{i(\omega'-\omega)nT}=\frac{2\pi}{T}\sum_{n=-\infty}^{\infty} \delta(\omega'-\omega+n\frac{2\pi}{T}) \qquad 15$$

so that

$$s'(\omega)=\sum_{n=-\infty}^{\infty} s(\omega+n\frac{2\pi}{T}) \qquad 16$$

---

[1]Consider the periodic signal $f(\omega)=\Sigma_n\delta(\omega+n\omega_o)$. It can be expanded in a Fourier series with complex coefficients $c_k=\frac{1}{\omega_o}\int_{-\omega_o/2}^{\omega_o/2} f(\omega)e^{-ik2\pi\omega/\omega_o}d\omega=\frac{1}{\omega_o}$. Thus $f(\omega)=\frac{1}{\omega_o}\Sigma_\kappa e^{ik2\pi\omega/\omega_o}$

In order that the two transforms, and thus the two signal, coincide, one needs $s(\omega)=0$ for $|\omega|\geq\frac{\pi}{T}$ then $s'(\omega)=s(\omega)$ and

$$s(t)=s'(t)=\sum_{n=-\infty}^{\infty}\frac{s_n\cdot\sin\left[\frac{\pi}{T}(t-nT)\right]}{\frac{\pi}{T}(t-nT)} \qquad 17$$

Thus, summarizing, if a signal is bandlimited, i.e. if it has a Fourier transform which is zero above a certain angular frequency ω_{max}, than the knowledge of its samples taken at a frequency $\nu_s\geq 2\nu_{max}=\frac{\omega_{max}}{\pi}$ is sufficient to reconstruct the signal at any time. $2\nu_{max}$ is often called the Nyquist frequency of the signal.

If the signal is sampled at a frequency less than the Nyquist one an error is made called the aliasing error. It is possible[2] to estimate an upper bound to the aliasing error :

$$|s(t)-s(t')|\leq\frac{1}{\pi}\int_{\pi\nu_s}^{\infty}|s(\omega)|d\omega \qquad 18$$

It is also useful to estimate what is the error one makes taking a truncated reconstruction of the signal

---

[2] $|s(t)-s'(t)|\leq\frac{1}{2\pi}\int_{-\infty}^{\infty}|s(\omega)-\Sigma_n s'(\omega+n\omega_{max})|d\omega\leq$

$$\leq\frac{1}{2\pi}\left[\int_{|\omega|>\omega_{max}}|s(\omega)|d\omega+\int_{|\omega|\leq\omega_{max}}\Sigma_{n\neq 0}|s(\omega+n\omega_{max})|d\omega\right]=$$

$$\frac{1}{2\pi}\left[\int_{|\omega|>\omega_{max}}|s(\omega)|d\omega+\int_{-\infty}^{\infty}|s(\omega)|d\omega-\int_{|\omega|<\omega_{max}}|s(\omega)|d\omega\right]=\frac{1}{\pi}\int_{|\omega|>\omega_{max}}|s(\omega)|d\omega$$

$$s''(t)=\sum_{n=-N}^{N}\frac{s_n\cdot\sin\left[\frac{\pi}{T}(t-nT)\right]}{\frac{\pi}{T}(t-nT)} \qquad 19$$

the r.m.s. error can be evaluated[3] and is given by:

$$\int_{-\infty}^{\infty}|s(t)-s''(t)|dt=T\Sigma_{|n|>N}|s(nT)|^2 \qquad 20$$

Discrete data are aquired, processed and stored in digital form. This means that they consist in a sequence of numbers known with a finite number of significant digits or bits[4]. If the aliasing error or the truncation error become smaller than the resolution, they become neglegible and the representation of s(t) by s'(t) or by s"(t) become accurate within the resolution.

Exemple: The gaussian signal in line 4 of table 1 is aquired with an 8 bits a/d converter (relative resolution $\approx\pm\frac{1}{500}$) adjusted so that the full scale corresponds to the peak amplitude of the signal. The aliasing error is less then $\frac{2}{\pi}\int_{\pi\nu_s}^{\infty}e^{-\frac{\omega\tau^2}{2}}d\omega=\sqrt{\frac{8}{\pi\tau}}\mathrm{Erf}(\pi\nu_s\tau)$, with

---

[3] $\delta s(t)=s(t)-s''(t)=\Sigma_{|n|>N}\frac{s_n\cdot\sin\left[\frac{\pi}{T}(t-nT)\right]}{\frac{\pi}{T}(t-nT)}$;

$$\int_{-\infty}^{\infty}|\delta s(t)|^2dt=\Sigma_{|n|,|m|>N}\, s_n s_m\int_{-\infty}^{\infty}\frac{\sin\left[\frac{\pi}{T}(t-nT)\right]\sin\left[\frac{\pi}{T}(t-mT)\right]}{\frac{\pi}{T}(t-nT)\quad\frac{\pi}{T}(t-mT)}dt)=T\Sigma_{|n|>N}|s_n|^2.$$

[4] Analog to digital converters can be found that convert data at 10-100 KHz sampling frequency with a resolution in the range of 15-25 bits or with sampling frequency up to the GHz with resolutions still of 6-8 bits. With N bit a number up to 2^{N-1} can be represented. This gives a relative resolution of $\pm\frac{\frac{1}{2}}{2^{N-1}}\approx\frac{1}{2N+1}$

Erf the error function, or $2\mathrm{Erf}(\pi\nu_s\tau)$ of the full scale. In order that this error is less than the resolution $\nu_s \approx \frac{3}{\pi\tau}$. The truncation error can be estimated by substituting the sum in eq. 20 by an integral and evaluating the mean error as $\sigma_t \approx \sqrt{\frac{1}{2NT}\int_{-NT}^{NT}|s(t)-s''(t)|^2 dt} \leq$
$\leq \sqrt{\frac{\mathrm{Erf}(\frac{\sqrt{2NT}}{\tau})}{2\sqrt{2\pi}.NT\tau}}$. With $T \approx \tau$ this gives $N \approx 3$. Thus the signal is represented just by its 7 center samples taken every τ sec.

The Discrete Fourier Transform

As already state the signal $s(\omega)$ plays a fundamental role in signal processing techniques. It is of great relevance then to have a practical method for its evaluation from the corresonding time signal s(t). Here we discuss a numeric method which is by far the most widely used in practical experimental chains at least for signal with frequencies below the GHz range.

As we have seen before, normally any continous signal can be reduced[5] to a set of N numbers s_n with $0 \leq n \leq N-1$. Of a sequence of N numbers one can define a transform, called the discrete Fourier transform as:

$$s(k) = \sum_{n=0}^{N-1} s_n \cdot e^{-i\frac{2\pi}{N}nk} \qquad 21$$

with an inversion formula given by

[5] It is worth to notice that strictly speaking a signal of finite duration cannot be band limited and thus cannot be sampled with arbitrary accuracy. The most serious problem with respect to that rises when the signal results from the actual truncation, at t=0 and $t=T_{max}$, of a signal which is not zero at those two points. It is possible to show that if a signal is differnt from zero only between t=0 and $t=T_{max}$ and if its derivatives up that of order n have limited values in this same interval, than $s(\omega) \rightarrow \omega^{n+1}$. Thus in order that the signal has small high frequency tails one needs that he has an high level of "tangency" at both t=0 and $t=T_{max}$. Discontinuities at the ends will cause slowly damped oscillations of $s(\omega)$ at infinity that will contribute to the integral in eq. 18. The practical way to reduce this truncation phenomenon is to multiply the data s_n by a suitably chosen "window" w_n that goes to zero smoothly at the ends of the measuring interval. A popular window is, for instance, the Hanning one defined by $w_n = \frac{1}{2}\left[1-\cos\left(\frac{2\pi n}{N-1}\right)\right]$

$$s_n=\frac{1}{N}\sum_{k=0}^{N-1} s(k)\cdot e^{i\frac{2\pi}{N}nk} \qquad 22$$

The sequence s(k) is periodic of period N so that one needs to know only the terms with $0 \le k \le N-1$. In addition $s(k)=s^*(N-k)$ and s(0) is real so that the transform contains only N independent real numbers as the signal s_n.

There exist very efficient algorithms, known as Fast Fourier Transforms (FFT), to calculate the discrete Fourier transform of a sequence. It is beyond the scopes of these notes to illustrate these algorithms, we whish however point out that, because the transform convert N numbers to N numbers, in principle the memory occupation needed for the calculation is just of two N components vectors, one for the input data and one for the output. It turns out, in addition, that the number of elementary operations, sum and product, needed to calculate a transform is proportional approximately to NlogN and not to N^2 as one would guess from the definition in eq 21. These two features, small memory occupation and small number of operation required, is the base of the efficiency of the algorithm.

From the coefficients s(k) of the FFT the spectrum of the truncated signal s"(t) in eq. 24 can be reconstructed. In fact substituting in eq. 19 s_n taken from eq. 22, and taking the Fourier transform of the resulting expression one gets:

$$s''(\omega)=\frac{T}{N}\sum_{k=0}^{N-1} s(k)\sum_{n=0}^{N-1} \cdot e^{i\frac{2\pi}{N}nk-\omega nT}=$$

$$23$$

$$=\frac{T}{N}\sum_{k=0}^{N-1} s(k)\frac{1-e^{iN(\frac{2\pi}{N}k-\omega T)}}{1-e^{i(\frac{2\pi}{N}k-\omega T)}} \qquad (|\omega|\le\frac{\pi}{T})$$

Notice that $s(k)=\frac{s"(\omega=k\frac{2\pi}{NT})}{T}$ so that the coefficients s(k) are the samples of the Fourier transform of the continous (truncated) signal taken at integer multiples of the angular frequency $\frac{2\pi}{NT}$ i.e. at integer multiples of a frequency which is the inverse of the total duration NT of the truncated signal.

The term $\frac{1-e^{iN\phi}}{1-e^{i\phi}}$ is an oscillating function of ϕ with period 2π. It has a large lobe at $\phi=0$ and a first zero at $\phi=\frac{2\pi}{N}$. It the bears close similarity to the function $\frac{\sin(x)}{x}$ that enters in the expansion of the time signal in eq. 17. Eq. 23 represent thus in a sense a frequency domain analogue of the sampling expansion of the time signal.

Systems

As already stated a physical instrument will be treated here as a system that performs some mathematical operation on one or more signals, the inputs, converting them to other signals, the outputs. Focusing on the case of one input i(t) and one output o(t), the action of a system on the input can be represented then as a functional operator $o=T\{i\}$. The full knowledge about the operator T can be obtained only through an experimental calibration procedure. For a generic operator this procedure can be extremely complicate as it involves measuring the output at any time for any possible value of the input at any other time. A significant semplification is obtained in two special but very important cases.

The first case of extraordinary importance is that of linear systems. A linear system is a system for which a principle of superposition holds:

$$T\{a_1 i_1+a_2 i_2\}=a_1 T\{i_1\}+a_2 T\{i_2\} \qquad 24$$

As any signal can expanded as a linear combination of some other properly chosed signals (orthonormal functions, Dirac delta etc) the calibration of the instrument reduces to measuring the output due to these signals only.

Linear systems are not just a mathematical curiosity. The response of the major part of the physical systems can often be approximated by a linear response if the input and the output signals have small enough variations $\delta i(t)$ and $\delta o(t)$ about some given function of time $i_o(t)$ and $o_o(t)$. In this small signal linear approximation the role of inputs and outputs is taken by the signal variations $\delta i(t)$ and $\delta o(t)$ while $i_o(t)$ and $o_o(t)$ enter in setting the system response.

The second important case is when the output at a given time t is a function of the input at the same time t

$$o(t)=f[i(t)] \qquad 25$$

where f(x) is an ordinary function of real variable. Such a system is usually called a system without memory and its calibration needs only the measurement of the function f(x). A system without memory can only approximate within a certain accuracy the behaviour of a physical instrument as an arbitrary fast response would violate at least special relativity. However one can always think to a real system as composed by an ideal system with no memory followed by a proper delay element.

In this section we will discuss some properties of linear systems and we will also consider one non linear systemwithout memory.

Linear Systems

Let first discuss linear systems that have only one input and one output. A generic input signal i(t) can be written as

$$i(t)=\int_{-\infty}^{\infty} i(t')\delta(t-t')dt'=T\lim_{T\to 0}\sum_{k=-\infty}^{\infty} i(kT)\delta(t-kT) \qquad 26$$

it consists then of a linear conbination of an infinite sequences of Dirac deltas shifted in time, with coefficients given by the input signals values.

Let call now h(t,t') the output of the system for an input consisting of a delta[1] centered at time t', δ(t-t'). Due to the principle of superposition in eq. 24 the output for the signal in eq. 26 will be:

$$o(t)=T\lim_{T\to 0}\sum_{k=-\infty}^{\infty} i(kT)h(t,kT)=\int_{-\infty}^{\infty} i(t')h(t,t')dt' \qquad 27$$

and the full knowledge of the system behaviour is obtained if h(t,t') is measured.

Time Invariant Systems

A system is called time invariant if for any i(t) and o(t), with o(t)=T{i(t)}, then o(t+τ)=T{i(t+τ)}. For a linear time invariant system h(t,t')=h(t-t') and is called the impulse response of the system. Eq. 27 becomes[2]:

---

[1] Not every linear system would respond with an ordinary function to a delta input. However for almost any system we will treat here, h(t,t') can be defined at least as a distribution.

[2] Actually, in order h(t,t')=h(t-t'), for many systems o(t) should be written as the sum of two pieces one $o_o(t)$ coming from the free evolution of the system and depending only on the internal status of the system itself or from some initial condition, and the other coming from the effect of the input. The system being linear, this two pieces just add up according to :

$$o(t)=o_o(t)+\int_{-\infty}^{\infty} h(t')i(t-t')dt'$$

$$o(t)= \int_{-\infty}^{\infty} h(t')i(t-t')dt' \qquad 28$$

The response of a linear, time invariant system to any signal can then been calculated simply knowing the function h(t). To measure this function in practice one has to apply to the system a pulse shorter than any relaxation time of the system itself and with an amplitude not too big to prevent the system to loose its linearity. Such a mesurement can be often difficult or even impossible in practice. An alternative way is to apply to the system a unit step signal θ(t) and measure the corresponding output signal $h_-(t)$. h(t) can then be derived from

$$\frac{dh_-}{dt} = \frac{d}{dt} \int_{-\infty}^{\infty} h(t')\theta(t-t')dt' = h(t) \qquad 29$$

. Eq. 29 has a very important consequence. If i(t),o(t) and h(t) admit a Fourier transform, then from the convolution theorem in line 3 of table 1 it follows:

$$o(\omega)=i(\omega)h(\omega) \qquad 30$$

which is the most inportant result of the linear response theory. The function h(ω) is called the frequency response of the system.

Systems in Series

Let consider a cascade of N systems in series, i.e. with the output $o_n(t)$ of the n^{th} system being the input $i_{n+1}(t)$ of the $(n+1)^{th}$ one, eq. 30 can be iterated to give

$$o_N(\omega)= \prod_{k=1}^{N} h_n(\omega) i_1(\omega) \qquad 31.$$

We will consider this issue in some detail when we will deal with systems described by ordinary differential equation at constant coefficients

Eq. 31, that involves an ordinary product of function, states that the order in which the systems are located within the cascade is irrelevant.

Stability

In order that the Fourier transform of the inpulse response exists h(t) has to obey the condition in Eq. 1. If this is the case it can be demonstrated[3] that the system output o(t) is bounded, $|o(t)| < \infty$, if the input is bounded, $|i(t)| < \infty$. A system with this property is called inconditionally stable. Thus only inconditionally stable systems have a frequency response.

Causal Systems

A real physical system can not violate the principle of causality. This means that the output at a given time t can only depend on the value of the input at times t'≤ t. Such system is called causal. Eq. 28 becomes in this case:

$$o(t)=\int_0^{\infty} h(t')i(t-t')dt' \qquad 32$$

This observation has some very important consequences. The first is that if the system is stable, so that h(ω) exists, than the real and the imaginary parts of this complex function are no longer independent but are related by a couple of integral equations called Kramers-Krönig dispersion relations or simply dispersion relations. To illustrate this relations consider that, because for causal systems h(t) is zero if t<0, then its Fourier transform is

$$h(\omega)=\int_0^{\infty} h(t)e^{-i\omega t} \qquad 33$$

---

[3] If $|i(t)|<\infty$ then $\int_{-\infty}^{\infty} |h(t')i(t-t')|dt' \leq \max(|i(t)|) \int_{-\infty}^{\infty} |h(t')|dt' < \infty$. On the other and if for any i(t) such that $|i(t)|<\infty$, $|o(t)|<\infty$ then applying the input $\frac{h(-t)}{|h(-t)|}$, which is bound, one gets the output $\int_{-\infty}^{\infty} |h(t)|dt$ that then is limited.

This has an important consequence: $h(\omega)$ remains a well defined function of ω even when this parameters is taken as a complex number $\omega=\omega'+i\omega''$ provided $\omega''\leq 0$. With this condition $e^{-i\omega t}$ in eq. 33 becomes $e^{-i\omega' t}e^{-|\omega''|t}$ and the integral goes to zero if $\omega''\to -\infty$. One can then evaluate the integral

$$\int_C \frac{h(\underline{\omega})}{\underline{\omega}-\omega}d\underline{\omega} \qquad 34$$

on a closed contour all contained in the lower half-plane (Fig 2) and get

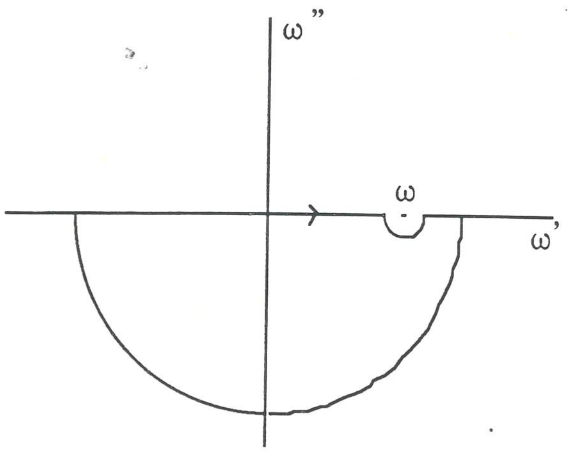

Fig 2. Integration contour to perform the integral in eq. 34.

$$P\int_{-\infty}^{\infty} \frac{h(\underline{\omega})}{\underline{\omega}-\omega}d\underline{\omega} + i\pi h(\omega)=0 \qquad 35$$

with P designating the principal value. Eq. 35 can be recasted as

$$Re\{h(\omega)\}=-\frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{Im\{h(\underline{\omega})\}}{\underline{\omega}-\omega}d\underline{\omega} \qquad 36a$$

and

$$\mathrm{Im}\{h(\omega)\} = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{\mathrm{Re}\{h(\underline{\omega})\}}{\underline{\omega}-\omega} d\underline{\omega} \qquad 36b$$

which are the dispersion relations.

In order to know $h(\omega)$ one then needs only to know either its real part or as well its imaginary one, the other being calculable from eq. 9a or 9b. Notice that the dispersion rules are true for whatever function obeys the condition $s(t)=0$ for $t<0$. Thus also for any signal that begins at $t=0$ one needs to know only either $\mathrm{Re}\{s(\omega)\}$ or $\mathrm{Im}\{s(\omega)\}$ between $0 \le \omega \le \infty$, to calculate $s(\omega)$ for every value of ω.

Exemple 2.1. A real capacitor is often modeled as an ideal one with a loss resistor in parallel. The current to voltage relation of the device is given by $V(\omega)=\frac{RI(\omega)}{1+i\omega RC}$. If the current is considered as the input then the frequency response is causal (verify) and obey eq. 36a and 36b. On the contrary if the role of V and I is reversed the system is not causal any more.

Non stable, causal systems and Laplace transforms.

If the system is causal but not stable then a Laplace transform of $h(t)$ can be defined as

$$h(p) = \int_{0}^{\infty} h(t) e^{-pt} dt \qquad 37$$

where p is a complex number. $h(p)$ is the Fourier transform of $h'(t)=h(t)e^{-p't}$ where p' is the real part of p. $h(p)$ exists provided that it exists a positive nunmber p_o such that, for $p' \ge p_o$, $h'(t)$ fullfills the condition in eq. 1 and can be Fourier transformed. $h(p)$ is called the transfer function of the system. If the system is stable then obviously p_o can be taken as $p_o=0$ and the transfer function for $p=i\omega$ coincides with the frequency response.

The convolution theorem (Table1 line 3) holds[4] for Laplace transforms too provided both function in the convolution are zero for negative times. Thus for an input such that i(t)=0 for t<0 one gets

$$o(p)=h(p)i(p) \qquad 38$$

A summary of Laplace transform properties can be found in textbooks. Here we remind only some of them we will use in the next:

$$\int_0^\infty \frac{dh}{dt} e^{-pt} dt = ph(p) - h(t=0^+) \qquad 39a$$

$$\lim_{p \to 0} ph(p) = \lim_{t \to \infty} h(t) \qquad 39b$$

$$\lim_{p \to \infty} ph(p) = \lim_{t \to 0} h(t) \qquad 39c$$

Example 2.2. The SQUID

For what we have to discuss here, a SQUID is a non linear system the input of which is the magnetic flux Φ threading the coil in Fig. 3a and the output of which is the voltage V in the same figure. The V-Φ characteristics for slow enough signals (fig. 3b) is an almost triangular periodic pattern with amplitude ΔV that depends on the amplifier gain and with period $\Delta\Phi=\phi_0\approx 2\cdot 10^{-15}$ Wb. For small signals $\delta\Phi$ around the point marked Φ_1 in Fig. 3b the system can be treated as a linear one with a voltage to flux gain of $G=2\Delta V/\phi_0$. The time delays introduced by the whole electronic chain can be represented, in this small signal approximation, assuming that the system has a single relaxation constant τ. The overall small signal frequency response is then

$$h(\omega)=\frac{G}{1+i\omega\tau} \qquad 40$$

To prevent the system to run out of its linearity range, the SQUID is often used in a feedback configuration (Fig. 3c) where the output voltage V is sent, through a suitable feedback resistor R to a second coil inductively coupled to the SQUID with mutual

[4] The proof can be found in any book mathematical methods like for instance: I. S. Sokolnikoff and R. M. Redheffer "mathematics of physics and modern engineering" Mc Graw-Hill

induction M. The total flux through the SQUID is then $\Phi_{tot}=\Phi-\frac{VM}{R}=\Phi-\beta V$, with $\beta=M/R$. Here we have assumed that M>0 and that the direction of the windings give the minus sign in front of the feedback term.

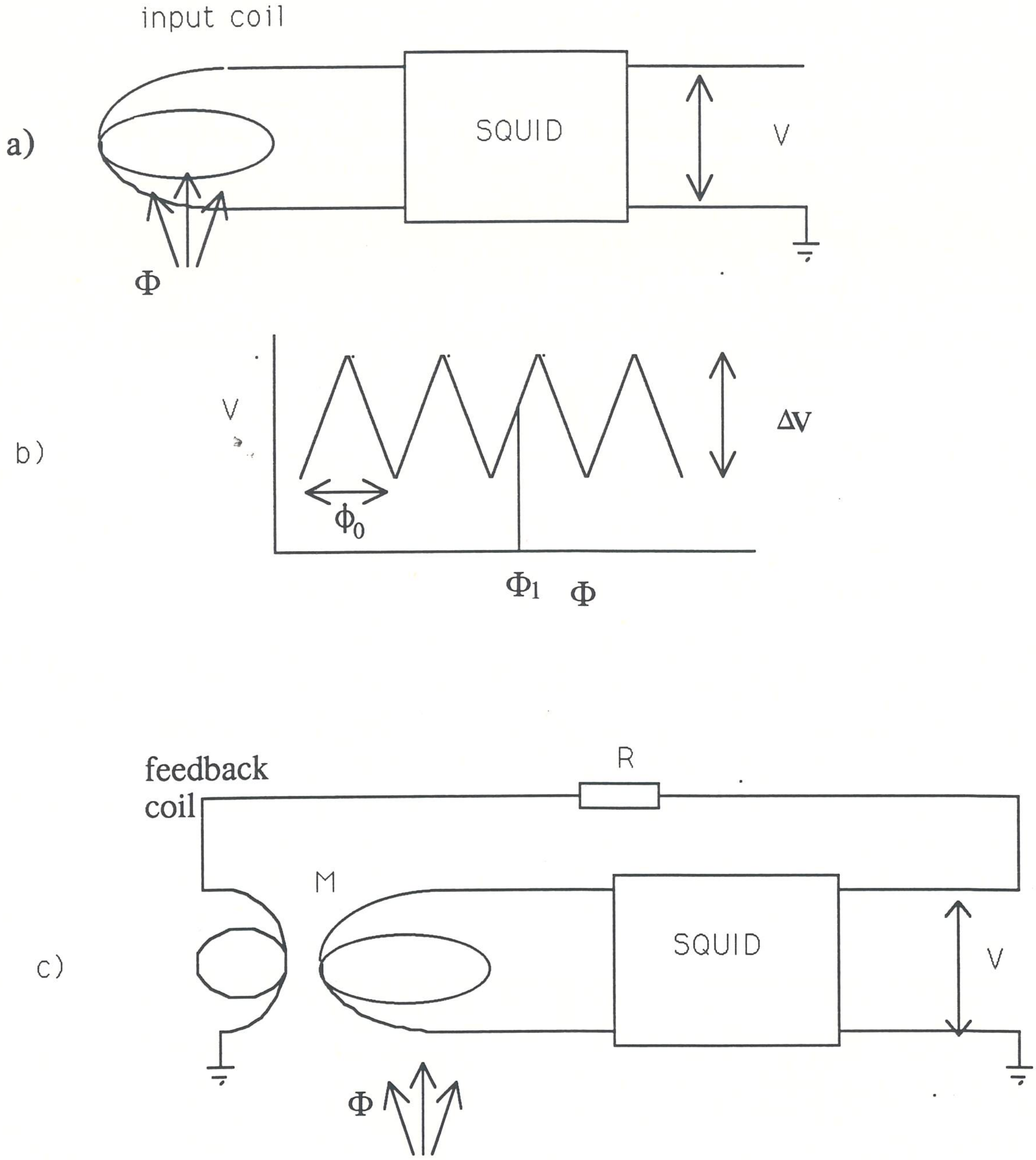

Fig. 3 The SQUID

The system can be than sketched as in Fig. 4

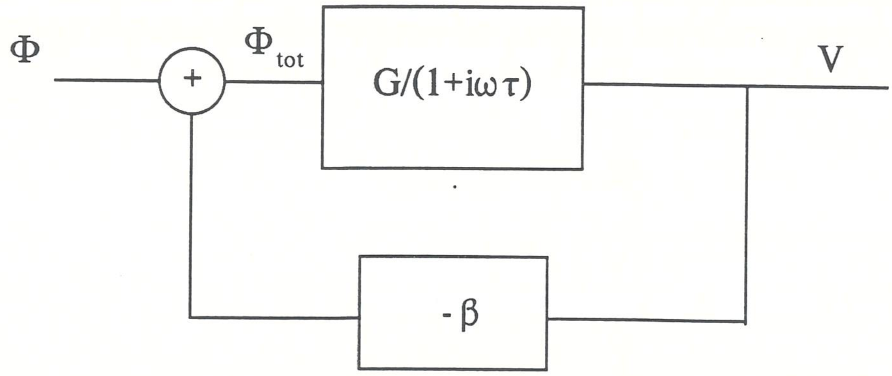

Fig 4. The feedback scheme

Using Fourier transforms one gets

$$V(\omega)=\frac{G}{1+i\omega\tau}(\Phi(\omega)-\beta V(\omega))=\frac{\frac{G}{1+i\omega\tau}}{1+\frac{\beta G}{1+i\omega\tau}}\Phi(\omega)=\frac{G'}{1+i\omega\tau'}\Phi(\omega) \quad 41$$

with $G'=\frac{G}{1+\beta G}$ and $\tau'=\frac{\tau}{1+\beta G}$.

Now it can be seen that if the input to the total system undergoes a variation $\delta\Phi$ the total input to the amplifier is $\delta\Phi-\beta\delta V=\frac{\delta\Phi}{1+\beta G}$. For $\beta G>>1$, a condition that can be fulfilled acting either on G or on β, this is a much smaller variation,that can be inside the linearity range of the system even if $\delta\Phi$ would not. The input to the feedback branch undergoes a variation $\delta VG'$ and obviously this variation has to be within the linearity range of the feedback branch.

The feedback loop has some other interesting features that are worth to be mentioned. First it can be seen that if $G\beta>>1$ then $G'=1/\beta$ and the overall gain of the system, for $\omega\tau<<1$, becomes independent of G. Without the feedback branch, any low frequency fluctuation of the gain δG would have been converted to an output fluctuation $\delta V=\delta G\cdot\Phi$ and thus the relative output fluctuation would have been $\delta V/V=\delta G/G$. With the feedback loop the effective gain is G' whose fluctuation is $\delta G'=\frac{\delta G}{(1+\beta G)^2}$ that translates to a relative output fluctuation $\delta V/V=\frac{\delta G/G}{1+\beta G}<<\delta G/G$.

A second observation is that the frequency behaviour of the feedback ampilfier has the same shape of the original one but the bandpass $1/\tau'$ is now larger than $1/\tau$. The product gain-bandpass is however unchanged as $G/\tau=G'/\tau'$.

Exemple 2.3 Systems described by a linear differential equation with constant coefficients.

Consider the coordinate x(t) of a one dimensional classical damped harmonic oscillator driven by a force f(t). The equation of motion is:

$$m\frac{d^2x}{dt^2}+\beta\frac{dx}{dt}+kx=f \qquad 42$$

The oscillator can be considered as a linear system whose output is x(t) and whose input is f(t). Taking the Laplace transform of both sides for $t\geq 0$ one gets.

$$(mp^2+\beta p+k)x(p)=f(p)+m\left[\left(\frac{dx}{dt}\right)_{t=0^+}+px(0^+)\right]+\beta x(0^+) \qquad 43$$

or

$$x(p)=\frac{f(p)/m}{p^2+p/\tau+\omega_o^2}+\frac{\left[\left(\frac{dx}{dt}\right)_{t=0^+}+px(0^+)\right]+x(0^+)/\tau}{p^2+p/\tau+\omega_o^2} \qquad 44$$

with $\omega_o^2=k/m$ and $\tau=m/\beta$. The transfer function of the system is then $h(p)=\frac{1}{m}\cdot\frac{1}{p^2+p/\tau+\omega_o^2}$ while the second term in eq. 44, represents the free evolution of the system and depends only on the initial conditions. It can be easily cheched that the system is stable and that it inpulse response is $h(t)=\frac{1}{\omega_1}e^{-\frac{t}{2\tau}}\sin(\omega_1 t)\theta(t)$ with $\omega_1^2=\omega_o^2-1/4\tau^2$.

Any system described by a linear differential equation with constant coefficients can be described in a similar way. If the equation is

$$\sum_{k=o}^{N} c_k\frac{d^k o}{dt^k}=\sum_{k=o}^{M} d_k\frac{d^k i}{dt^k} \qquad 45$$

and if we take for simplicity zero initial conditions then

$$h(p)=\frac{\sum_{k=o}^{M} d_k p^k}{\sum_{k=o}^{N} c_k p^k} \qquad 46$$

from eq. 39c it follows that in order the system to be stable one needs M<N.

Multiple inputs and outputs.

All the results obtained above can be generalized to the case of a system with N inputs and M outputs. the input-output relation in the time domain will be, for a time-invariant system,

$$o_k(t)=\sum_{l=o}^{N} \int_{-\infty}^{\infty} h_{kl}(t')i_l(t-t')dt' \qquad 47$$

and in the frequency domain, if the system is stable,

$$o_k(\omega)=\sum_{l=o}^{N} h_{kl}(\omega)i_l(\omega) \qquad 48$$

A simple exemple of a multiple inputs-outputs system is the electric(or mechanic or electromechanic) quadrupole (Fig 5). This is a system with a voltage and a current (or a force and a velocity) as inputs and a voltage and a current (or again a force and a velocity) as outputs. The role of inputs and outputs can be interchanged but obviously the causality is obtained only with one of the possible choices.

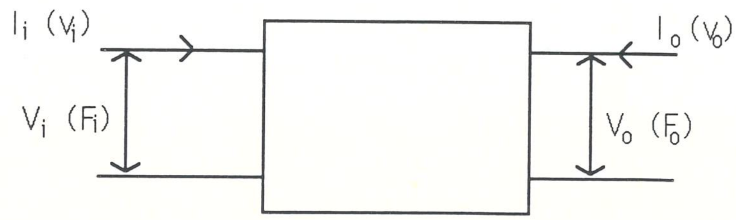

Fig 5. Schematic of an electromechanic quadrupole.

Exemple 2.4 The capacitor as an electromechanical quadrupole.
Consider the electromechanic displacement transducer in Fig. 4

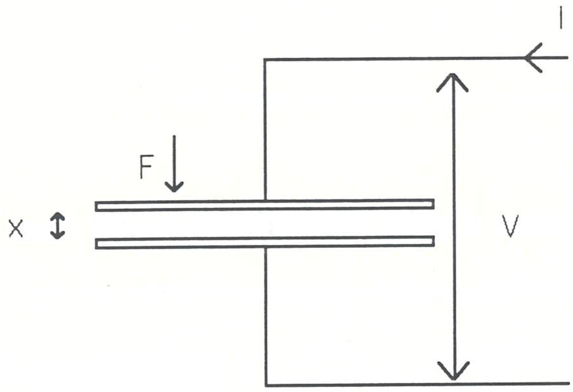

Fig 6 A capacitor as an electromechanic transducer

If the capacitor is charged with a constant electric field E_o (constant charge operation) for a small variation δx of the distance x between the plates, the voltage V=Ex undergoes a variation $\delta V=E_o\delta x$ while for a small change δq of the charge q=CEx (C is the capacity) the voltage changes by $\frac{\delta q}{C_o}$, with C_o the unperturbed capacity. The force on the plates is $F=\frac{1}{2}E\cdot q=\frac{q^2}{8\pi\varepsilon_o S}$, with S the area of the plates. So a small variation δq brings about a change in the force $\delta F=\frac{q_o\delta q}{4\pi\varepsilon_o S}=E_o\delta q$ while δx has no effect on it. With these approximation the system is linearized and, reminding that the velocity is $v=\frac{dx}{dt}$ while the current is $I=\frac{dq}{dt}$, one gets

$$\delta V(\omega)=\frac{E_o\delta v(\omega)}{i\omega}+\frac{\delta I}{i\omega C_o}$$

49

$$\delta F(\omega)= \qquad +\frac{E_o\delta I}{i\omega}$$

Exercise. If a resistor is put in parallel to the capacitor, what is the force to velocity relation in the frequancy domain

The Mixer

Here I will discuss one exemple of system without memory that has great importance in many applications. The system performs the ordinary product of two input signals and is called a multiplier or a mixer:

$$o(t)=i_1(t)\cdot i_2(t)$$

Real mixers introduce obviously finite delays but can be usually treated as ideal systems followed for instance by a low-pass element with impulse response $h(t)=\frac{e^{-t/\tau}}{\tau}\theta(t)$.

If the same signal is sent to both inputs the mixer performs the square of the input and it is called a square wave detector.

The Fourier transform of the mixer output is the convolution of the Fourier transforms of the inputs:

$$o(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty} i_1(\omega')i_2(\omega-\omega')\, d\omega' \qquad 50$$

this is the principal property of a mixer which is used to move along the frequency axis the Fourier transform of any signal.

A very well known application of this property is that to the synchronous transmission and detection of signals. Suppose that one of the two input signals is a narrow band signal as that in eq. 9 and the signal at the other input is a pure sinusoidal function $\sin(\omega_1 t)$. The output signal will then be:

$$o(t)=a(t)\sin^2(\omega_1 t)+b(t)\sin(\omega_1 t)\cos(\omega_1 t)=$$

$$=\frac{1}{2}a(t)-\frac{1}{2}a(t)\cos(2\omega_1 t)+\frac{1}{2}b(t)\sin(2\omega_1 t) \qquad 51$$

whose Fourier transform is

$$o(\omega)=\frac{a(\omega)}{2}-\frac{a(\omega-2\omega_1)+a(\omega+2\omega_1)}{4}+\frac{b(\omega-2\omega_1)-b(\omega+2\omega_1)}{4i} \qquad 52$$

Because both $b(\omega)$ and $a(\omega)$ are limited to frequencies $<<\omega_1$ the output has a line around $\omega=0$ and two lines around $\omega=\pm 2\omega_1$. If the output is filterd by a suitable low pass filter with a roll-off frequency much less than $2\omega_1$, the output $o'(t)$ will have a spectrum

$$o'(\omega)=\frac{o(\omega)}{1+i\omega\tau}\approx\frac{a(\omega)}{2}\frac{1}{1+i\omega\tau} \qquad 53$$

and will substantially recover the signal a(t) i.e. the conponent of the narrow band signal in-phase with $\sin(\omega_1 t)$. It is straightforward to calculate that if the mixer is driven by the signal $\cos(\omega_1 t)$, then it is the signal b(t) which is recovered. The system composed by a mixer driven by a pure harmonic signal (or by a square wave signal; see below) and a low pass filter is often called a phase sensitive detector (PSD)

To multiply a narrow band signal by a a periodic signal at the carrier frequency is often called a down conversion, as the signal is moved from the frequency region around ω_1, the carrier frequency, to that around $\omega=0$. The opposite operation, the up-conversion of a low frequency signal, is also possible again by sending the low frequency signal s(t) to a mixer together with a pure sinusoidal or cosinusoidal carrier. In this case the Fourier transform of the output will be

$$o(\omega)=\frac{s(\omega-\omega_1)-s(\omega+\omega_1)}{2i} \qquad 54a$$

for a sinusoidal carrier and

$$o(\omega)=\frac{s(\omega-\omega_1)+s(\omega+\omega_1)}{2} \qquad 54b$$

for a cosinusoidal one. If the input signal was a low frequency one the output will be narrow-band around ω_1.

The up-conversion is mainly used in applications where the transmission or the detection of the signal is easier, (lower noise, lower attenuation etc) at frequencies higher than those where the signal energy is originally concentrated. After the transmission or the detection, the signal can be eventually down-converted by a PSD driven by a local oscillator and the original behaviour recovered (Fig. 7).

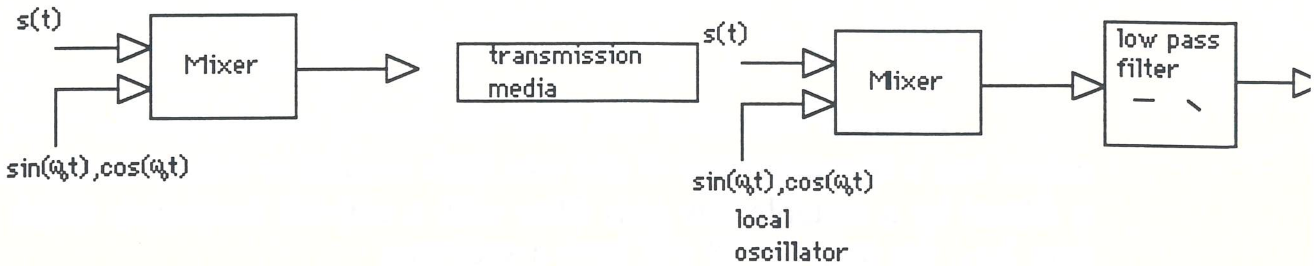

Fig 7 Modulation-demodulation scheme

Exemple. Signal chopping and lock-in amplifier.

The light coming from a point source is detected by a detector subject to low frequency disturbances like drifts or flicker noise. If a rotating wheel with a certain nunber of equally spaced windows is inserted on the light path, the intensity of light coming from the source and reaching the detector will be a square wave whose period will be T/N, with T the rotation period of the wheel and N the number of windows, and with amplitude proportional to the intensity of the point source. The signal can then be recovered by a PSD. In ordinary applications of this kind the PSD is not driven by an harmonic signal but instead by a square wave signal (lock-in amplifiers) which is much easier to generate starting from whatever periodic waveform. If the signal to be recovered is itself a square wave, this is of no consequence (verify). If however only the fundamental harmonic of the narrowband signal has to be recovered, the signal has to be prefiltered by some lowpass filter suppressing high order harmonics. In fact a square wave periodic signal can be expanded as a series of odd harmonics of the signal frequency , $\mathrm{sign}[\sin(\omega_1 t)] = \sum_{l=o}^{\infty} \frac{\pi}{2k+1} \sin[(2k+1)\omega_1 t]$ so that the mixing to a square wave down-converts all the odd harmonics in the input signal , though eachone attenuated by a factor $1/(2k+1)$.

A phase sensitive detector can also be used to measure the Fourier transform of a signal. In fact the output spectrum being $o_s(\omega) = \frac{s(\omega-\omega_1)-s(\omega+\omega_1)}{2i(1+i\omega\tau)}$ (sine wave drive) or $o_c(\omega) = \frac{s(\omega-\omega_1)+s(\omega+\omega_1)}{2(1+i\omega\tau)}$ (cosine wave drive), if τ is chosen that long that $s(\omega)$ can be considered a constant over any frequency interval $\Delta\omega < 1/\tau$, then one gets that $o_s(0) \approx -\mathrm{Im}\{s(\omega_1)\}$, $o_c(0) \approx \mathrm{Re}\{s(\omega_1)\}$.

The Noise

Physical systems are affected by random noise. In the framework of the theory of probability this translates to two facts. On one side a single repetition of an experiment has to be considered an element of a statistical ensemble. As a consequence the value of any signal at a given time t, s(t), can only be predicted in a statistical way. On the other hand, within the single repetition of the experiment, s(t) is a fixed function of time that can be measured and, for instance, recorded on a magnetic disk.

In a statistical experiment, whenever to every possible outcome ξ of the experiment a number $v(\xi)$ is assigned, one says that a **random variable** $\mathbf{v}$ has been defined. If to every ξ a function of a parameter t, let call it $\mathbf{x}(t,\xi)$, is instead assigned , we say that a stochastic process $\mathbf{x}(t)$ has been defined. Thus a noisy signal in a physical apparatus is considered in this framework as a stochastic process.

The statistical properties of a random variable v are known when its probability density function $f_{\mathbf{v}}(x)$ is given. A stochastic process consists of a continous ensemble of random variables which are the values of the process for all the possible choices of t. Its statistical properties are thus known only when it is given the infinite set of probability densities $f_{\mathbf{x}(t)}(x)$, $f_{\mathbf{x}(t),\mathbf{x}(t')}(x,y)$, $f_{\mathbf{x}(t),\mathbf{x}(t'),\mathbf{x}(t'')}(x,y,z)$ etc.. Despite this, most of the information in physical problems is contained in low order densities and in their moments[1] . Among these the most important ones, that we will widely use in the following, are the mean value[2]

$$\eta(t)=<\mathbf{x}(t)>= \int_{-\infty}^{\infty} x f_{\mathbf{x}(t)}(x)dx \qquad 55$$

the autocorrelation

---

[1] The density $f_{\mathbf{x}(t_0),\mathbf{x}(t_1),,,,,\mathbf{x}(t_n)}(x_o,x_1,,,,,x_n)$ is called a density of n^{th} order. Consider a set of n random variables $\mathbf{v}_1,,,\mathbf{v}_n$ and their probability density $f_{\mathbf{v}_1,,,\mathbf{v}_n}(x_1,,,,,x_n)$, a k^{th} order moment is defined as the ensemble average

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx_1...dx_n \, x_1^{m_1}\cdot\cdot x_n^{m_n} f_{\mathbf{v}_1,,,\mathbf{v}_n}(x_1,,,,,x_n) \quad \text{with} \quad \sum_{i=1}^{n} m_i=k.$$

[2] I will indicate with <> the ensemble mean value of any quantity

$$R(t,t')=<\mathbf{x}(t)\mathbf{x}(t')>=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{\mathbf{x}(t)\mathbf{x}(t')}(x,y)\,dx\,dy \qquad 56$$

and the autocovariance

$$C(t,t')=<[\mathbf{x}(t)-\eta(t)][\mathbf{x}(t')-\eta(t')]=R(t,t')-\eta(t)\eta(t') \qquad 57$$

notice that $<\mathbf{x}^2(t)>=R(t,t)$, that the variance $\sigma^2_{\mathbf{x}(t)}=<\mathbf{x}^2(t)>-\eta^2(t)=C(t,t)$ and that the autocovariance is the autocorrelation of the zero mean process $\mathbf{x}(t)-\eta(t)$.

Example 3.1 Shot Noise

Consider the time interval $0\leq t\leq T$ and choose one time instant t_1 at random. With "at random" I mean that the random variable t_1 has a uniform distribution so that the probability that $t_a \leq t_1 \leq t_b$ is $P\{t_a \leq t_1 \leq t_b\}=\frac{t_b-t_a}{T}$ for $t_b\geq t_a\geq 0$. If the experiment is repeated N times, so that the N random variables $t_1,,,,,,t_N$ are independent, then the probability to get k points in the interval $t_a \leq t_1 \leq t_b$ will be given by the binomial probability

$$P(k,N)=\frac{N!}{k!(N-k)!}\left(\frac{t_b-t_a}{T}\right)^k\left(1-\frac{t_b-t_a}{T}\right)^{N-k} \qquad 58$$

while the probabilty to get k points in the above interval and k' in the interval $t_c \leq t \leq t_d$, assuming that these two intervals do not overlap, is given by

$$P(k,k',N)=\frac{N!}{k!k'!(N-k-k')!}\left(\frac{t_b-t_a}{T}\right)^k\left(\frac{t_d-t_c}{T}\right)^{k'}\left(1-\frac{t_b-t_a}{T}-\frac{t_d-t_c}{T}\right)^{N-k-k'} \qquad 59$$

If N and T both $\rightarrow\infty$ but $N/T=\lambda$, then taking the limits of eq. 58 and 59 one gets

$$P(k)=e^{-\lambda(t_b-t_a)}\frac{[\lambda(t_b-t_a)]^k}{k!} \qquad 60$$

and

$$P(k,k')=e^{-\lambda(t_b-t_a)}\frac{[\lambda(t_b-t_a)]^k}{k!}e^{-\lambda(t_d-t_c)}\frac{[\lambda(t_d-t_c)]^{k'}}{k'!}=$$

$$=P(k)\cdot P(k') \qquad 61$$

Thus the random variable k is Poisson distributed, and the number of points in two non overlapping time intervals are independent random variables.

Consider now the stochastic process:

$$\mathbf{x}(t)=\Sigma_{i,t_i\leq t}\,\theta(t-t_i) \qquad 62$$

that counts the number of points such that $t_i \leq t$. This process is a good model for all physical situations where a beam of particles hits a target or a detector located at a certain position in space. t_i is then the arrival time of the i^{th} particle of the beam and $\mathbf{x}(t)$ is the total number of particles that have already hit the detector at time t.

For a specific repetition of the experiment $\mathbf{x}(t)$ is just a never decreasing staircase curve made of steps of random time duration and unit height, while for a fixed time t, it is a Poisson distributed random variable.

The mean value of $\mathbf{x}(t)$ is $\eta(t)=\lambda t$. The autocorrelation can be evaluated using the property that the numbers of points in two non overlapping intervals are independent variables. This means that their joint probability density is just the product of their respective densities and, as a consequence, the mean value of their product is just the product of their mean values. The number of points between 0 and t is $\mathbf{x}(t)$ while the number of points between t and t', assuming $t \leq t'$, is $\mathbf{x}(t')-\mathbf{x}(t)$. Thus

$$<[\mathbf{x}(t')-\mathbf{x}(t)]\mathbf{x}(t)]>=R(t',t)-R(t,t)=\lambda(t'-t)\cdot\lambda t \qquad 63$$

and considering that $R(t,t)=<\mathbf{x}^2(t)>=\lambda t+\lambda^2 t^2$ (from basic properties of Poisson distribution), one gets

$$R(t't)=\lambda^2 t't+\lambda t \qquad 64$$

and

$$C(t',t)=\lambda t \qquad 65$$

A special class of stochastic processes is that of normal processes. A process is called normal if the joint probability density of any order is normal. With normal I mean that:

$$f_{\mathbf{x}(t_0),\mathbf{x}(t_1),,,,,\mathbf{x}(t_n)}(x_o,x_1,,,,,x_n)=$$

$$=\frac{|\mu|}{(2\pi)^{N/2}}\cdot\exp\left(-\frac{1}{2}\sum_{i,j=1}^{n}\mu_{ij}[x_i-\eta(t_i)][x_j-\eta(t_j)]\right) \qquad 66$$

where the matrix μ_{ij} is the inverse of $C(t_i,t_j)$ and $|\mu|$ is its determinant.

As it can be seen from eq. the statistical properties of a process of this kind are all contained in the mean value and in the autocorrelation.

Example 3.2 High density shot noise

In the limit $\lambda\rightarrow\infty$, in exemple 3.1, the Poisson distribution tends to the Gaussian one. Thus in this limit $\mathbf{x}(t)$ is a normal random variable. On the other side the process $\delta\mathbf{x}(t)=\mathbf{x}(t+\tau)-\mathbf{x}(t)$ and $\mathbf{x}(t)$ are statistically independent so their joint probability density is just the product of two Gaussians and is then a Gaussian. Whatever linear transformation of Gaussian random variables still gives Gaussian variables so that $\mathbf{x}(t+\tau)=\delta\mathbf{x}(t)+\mathbf{x}(t)$ and $\mathbf{x}(t)$ are joint Gaussian.

Stationary processes.

If the statistical properties of a process are not affected by a shift of the time origin, $t\rightarrow t+\Delta t$, then the process is called stationary. For a stationary process, the density functions of any order can only depend on time delays and not on the single time arguments of the process. Thus the first order density, $f_{\mathbf{x}(t)}(x)$, has to be independent of time, $f_{\mathbf{x}(t)}(x)=f_{\mathbf{x}}(x)$, and the second order one $f_{\mathbf{x}(t_0),\mathbf{x}(t_1)}(x_o,x_1)$ can only depend on t_o-t_1, $f_{\mathbf{x},\mathbf{x}}(x,y,t_o-t_1)$. As a consequence the mean value is independent of time $\eta(t)=\eta$ and the autocorrelation is $R(t,t')=R(\tau)$ with $\tau=t-t'$.

A stochastic processes with time independent mean value and with the autocorrelation depending only on the delay t-t' is called wide sense stationary. A stationary process is wide sense stationary. If a process is normal and wide sense stationary it is also stationary.

The autocorrelation of a stationary process has some simple properties that is worth mentioning. From the defintion eq. it follows that

$$R(\tau)=R(-\tau). \qquad 67$$

from $<[\mathbf{x}(t+\tau)\pm\mathbf{x}(t)]^2>=2R(0)\pm2R(\tau)\geq0$ one gets that

$$R(0)\geq|R(\tau)| \qquad 68$$

Notice that if for $\tau > \tau_o$ $R(\tau)$ vanishes, then $\mathbf{x}(t)$ and $\mathbf{x}(t+\tau)$ become uncorrelated and, if the process is normal, they become also independent.

In order to give a physical interpretation of the autocorrelation, let consider a stationary gaussian process with zero mean. It can be shown that the conditional probability density of the process $\mathbf{x}(t+\tau)$ at time $t+\tau$, with the condition of having found the value $\mathbf{x}(t)$ at time t has a mean value $<\mathbf{x}(t+\tau)|\mathbf{x}(t)>=\mathbf{x}(t)R(\tau)/R(0)$ and variance $\sigma^2=R(0)[1-R^2(\tau)/R^2(0)]$. Thus for $\tau\to 0$ and $R(\tau)/R(0)\to 1$ $\mathbf{x}(t+\tau)$ fluctuates with very small uncertainty around the previous value $\mathbf{x}(t)$. As soon as $R(\tau)/R(0)$ approaches zero, both the mean value goes to zero, which is the unconditional mean value of the process, and the variance goes to R(0), which again is its unconditional value. The autocorrelation is then considered often as a measure of the "memory" of the process.

Exemple 3.3: Random Telegraph Signal

Consider a physical system that has only two possible states. The states are identified by the value of some physical quantity x that can take the values x=1 in one of the two states and x=0 in the other one. We will assume that if the system is found at time t in the state x=1(0) its probability to be found in the state x=0(1) after a very short time τ is $p_-\tau$ ($p_+\tau$). These two probabilities are conditional probabilities. In order to get the total probability for the system to be found in the state x=1 at a time $t+\tau$ one has to multiply the total probability $P_1(t)$ of having found the system at time t in the x=1 state, by the probability $(1-p_-\tau)$ that the system had remained in that state during the time $t,t+\tau$. To that one shoud then add the probability of having found the system in the state x=0 at time t, $P_0(t)=1-P_1(t)$ times the probability $p_+\tau$ of jumping from this state to the other. Then $P_1(t+\tau)=P_1(t)(1-p_-\tau)+(1-P_1(t))p_+\tau$ or $\frac{P_1(t+\tau)-P_1(t)}{\tau}+P_1(t)\cdot(p_-+p_+)=p_+$. Taking the limit for $\tau\to 0$ one gets $\frac{dP_1}{dt}+P_1(p_++p_-)=p_+$. The solution is

$$P_1(t)=P_1(0)e^{-\lambda t}+P_\infty\left(1-e^{-\lambda t}\right) \tag{69}$$

with $\lambda=p_++p_-$ $P_\infty=\frac{p_+}{p_++p_-}$

If $P_1(0)=P_\infty$ then $P_1(t)=P_1(0)=P_\infty$.

Let now define the process $\mathbf{x}(t)$ given by the istantaneous value of x at time t. $\mathbf{x}(t)$ can only take the values x=0 or x=1. In the case of the time independent solution $P_1(t)=P_\infty$, the mean value of the process x(t) is independent of time and is $\eta=P_\infty$. If instead the process is "prepared" in the state x=1 at t=0 , then $P_+(0)=1$ and the mean value relaxes to its stationary value in a time $\approx 1/\lambda$.

Let now calculate the autocorrelation. To do that we have to calculate the probability $P_{11}(t,\tau)$ that the system is found at x=1 both at time t and at time $t+\tau$. The autocorrelation is then $R(t,t+\tau)=P_{11}(t,\tau)$ as in any other case $x(t)x(t+\tau)=0$. Now one can calculate that

$$P_{11}(t,\tau+\delta t)=P_{11}(t,\tau)(1-p_-\delta t)+P_{10}(t,\tau)p_+\delta t$$

and

$$P_{10}(t,\tau+\delta t)=P_{11}(t,\tau)p_-\delta t+P_{10}(t,\tau)(1-p_-\delta t)$$

where $P_{10}(t,\tau)$ is the probability to find the system with x=1 at t and x=0 at t+τ Taking the limit for $\delta t \to 0$ we get

$$\frac{dP_{11}(t,\tau)}{d\tau} + p_- P_{11}(t,\tau) - p_+ P_{10}(t,\tau) = 0 \qquad 70a$$

and

$$\frac{dP_{10}(t,\tau)}{d\tau} + p_+ P_{10}(t,\tau) - p_- P_{11}(t,\tau) = 0 \qquad 70b$$

that can be integrated using Laplace transforms to give

$$P_{11}(t,i\omega)(i\omega+p_-) - P_{10}(t,i\omega)p_+ = P_{11}(t,\tau=0^+) \qquad 70c$$

$$-P_{11}(t,i\omega)p_- \quad + P_{10}(t,i\omega)(i\omega+p_+) = P_{10}(t,\tau=0^+) \qquad 70d$$

Now $P_{11}(t,\tau=0^+)=P_1(t)$ and $P_{10}(t,\tau=0^+)=0$ so that

$$P_{11}(t,i\omega)=P_{11}(i\omega)=P_1(t)\left[\frac{1}{i\omega}\cdot P_\infty + \frac{1}{i\omega+\lambda}\cdot(1-P_\infty)\right] \qquad 71$$

that, in the time domain becomes, for $\tau \geq 0$,

$$P_{11}(t,\tau)=R(t,\tau)=P_1(t)\left[P_\infty+(1-P_\infty)e^{-\lambda\tau}\right] \qquad 72$$

while the autocovariance, after some calculation, is

$$C(t,\tau)=P_1(t)e^{-\lambda\tau}\{\,1-P_\infty+[P_\infty - P_1(0)]\,e^{-\lambda t}\} \qquad 73$$

Notice that eq. 72 and 73 both hold even for negative τ provided $|\tau|$ is substituted to τ in the exponentials and provided also that $t+\tau \geq 0$. If $t \to \infty$ then R depends only on $|\tau|$ and the process is wide sense stationary . A process of this kind is called asyntotically wide sense stationary. It can be demonstarted that it is asyntotically stationary in strict sense too.

It is worth now to notice that if the system is found at t=0 in the state x=1, after Δt seconds the interval in which it is probably found is $\eta(\Delta t) \pm \sigma(\Delta t) = \eta(\Delta t) \pm \sqrt{C(\Delta t,0)}$. For $\lambda\Delta t \ll 1$ this gives $\eta(\Delta t) \pm \sigma(\Delta t) = 1 \pm \sqrt{(1-P_\infty)\lambda\Delta t}$ so that the system preserve the memory of his previous state for a time $1/\lambda$ which is the time decay constant of its stationary autocorrelation.

If the system we are speaking of is in thermal equilibrium, then

$\frac{P_+}{P_-} = \frac{p_+}{p_-} = e^{-\frac{F_1-F_o}{k_B T}}$ and $C(\tau)=\frac{e^{-\frac{F_1-F_o}{k_B T}}}{\left[1+e^{-\frac{F_1-F_o}{k_B T}}\right]^2} e^{-(p_-+p_+)|\tau|}$, where $F_{1,0}$is the free energy of the state x=1,0. We will use this result later on.

Power Spectrum

If a process is wide sense stationary, then its autocorrelation is a function of the single variable τ and can be Fourier transformed. The Fourier transform of the autocorrelation

$$S(\omega)= \int_{-\infty}^{\infty} R(\tau)e^{-i\omega\tau}d\tau \qquad 74$$

is called the power spectrum or the power spectral density of the stochastic process and plays a role of paramount importance in noise analysis (see below). Notice that, because $R(\tau)=R(-\tau)$, then $S(\omega)\geq 0$. Also notice that

$$<\mathbf{x}^2>=R(0)=\frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)d\omega \qquad 75$$

At the and of this section, let briefly extend the basic definitions to the case of more than one stochastic process. The statistical properties of two stochastic processes $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are known when the joint probability densities of any order are given. As in the case of one single stochastic process, most of the information is often contained in the densities and in the moments of lowest order such as the cross correlation:

$$R_{\mathbf{xy}}(t,t')=<\mathbf{x}(t)\mathbf{y}(t')> \qquad 76$$

and the crosscovariance

$$C_{\mathbf{xy}}(t,t')=R_{\mathbf{xy}}(t,t')-\eta_{\mathbf{x}}(t)\eta_{\mathbf{y}}(t') \qquad 77$$

two processes are joint stationary if their joint statistics are unaffected by a time origin shift and they are only wide sense stationary if only their mean values and their auto and crosscorrelations are unaffected by the shift. In this case the crosscorrelation can be Fourier transformed and a cross-power spectrum $S_{xy}(\omega)$ can be defined. Further properties of the crosscorrelation will be discussed in the following.

Stochastic processes and linear systems

As already stated, within a single repetition of an experiment, a stochastic process is a well defined function of its parameter. It can then be treated as any other signal and used as an input to a linear system. Suppose the system is described by a non stationary inpulse response h(t,t'), then the output of the system will just described by eq. 27 with **x**(t) in place of i(t). The output, let's call it **y**(t), will now be a stochastic process. Indeed, as in any other repetition of the experiment **x**(t) will change in a non fully predictable way, so will do **y**(t).

To evaluate the mean value of **y**(t) consider that an integral is just a limit of a linear combination and that the mean value of a linear combination of random variables is just the linear combination of their mean values. As a consequence:

$$<\mathbf{y}(t)>=<\int_{-\infty}^{\infty}h(t,t')\mathbf{x}(t')dt'>=\int_{-\infty}^{\infty}h(t,t')<\mathbf{x}(t')>dt' \qquad 78$$

With the same rule one can evaluate the cross correlation between the input and the output

$$<\mathbf{x}(t)\mathbf{y}(t')>=\int_{-\infty}^{\infty}h(t',t'')<\mathbf{x}(t)\mathbf{x}(t'')>dt'' \qquad 79$$

and the autocorrelation of the output

$$<\mathbf{y}(t)\mathbf{y}(t')>=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}h(t,t'')h(t',t^*)<\mathbf{x}(t'')\mathbf{x}(t^*)>dt''dt^* \qquad 80$$

The same reasoning can be extended to whatever moment of inputs and outputs of any linear system characterized by a linear operator L_t acting on a function of the argument t :

$$<L_{t_1}\{i(t_1)\}L_{t_2}\{i(t_2)\}L_{t_n}\{i(t_n)\}>=L_{t_1}L_{t_2}L_{t_n}\{<i(t_1)i(t_2)i(t_n)>\} \qquad 81$$

Thus for instance if the system performs the ordinary derivative

$$y(t)=\frac{dx}{dt} \qquad 82$$

then

$$<\mathbf{y}(t)>=\frac{d<\mathbf{x}>}{dt} \qquad 83$$

and

$$<\mathbf{x}(t),\mathbf{y}(t')>=\frac{\partial R_{\mathbf{xx}}(t,t')}{\partial t'} \qquad 84$$

$$<\mathbf{y}(t),\mathbf{y}(t')>=\frac{\partial^2 R_{\mathbf{xx}}(t,t')}{\partial t\ \partial t'} \qquad 85$$

If both the system is time invariant at the the input is stationary, then the output is stationary and equations from 78 to 85 become

$$<\mathbf{y}>=<\mathbf{x}>\int_{-\infty}^{\infty}h(t)dt=h(\omega=0) \qquad 78a$$

$$<\mathbf{x}(t)\mathbf{y}(t+\tau)>=R_{\mathbf{xy}}(\tau)=\int_{-\infty}^{\infty}h(t')R_{\mathbf{xx}}(\tau-t')dt'' \qquad 79a$$

$$<\mathbf{y}(t)\mathbf{y}(t+\tau)>=R_{\mathbf{yy}}(\tau)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}h(t')h(t'')R_{\mathbf{xx}}(\tau-t''+t')dt'dt'' \qquad 80a$$

$$<\mathbf{y}(t)>=\frac{d<\mathbf{x}>}{dt}=0 \qquad 83a$$

$$<\mathbf{x}(t),\mathbf{y}(t+\tau)>=\frac{dR_{\mathbf{xx}}(\tau)}{dt} \qquad 84a$$

$$<\mathbf{y}(t),\mathbf{y}(t+\tau)>=-\frac{d^2R_{\mathbf{xx}}(\tau)}{d\tau^2} \qquad 85a$$

Eq. 79a and 80a can be Fourier transformed to obtain the two fundamental results:

$$S_{\mathbf{xy}}(\omega)=h(\omega)S_{\mathbf{xx}}(\omega) \qquad 86$$

$$S_{\mathbf{yy}}(\omega)=h^*(\omega)h(\omega)S_{\mathbf{xx}}(\omega)=\left|h(\omega)\right|^2 S_{\mathbf{xx}}(\omega) \qquad 87$$

Exemple 3.4 The shot noise current.

Taking the derivative of the process in eq. 62 one obtains a new process $\mathbf{y}(t)$:

$$\mathbf{y}(t)=\sum_{i,t_i\leq t}\delta(t-t_i) \qquad 88$$

if $q\mathbf{x}(t)$ represents the total charge arrived on a detector due to a flux of elementary carriers of charge q, then $q\mathbf{y}(t)$ represents the electric current trough the detector surface.

The mean value of $\mathbf{y}(t)$ is

$$\eta_{\mathbf{y}}(t)=\frac{d\lambda t}{dt}=\lambda \qquad 89$$

to calculate the autocorrelation of $\mathbf{y}(t)$ let rewrite that of $\mathbf{x}(t)$ as $R_{\mathbf{xx}}(t,t')=\lambda^2tt'+\lambda[t'-(t'-t)\theta(t'-t)]$; applying eq. 85 we get

$$R_{\mathbf{yy}}(t,t')=\lambda^2+\lambda\frac{d}{dt}\left\{\frac{d[t'-(t'-t)\theta(t'-t)]}{dt'}\right\}= \qquad 90$$

$$=\lambda^2+\lambda\frac{d}{dt}[(1-(t'-t)\delta(t'-t)-\theta(t'-t)]=\lambda^2+\lambda\delta(t'-t)$$

The results above have been obtained for both t and t' >0. However the beginning of the process can now be pushed backward to -∞ and thus the process has become (wide sense) stationary.

Notice that the "fluctuating current" $y(t)-\lambda$ has zero mean and autocorrelation $\lambda\delta(\tau)$. The power spectrum is then $S_{yy}(\omega)=\lambda$. A frequency independent power spectrum is called a white noise. If $\lambda\rightarrow\infty$ than the process is also normal.

If the detector has a finite response time, one can then model the system as an ideal shot noise filtered by a low pass filter with inpulse response $h(\omega)=\frac{1}{1+i\omega\tau}$. The filter output z(t) will then be

$$z(t)=\sum_{i,t_i\leq t}\theta(t-t_i)\frac{e^{-(t-t_i)/\tau}}{\tau}$$

with mean value λ. The power spectrum of the fluctuating part is just $S_{zz}(\omega)=\frac{\lambda}{1+\omega^2\tau^2}$

Notice that the root mean square deviation of the fluctuating part of the unfiltered signal is $<y^2>=\lambda\delta(0)=\infty$ while that of the filtered one is

$$<z^2>=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\lambda}{1+\omega^2\tau^2}d\omega=\frac{\lambda}{2\tau}$$

The detector measures then a mean current $I=q\lambda$ with root mean square fluctuation $\sigma_I=q\sqrt{\frac{\lambda}{2\tau}}=\sqrt{\frac{qI}{2\tau}}$. By measuring σ_I, I and τ the charge of the carrier q can be obtained.

Exemple 3.5 The spectral density as a density of noise energy per unit angular frequency.

To clarify the physical meaning of the power spectrum of a stationary noise, assume that a stationary random force f(t) with spectrum $S_f(\omega)$ drives the harmonic oscillator of example 2.3.

The frequency response of the oscillator is

$$h(\omega)=\frac{1}{m}\cdot\frac{1}{\omega_o^2-\omega^2+i\omega/\tau} \qquad 91$$

If the Q-factor of the oscillator is Q>>1, $|h(\omega)|$ is a narrow line that peaks at resonance, $\omega=\omega_o$, at the value $|h(\omega)|\approx\frac{\tau}{m\omega_o}$ and has a bandwith $\delta\omega\approx 1/\tau$. The oscillator is then a good narrow band filter, though it amplifies any force signal with angular frequency ω_o by the factor $\frac{\tau}{m\omega_o}$. To get rid of this unwanted amplification we will divide the oscillator coordinate by the same factor so that the overall response of the filter is now

$$h(\omega)=\frac{\omega_o/\tau}{\omega_o^2 - \omega^2 + i\omega/\tau} \qquad 91a$$

Applying eq. 87 one obtains, for the power spectrum of the filter output $\mathbf{x}(t)$

$$S_x(\omega)=\frac{\omega_o^2}{\tau^2}\cdot\frac{1}{(\omega_o^2 - \omega^2)^2 + \omega^2/\tau^2} S_f(\omega) \qquad 92$$

If $S_f(\omega)$ is a slowly changing function of ω, then it can be considered a constant within the range $\omega_o \pm 1/\tau$ and eq. 92 becomes

$$S_x(\omega)=\frac{\omega_o^2}{\tau^2}\cdot\frac{1}{(\omega_o^2 - \omega^2)^2 + \omega^2/\tau^2} S_f(\omega_o) \qquad 92a$$

If $f(t)$ has zero mean, $\mathbf{x}(t)$ will have zero mean too. Its r.m.s deviation can be then evaluated from

$$\langle \mathbf{x}^2 \rangle = \frac{1}{2\pi}\cdot\int_{-\infty}^{\infty} S_x(\omega)d\omega = \frac{S_f(\omega_o)}{2\tau} \qquad 93$$

and is then $\sigma_{\mathbf{x}}=\sqrt{\frac{S_f(\omega_o)}{2\tau}}$. Thus the total "noise energy" of the force signal in a frequency band $\delta\omega=\pm 1/\tau$ around the frequency $\omega=\omega_o$, is $S_f(\omega_o)\cdot\delta\omega$ and the spectral density plays the role of an energy density per unit angular frequency.

Thermal Noise

We now will discuss a source of noise which has great importance in many experimental apparatuses. I will use the discussion of this noise phenomenon has a mean to illustrate various aspects of the ideas we have discussed up to now. I will focus on linear electrical and mechanical devices though the results obtained are of a more general relevance. I will briefly discuss at the end a general theory that applies to whatever physical system in thermal equilibrium.

Let consider first a small mass moving in a viscous fluid. The mass is small compared to the precision of the measurement of its position but it is larger than the molecular scale. The scattering of the fluid molecules give random inpulses to the particle that is then subject to a stochastic force **f**(t). As a consequence, the position of the particle **x**(t) and its velocity **v**(t) are both stochastic processes. The force has a non zero mean value: it is known that, if the particle moves relative to the fluid with velocity **v**(t), then the mean value of the force is $<f(t)>=-\beta v(t)$. The force can then be separated in a mean force $<f(t)>$ and in a random component $f_r(t)=f(t)-<f(t)>$

As the time scale of the observation of the particle motion is usually much longer then that of the molecular collisions, the force can be assumed to result from many statistically independent collisions. This has two consequencies: on one side the statistical memory of the process will decay much faster than any measurement time, so that the autocorrelation of $f_r(t)$ can be assumed to be delta-like, $R(\tau)=P\delta(\tau)$, with P a constant. On the other side it is likely that the force is a normal process as a consequence of the central limit theorem that states that a linear combination of independent random variables tends rapidly to be distributed normally when the number of variables increases. Thus, summarizing, the force $f_r(t)$ is a white gaussian noise with zero mean and power spectrum $S_f(\omega)=P$.

The velocity of the particle has to obey Newton's law, $m\frac{dv}{dt}=f(t)$ that translates into:

$$\frac{dv}{dt}+\frac{\beta}{m}v=\frac{f_r}{m} \qquad 94$$

The velocity is then the output of a linear system the input of which is the random white force f_r. The transfer function of the system is

$$h(\omega)=\frac{1}{m}\cdot\frac{\tau^*}{1+i\omega\tau^*} \qquad 95$$

with $\tau^*=\frac{m}{\beta}$. The power spectrum of the velocity is thus

$$S_v(\omega)=\frac{1}{m^2}\cdot\frac{P\tau^{*2}}{1+\omega^2\tau^{*2}} \qquad 96$$

the mean square velocity is readily obtained from

$$<v^2>=\frac{1}{2\pi}\cdot\int_{-\infty}^{\infty}S_v(\omega)d\omega=\frac{P\tau^*}{2m^2} \qquad 97$$

and the mean kinetik energy $<E_k>=\frac{P\tau^*}{4m}$.

Now, according to the equipartition theorem, the mean kinetic energy of a free particle at thermodinamic equilibrium is $<E_k>=\frac{1}{2}k_BT$ so that $P=2k_BTm/\tau^*=2k_BT\beta$. Thus, within this model

$$S_f(\omega)=2k_BT\beta \qquad 98a$$

and

$$S_v(\omega)=\frac{\tau^*}{m}\cdot\frac{2k_BT}{1+\omega^2\tau^{*2}} \qquad 98b$$

If a force $f_o\cos(\omega t)$ is applied to the particle, the velocity will then be $v(t)=\frac{f_o\tau^*}{m}\ \frac{\cos(\omega t)+\omega\tau^*\sin(\omega t)}{1+\omega^2\tau^{*2}}$ and the mean energy dissipated per cycle will be $E=\frac{1}{2}\frac{f_o^2\tau^*}{m}\ \frac{1}{1+\omega^2\tau^{*2}}$. Thus the fluctuation of the velocity contains the same coefficient that links the energy dissipation to the square of the amplitude of the driving frorce. This is only a special result of a more general theorem, called the fluctuation dissipation theorem, that we will discuss later on.

Exemple 3.5. Brownian noise in the harmonic oscillator.

We can apply the result we already obtained for a free particle in a viscous fluid to discuss the slightly more complicate case of a particle in a viscous fluid which is also subject to the action of an elastic force. This is a good model for any mechanical damped oscillator at thermal equilibrium.

The equation of motion of the harmonic oscillator is eq. 42. The oscillator is than a linear system with frequency response $h(\omega)=\frac{1}{m}\cdot\frac{1}{\omega_o^2-\omega^2+i\omega/\tau}$. If the stochastic force due to the damping is used as the input the output will have a spectrum:

$$S_x(\omega)=\frac{2\beta k_BT}{m^2}\cdot\frac{1}{(\omega_o^2-\omega^2)^2+\omega^2/\tau^2} \qquad 99$$

The autocorrelation can be calculated from 99 and is

$$R(t)=\frac{k_BT}{m\omega_o^2}\cdot e^{-t/2\tau}\left[\cos(\omega_1 t)+\frac{1}{2\omega_1\tau}\sin(\omega_1)\right] \quad 100$$

One can immediately check that the average pontential energy $<U>=\frac{1}{2}\cdot m\omega_o^2<x^2>=\frac{1}{2}k_BT$ so that the method is self consistent.

The oscillator noise is a " narrowband" noise: Its spectrum is made of two lines centered around $\omega_1=\pm\omega_o\sqrt{1-1/4Q^2}$.

Thermal noise in linear networks

The dissipation that gives origin to Ohm's low in ordinary resistors also take place because a large amount of uncorrelated scattering events as in the case of the particle moving in the fluid. It is again a reasonable model to add to the mean value voltage V(t)=RI(t) a random zero mean gaussian fluctuating voltage **V**(t) with autocorrelation $R(\tau)=P\delta(\tau)$. In the language of the linear circuit theory this translates to thinking at the resistor as an ideal element followed by a random voltage generator that generates the **V**(t) (Fig 8a).

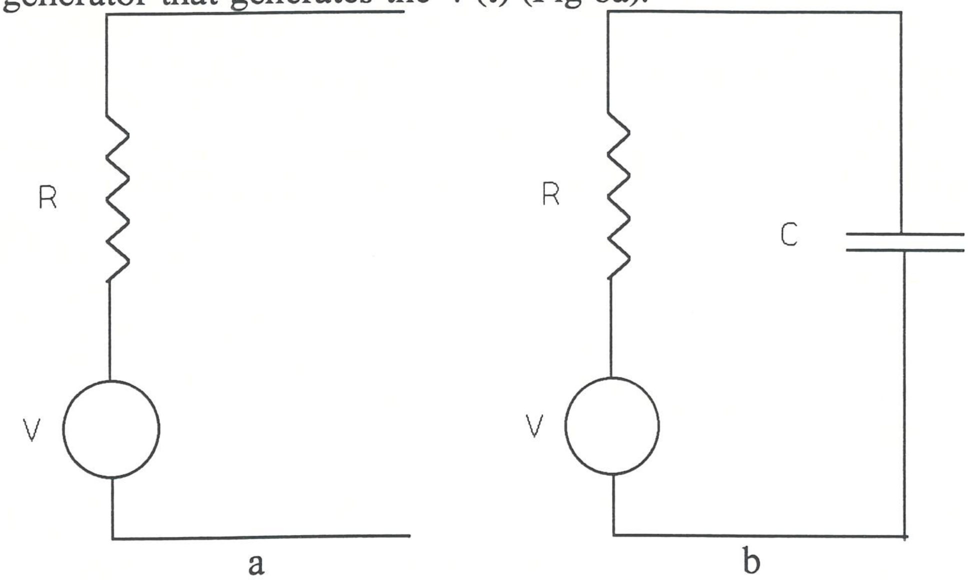

Fig 8

Let now put this element in series to an ideal capacitor of capacitance C, as in Fig 8b. **V**(t) can then be considered as an input to a linear circuit whose output is the voltage $\mathbf{V_c}(t)$ across the capacitor. The frequency response of the linear system is $h(\omega)=(1+i\omega RC)^{-1}$ so that the power spectrum of the output is $S_{Vc}(\omega)=\frac{P}{1+i\omega RC}$. The mean square value of

$V_c(t)$ is $<V_c^2> = \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} \frac{P d\omega}{1+i\omega RC} = \frac{P}{2RC}$. Again inposing that the mean energy stored in the capacitor has to be equal to $\frac{1}{2}$ kT, one gets:

$$P=2kTR \qquad 101$$

Notice that eq. 98a can be derived from eq. 101 on the basis of the electromechanical equivalence voltage→ force current→velocity.

If a resistor can be modeled as in Fig. 8a, than by the voltage current substitution it can also be modeled as an ideal resistor with a current generator in parallel that generates a random current $I(t)=V(t)/R$. The spectral density of the current is than $S_I(\omega)=S_V(\omega)/R^2=2kT/R$.

To calculate the voltage and current noise in more complicate circuits, one can then associate to any resistor a white noise generator with spectrum given by eq. 101, the noise voltages coming from different generators being independent. It is useful, in this kind of calculations, to know the following two rules: if $x(t)$ and $y(t)$ are two stationary random processes and $z(t)=ax(t)+by(t)$, then

$$R_{zz}(\tau)=a^2R_{xx}(\tau)+b^2R_{yy}(\tau)+ab[R_{xy}(\tau)+R_{xy}(-\tau)] \qquad 102a$$

and

$$S_{zz}(\omega)=a^2S_{xx}(\tau)+b^2S_{yy}(\tau)+2abRe\{S_{xy}(\omega)\} \qquad 102b.$$

Example 3.6

Consider the circuit in Fig 9

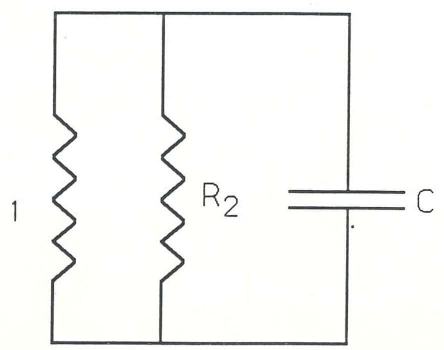

Fig 9

The resistor R_1 generates the voltage V_1 and the resistor R_2 gives origin to the voltage V_2. The voltage across the capacitor in the frequency domain would be, for ordinary signals $V_c = \frac{V_1(\omega)R_2}{R_1+R_2+i\omega CR_1R_2} + \frac{V_2(\omega)R_1}{R_1+R_2+i\omega CR_1R_2}$. Because the two contributions are statistically independent, the spectrum of V_c will just be the sum of the two and thus will be $S_{Vc}(\omega)= \frac{2kTR_1R_2^2}{(R_1+R_2)^2+(\omega CR_1R_2)^2} + \frac{2kTR_2R_1^2}{(R_1+R_2)^2+(\omega CR_1R_2)^2}$. It is immediate to check that $S_{Vc}(\omega)= \frac{2kTR_{//}}{1+(\omega CR_{//})^2}$ with $R_{//}$the parallel resistance of the two resistors.

Alternatively, if one wants to know the voltage noise spectrum between two points of a linear network one can consider the result in the following Suppose between two point of a network there is an impedance $Z(\omega)$. Suppose that a resistor is connected to the port. The thermal noise voltage $V_R(t)$ due to the resistor will induce a random current $I_{RZ}(t)$ in the port and a voltage $V_{RZ}(t)$ across it. The mean power dissipated by the thermal noise would be $P=\langle V_{RZ}(t)I_{RZ}(t)\rangle=R_{VI}(0)$. Now the voltage $V_{RZ}(t)$ can be considered as the output of a linear device with frequency response $Z(\omega)$ whose input is $I_{RZ}(t)$. As a consequence the cross-spectrum of $V_{RZ}(t)$ and $I_{RZ}(t)$ will be $S_{VI}(\omega)=S_{II}(\omega)Z(\omega)$. On the other hand $I_{RZ}(t)$ is the output of a linar filter whose frequency response is $(Z(\omega)+R)^{-1}$ and the input of which is $V_R(t)$. so that $S_{II}=\frac{2kTR}{|Z(\omega)+R|^2}$. The mean power

$$P_{R\rightarrow Z}=R_{VI}(0)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{2kTR}{|Z(\omega)+R|^2}Z(\omega)d\omega=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{2kTR}{|Z(\omega)+R|^2}\mathrm{Re}\{Z(\omega)\}d\omega,$$

where I have used the fact that $\mathrm{Im}\{Z(\omega)\}$ is an odd function of ω and does not contribute to the integral.

The resistors present in the network will produce a random voltage that can be tought as due to a voltage generator in parallel to $Z(\omega)$ with spectral density $S_{zz}(\omega)$. This generator will dissipate power into the resistor. The power can be calculated as above interchanging the role of the resistor and the impedance.One than gets$P_{R\rightarrow Z}=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{S_{zz}(\omega)}{|Z(\omega)+R|^2}Rd\omega$.

At thermal equilibrium $P_{R\rightarrow Z}=P_{Z\rightarrow R}$ and then

$$S_{zz}(\omega)=2kT\mathrm{Re}\{Z(\omega)\} \qquad 103$$

The fundamental result in eq. 103 can be translated in term of current noise assuming a current generator in parallel with $Z(\omega)$ and with spectral density $S(\omega)=2kTRe\{Z(\omega)\}/|Z(\omega)|^2=2kTRe\{Y(\omega)\}$ with $Y(\omega)$ the admittance of the port.

The Fourier transform of $S_{zz}(\omega)$, the autocorrelation of the noise $R_{zz}(\tau)$, is the Fourier transform of $kT2Re\{Z(\omega)\}=kT[Z(\omega)+Z^*(\omega)]$. Calling $Z(\tau)$ the transform of $Z(\omega)$ then the autocorrelation of the noise is $R(\tau)=kT[Z(\tau)+Z(-\tau)]$. Because linear devices in electromagnetic network are causal $Z(\tau)$ vanishes for $\tau<0$ so that the autocorrelation is $R(\tau)=kTZ(|\tau|)$. $Z(\omega)$ can in addition be considered a Laplace transform taken for $p=i\omega$. Then eq. 39c tells that $\lim_{\omega\to\infty} i\omega Z(\omega)=Z(t=0)$. This result can be used to calculate the mean square fluctuation of the voltage across the impedance as

$$<\mathbf{V}_z{}^2(\omega)>=\frac{kT}{C} \qquad 104$$

where we have defined $\frac{1}{C}=\lim_{\omega\to\infty} i\omega Z(\omega)$. A similar theorem holds for currents substituting admittances to impedances and inductances to capacitances.

Narrow band noise

Consider the process

$$\mathbf{x}(t)=\mathbf{a}(t)\cos(\omega_o t)-\mathbf{b}(t)\sin(\omega_o t) \qquad 105$$

where $\mathbf{a}(t)$ and $\mathbf{b}(t)$ are two zero mean joint stationary stochastic processes and ω_o is a number. The mean value of the process is zero and the autocorrelation is:

$$R_{xx}(t,t')=R_{aa}(\tau)\cos(\omega_o t)\cos(\omega_o t')+R_{bb}(\tau)\sin(\omega_o t)\sin(\omega_o t')- R_{ab}(\tau)\cos(\omega_o t)\sin(\omega_o t')-R_{ab}(-\tau)\sin(\omega_o t)\cos(\omega_o t') \qquad 106$$

Thus the process $\mathbf{x}(t)$ can be wide sense stationary only if $R_{aa}(\tau)=R_{bb}(\tau)$ and $R_{ab}(\tau)=-R_{ab}(-\tau)$. If this is the case then the autocorrelation of $\mathbf{x}(t)$ is

$$R_{xx}(\tau)=R_{aa}(\tau)\cos(\omega_o\tau)+R_{ab}(\tau)\sin(\omega_o\tau) \qquad 107$$

and the spectrum is

$$S_{xx}(\omega)=\frac{1}{2}[S_{aa}(\omega-\omega_o)+S_{aa}(\omega+\omega_o)] - -\frac{1}{2i}[S_{ab}(\omega-\omega_o)-S_{ab}(\omega+\omega_o)] \qquad 108$$

If the processes $\mathbf{a}(t)$ and $\mathbf{b}(t)$ have spectra limited to frequencies much less than ω_o, then $\mathbf{x}(t)$ will have a spectrum made of two lines around $\pm\,\omega_o$ and will then consist basically of a sinusoid at frequency ω_o of random phase and amplitude.

It turns out that any stationary process can be represented as in eq. 105. To show this assume that a process $\mathbf{x}(t)$ is given by eq. 105 and that a second process $\mathbf{y}(t)$ is given by

$$\mathbf{y}(t)=\mathbf{a}(t)\sin(\omega_o t)+\mathbf{b}(t)\cos(\omega_o t) \qquad 105a$$

the autocorrelation of this second process will be

$$R_{yy}(\tau)=R_{aa}(\tau)\cos(\omega_o\tau)+R_{ab}(\tau)\sin(\omega_o\tau) \qquad 107a$$

while the crosscorrelation with $\mathbf{x}(t)$ is

$$R_{xy}(\tau)=-R_{aa}(\tau)\sin(\omega_o\tau)+R_{ab}(\tau)\cos(\omega_o\tau) \qquad 108$$

Eq. 107 and 108 can be seen as an orthogonal transformation of $R_{aa}(\tau)$and $R_{ab}(\tau)$ to $R_{xx}(\tau)$ and $R_{xy}(\tau)$. The transformation can be inverted so that

$$R_{aa}(\tau)=R_{xx}(\tau)\cos(\omega_o\tau)-R_{xy}(\tau)\sin(\omega_o\tau) \qquad 109a$$

$$R_{ab}(\tau)=R_{xx}(\tau)\sin(\omega_o\tau)+R_{xy}(\tau)\cos(\omega_o\tau) \qquad 109a$$

Thus, whatever the choice of the odd function $R_{xy}(\tau)$, both $R_{aa}(\tau)$ and $R_{bb}(\tau)$ can be calculated. If one, for instance, takes $R_{xy}(\tau)=0$ then

$R_{aa}(\tau)=R_{xx}(\tau)\cos(\omega_o\tau)$ and $R_{ab}(\tau)=R_{xx}(\tau)\sin(\omega_o\tau)$. In this last case evidently

$$S_{aa}(\omega)=S_{bb}(\omega)=\frac{1}{2}[S_{xx}(\omega-\omega_o)+S_{xx}(\omega+\omega_o)] \qquad 110a$$

$$S_{ab}(\omega)=\frac{1}{2i}[S_{xx}(\omega-\omega_o)-S_{xx}(\omega+\omega_o)] \qquad 110b$$

As an example consider the thermal noise of the harmonic oscillator in eq. 99. If one takes

$$S_{aa}(\omega)=\left\{\frac{k_BT}{m\tau\omega_o^2}\right\}\cdot\left\{\frac{1}{\omega^2+1/4\tau^2}\right\} \qquad 111a$$

and

$$S_{ab}(\omega)=\left\{\frac{k_BT}{m\tau\omega_o^2}\right\}\cdot\left\{\frac{-i\omega/\omega_1}{\omega^2+1/4\tau^2}\right\} \qquad 111b$$

then $\mathbf{x}(t)$ can be written as $\mathbf{x}(t)=\mathbf{a}(t)\cos(\omega_1 t)-\mathbf{b}(t)\sin(\omega_1 t)$. Thus the process consists of two sinusoids shifted by $\pi/2$ the amplitudes of which are two stationary random processes. $R_{aa}(t)$ and $R_{ab}(t)$ are given by:

$$R_{aa}(t)=\left\{\frac{k_BT}{m\omega_o^2}\right\}\cdot e^{-|t|/2\tau} \qquad 112a$$

and

$$R_{ab}(t)=\left\{\frac{k_BT}{m\omega_o^2}\right\}\cdot\left\{\frac{1}{2\tau\omega_1}\right\}e^{-|t|/2\tau}[2\theta(t)-1] \qquad 112b$$

Because $R_{ab}(0)$ is zero the two random variables $\mathbf{a}(t)$ and $\mathbf{b}(t)$ taken at the same time are independent. This is not true for the processes themselves that become independent only in the limit of high merit factor $\omega_1\tau \rightarrow \infty$.

1/f Noise

As a last exemple of a noise process which has great relevance in applications let discuss the so called 1/f noise. Historically this kind of process, that plagues any physical instrumentation, was first studied with much effort in metal film resistors and in semiconductors. In this systems the process is a voltage noise that only shows up when the device is biased by a current I. The measured spectral density of this voltage is well represented by:

$$Sv(\omega)=\frac{I^2R^2\alpha 2\pi}{N_c|\omega|^\delta} \qquad 113$$

where R is the resistance, N_c is the charge carrier total number and α is a constant that for metallic film resistor is of the order of 10^{-2}. The exponent δ is always a number close to 1. At least in some case it has been shown that the noise comes from equilibrium fluctuations of the resistivity.

Fluctuating physical quantities showing 1/f spectra have successively been shown to exist in many very different physical situations. Also thermal noise with 1/f spectra has been measured in quite a few magnetic systems[?].

To discuss the physics of the 1/f noise goes beyond the scope of these notes. I want to elaborate here only on some features of this kind of noise which are relevant to the signal detection theory.

A spectrum like that in eq. 113 can not even have a Fourier transform without some cutoff at both high and low frequencies. To provide these cutoffs in a natural way let consider that in many physical systems the 1/f noise is though to come from the incoherent superposition of the fluctuations of many independent subsystems each one showing an autocorrelation with a simple exponential relaxation. Though this model could not be of general validity, it is useful to provide a well behaved mathematical description. Suppose that the noise is then the sum of a series of stochastic processes each one with autocorrelation

$$R(t)=x^2e^{-|t|/\tau} \qquad 114$$

and with spectrum

$$S(\omega)=\frac{2x^2\tau}{1+(\omega\tau)^2} \qquad 115$$

Assume also that the time constants τ have a distribution $f(\tau)=\dfrac{\frac{1}{\tau}}{\ln(\tau_1/\tau_2)}$ for $\tau_1 \leq \tau \leq \tau_2$ and zero elsewhere. This distribution corresponds to a uniform distribution of the logarithm of the time constants. The total noise spectrum will be:

$$S(\omega)=\frac{2x^2}{\ln(\tau_2/\tau_1)}\int_{\tau_1}^{\tau_2}\frac{1}{1+(\omega\tau)^2}d\tau =$$

$$= \frac{2x^2}{\ln(\tau_2/\tau_1)}\frac{1}{|\omega|}\cdot[\mathrm{atn}(\omega\tau_2)-\mathrm{atn}(\omega\tau_1)]\approx\frac{\pi}{2}\frac{2x^2}{\ln(\tau_2/\tau_1)}\frac{1}{|\omega|} \qquad 116$$

where we have assumed $\omega\tau_1<<1<<\omega\tau_2$.

The autocorrelation of such a system is:

R(τ)=\F(x2,ln(τ2/τ1))\I(τ1,τ2,\F(e\S\UP10(-|t|/τ),τ)dτ) =\O(\S\UP18(=x2 for t=0),\S\DO18(≈ \F(x2,ln(τ2/τ1))[ln(τ2/|t|)- 0.5772] for τ1<< |t| <<τ2))117

Notice that if $\tau_2 >>|t|$ the residual logarithmic dependence on $|t|$ is just a small correction to the overall autocorrelation

An interesting quantity that gives an approximate idea of the dynamic of a process is the mean square variation on a time T defined as

$$\Delta^2x(T)=<[x(t+T)-x(t)]^2> = 2R(0)-2R(T) \qquad 118$$

For our 1/f noise

$$\Delta^2x(T) \approx \frac{2x^2}{\ln(\tau_2/\tau_1)}\cdot[\ln(T/\tau_1)+0.5772] \qquad 119$$

so that the process displays only a slight logarthmic dependence on T and its variations are substantially independent of the time scale of the observation.

Noise in linear two ports

I will now briefly discuss one example of multiple inputs and outputs device which is of great importance for the following. The device is a two port linear element that, for sake of clarity, I will assume to be an electrical device. With that I mean that inputs and outputs will consist of voltages and currents. In the frequency domain the relation between voltages and currents for a noiseless two port is

$$V_1(\omega)=Z_{11}(\omega)I_1(\omega)+Z_{12}(\omega)I_2(\omega)$$
$$V_2(\omega)=Z_{21}(\omega)I_1(\omega)+Z_{22}(\omega)I_2(\omega) \qquad 120$$

where the symbols refer to fig. 10. Eq. 120 is just one of the possible representation as the role of the inputs and of tyhe outputs can be interchanged.

The effect of the noise can be as usual represented by means of the two voltage noise generators (fig. 10a) producing the noise voltages $\mathbf{V}_{n1}(t)$ and $\mathbf{V}_{n2}(t)$. The total input and output voltages will be given by the sum of the noiseless signal part calculated from eq. 120 and of the respective noise voltage. The two voltage noise generator need not to be independent or uncorrelated as they can origin from the same physical mechanism inside the two port device.

An equivalent, and more commonly used, way to describe the noise of the two port is in term of noise generators only in the input port (here the port 1). In this case,

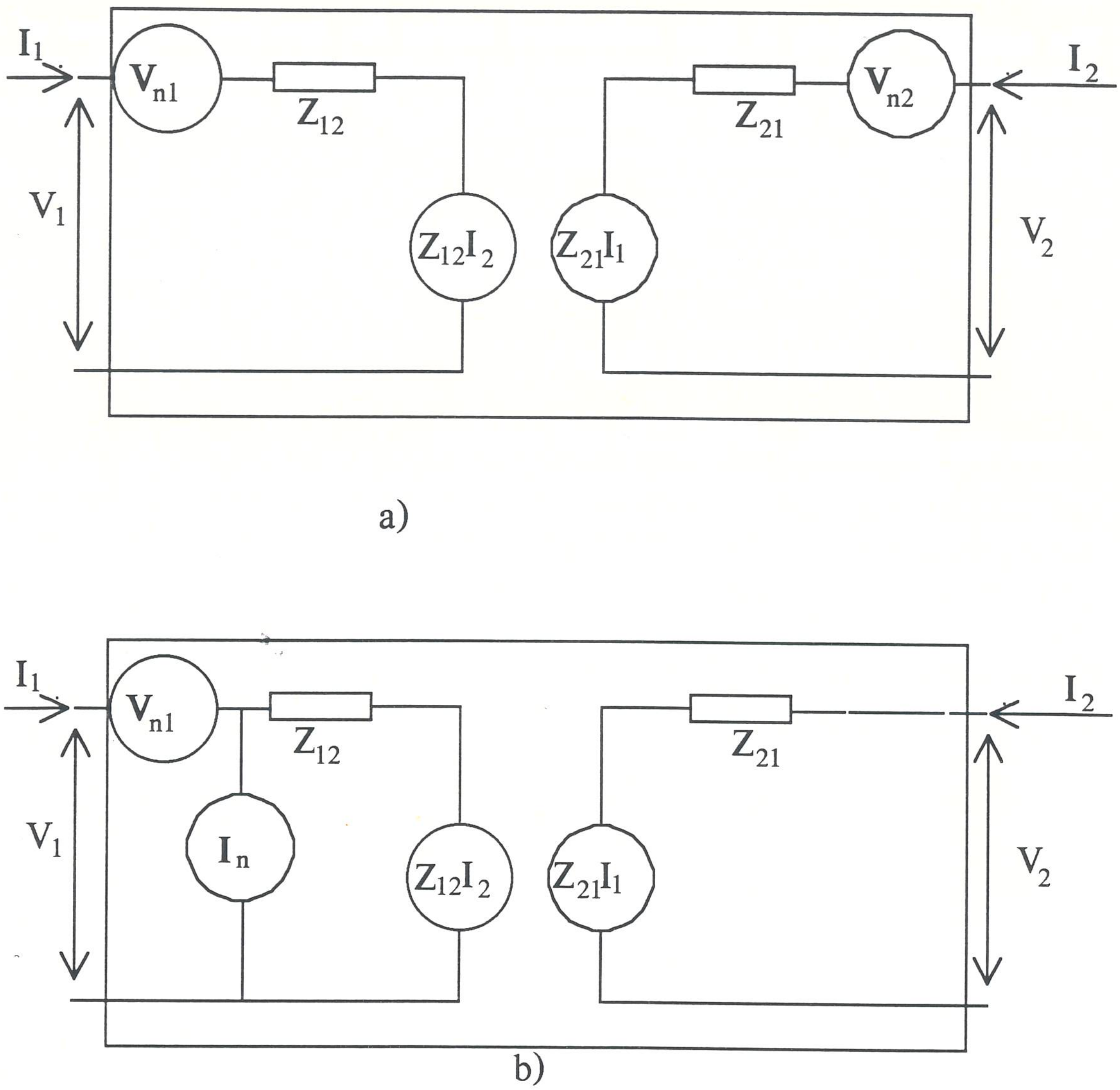

Fig 10

using the ordinary circuit theory, the two voltage generator of fig. 10a can be converted to a current generator in parallel to the input port and to a voltage generator in series to it (fig 10b).

Let call $S_V(\omega)$ and $S_I(\omega)$ the spectral densities of the voltage and current noise respectively. The two generators again need not to be uncorrelated so that also a cross-spectrum $S_{VI}(\omega)$ has to be introduced. Suppose that a voltage source with impedance $Z_s(\omega)$ is connected to the input port and that a load of impedance $Z_l(\omega)$ is connected to the output.. The output voltage noise due to the noise sources can be calculated, in the frequency domain, treating the noise processes as ordinary signals and calculating, by means of ordinary circuit theory, the frequency response that links the output to the corresponding current and voltage generators.

The spectral density is then calculated taking the spectral densities of the noise generators and multplying them by the square modulus of the relative transfer functions already calculated. Using this rule one first calculate , from elementary circuit theory, that:

$$V_2(\omega)=\frac{Z_l(\omega)[Z_s(\omega)I_n(\omega)-V_n(\omega)]}{[Z_{11}(\omega)+Z_s(\omega)][Z_{22}(\omega)+Z_l(\omega)] - Z_{12}(\omega)Z_{21}(\omega)} = \qquad 121$$

$$=G(\omega)[Z_s(\omega)I_n(\omega)-V_n(\omega)]$$

where the definition of $G(\omega)$ is self evident.

The output spectral density $S_{V2}(\omega)$ is then thus $|G(\omega)|^2$ times the spectral density of the linear combination $V'(t)=I'_n(t)+V_n(t)$, where $I'_n(t)$ is the output of a system with transfer function $Z_s(\omega)$ and input $I_n(t)$. As a consequence $S_{V'}(\omega)=S_V(\omega)+S_{I'}(\omega)+2Re\{S_{VI'}(\omega)\}$ = $= S_V(\omega)+|Z_s(\omega)|^2S_I(\omega)+2Re\{Z_s(\omega)S_{VI}(\omega)\}$. Thus the spectral density of the output is

$$S_{v2}(\omega) = |G(\omega)|^2 \left(S_V(\omega)+|Z_s(\omega)|^2S_I(\omega)+2Re\{Z_s(\omega)S_{VI}(\omega)\}\right)$$

122

It is often convenient, for reason that will be clear later on, to separate the current noise as $I_n(t)=I_{no}(t)+Y(t)*V_n(t)$. Here $I_{no}(t)$ is a process uncorrelated with $V_n(t)$, $Y(t)$ is the impulse response of a properly selected admittance, called the correlation admittance, and the star denotes the convolution operation.. $Y(\omega)$ has to be chosen such that $S_{VI}(\omega)=Y(\omega)S_V(\omega)$ and is then $Y(\omega)=S_{VI}(\omega)/S_v(\omega)$. As a consequence $S_I(\omega)=S_{Io}(\omega)+|Y(\omega)|^2S_v(\omega)=S_{Io}(\omega)+|S_{VI}(\omega)|^2/S_V(\omega)$.

It is also convenient to express the spectral densities as

$$S_{Io}(\omega)=\frac{k_BT_n(\omega)}{R_n(\omega)} \qquad 123a$$

$$S_V(\omega)=k_BT_n(\omega)R_n(\omega) \qquad 123b$$

$T_n(\omega)$ and $R_n(\omega)$ are called respectively the noise temperature and the noise resistance of the device.

With all these definition the noise at the output port can be written as:

$$S_{v^2}(\omega) = |G(\omega)|^2 \cdot \left[k_B T_n(\omega) \left\{ R_n(\omega)|1+Y(\omega)Z_s(\omega)|^2 + \frac{|Z_s(\omega)|^2}{R_n(\omega)} \right\} \right] \quad 124$$

The signal $V_s(\omega)$ at the input port gives origin at the output to a signal with Fourier transform $V_o(\omega)=G(\omega)V_s(\omega)$. The spectral density in the square brackets of eq. 124 can be tought then as that of an equivalent input voltage noise.

In many applications, as the noise in eq.124 depends on the source parameters, it is convenient to express the noise performance of the two port device by the ratio between the total noise at the output and the noise due to the source itself. This ratio is called the noise figure of the device. If the source is just adding his thermal noise $S_s(\omega)=2k_BTRe\{Z(\omega)\}$, then the noise at the output will be the incoherent sum of this term and that in eq. 124. The noise figure will then be:

$$F = 1 + \frac{\left[k_B T_n(\omega) \left\{ R_n(\omega)|1+Y(\omega)Z_s(\omega)|^2 + \frac{|Z_s(\omega)|^2}{R_n(\omega)} \right\} \right]}{2k_BTRe\{Z(\omega)\}} \quad 125$$

If $Y(\omega)=0$ eq. 125 greatly simplifies and has an easily calculable minimum when $Z_s(\omega)=R_n(\omega)$. In this case $F = 1+\frac{T_n(\omega)}{T}$ and the noise density due to the two port device is $2k_BT_n(\omega)R_n(\omega)$. When this noise minimum is attained, the source is called to be noise matched to the two port device.

Estimation of Noise Parameters

As already stated in the introduction, a crucial step of the calibration of an apparatus is the estimation of its noise parameters. A full calibration would imply the estimate of the joint probability densities of any order of all the random processes one has to deal with. This is not only impossible in practice but even useless as much of the information is contained in the lower moments or in the low order probability densities. Here I will discuss the estimation of the mean, of the mean square deviation and of the power spectrum. I will not treat the estimate of the autocorrelation as the equivalent power spectrum estimate is by far much more widely used and practical.

Estimate of the mean of a random process.

The estimate of the mean value of a random process coincides in many case with the estraction of the signal itself from the zero mean fluctuating part of the process. As a consequence the following sections will deal with this subject in much more of a detail. Here I only briefly discuss the estimate of the constant mean value η of a stationary process as a mean to illustrate the significance of the statistical estimate itself and some properties of the time integral of a stationary process. Here, and for the rest of the section, it is assumed that to estimate the various parameters of the process, mean, spectra etc., one can only use a single sample of the process known for $0 \leq t \leq T$.

As a first estimate of η one can consider the value of the process $\mathbf{r}(t)$ itself for a given value of t. If the process is normal then

$$\eta=\mathbf{r}(t)\pm\sqrt{R(0)} \qquad 127$$

with 0.68% probability according to the standard error formula for normal random variables.

To improve the precision of the estimate one has to resort to some form of average. As the ensemble average is not available let consider the result of time averaging. Take the random variable

$$\mathbf{I}=\frac{1}{T}\int_0^T \mathbf{r}(t)dt \qquad 128$$

it is easy to calculate that $<I>=<r>=\eta$ so that **I** fluctuates around η and is thus a possible estimate of this parameter. To evaluate the variance of **I** consider that:

$$\sigma_I^2=\frac{1}{T^2}\int_0^T\int_0^T <r(t)r(t')>dtdt'-<r>^2 =$$

129

$$=\frac{1}{T^2}\int_{-T}^{T} d\tau C_r(\tau)\int_0^{T-|\tau|} d\tau' = \frac{1}{T}\int_{-T}^{T} C_r(\tau)(1-\frac{|\tau|}{T})d\tau$$

where I have changed from the variable t and t' to τ=t-t' and τ'=t'.

The calculation cannot proceed any further without adding some information. Let then consider some interesting case. Let the noise be white up to a certain angular frequency $\omega=1/\tau^*$. The autocovariance will be $R(\tau)=R(0)e^{-\tau/\tau^*}$. Eq. 129 gives then

$$\sigma^2=2R(0)\frac{\tau^*}{T}\left[1+\frac{\tau^*}{T}\left(e^{-(T/\tau^*)}-1\right)\right]. \qquad 130$$

If $T>>\tau^*$ then

$$\sigma=\sqrt{\frac{2R(0)\tau^*}{T}} \qquad 131$$

which is reduced in respect to the non averaged case by the factor $\sqrt{\frac{2\tau^*}{T}}$. If $T<<\tau^*$ then $\sigma\rightarrow R(0)$ and the zero integration case is obviously recovered. Notice that, for the chosen autocorrelation function, $2R(0)\tau^*=S(0)$, where $S(\omega)$ is here the spectrum of the zero mean process $r(t)-\eta$, so that, for large T,

$$\sigma^2=\frac{S(0)}{T} \qquad 132$$

which is just the total energy contained in the spactrum in the frequency band $\omega=0\pm1/2T$. This last result is of more general relevance as can be readily checked expressing the autocovariance in term of its spectrum in eq. 129

$$\sigma^2=\frac{1}{2\pi}\cdot\int_{-\infty}^{\infty} S(\omega)\left(\frac{\sin(\omega T/2)}{\omega T/2}\right)^2 d\omega. \qquad 133$$

and considering that $(\sin x/x)^2$ is a function with a major lobe in the interval $0\pm\pi$.

If the autocorrelation is that of the 1/f noise in eq. 117, after some calculation one gets $\sigma^2\approx\frac{x^2}{\ln(\tau_2/\tau_1)}[\ln(\tau_2/T)-.577]$ that, for $\tau_2 >> T$, has no significant improvement with T. A commonly occurring situation is one where both an almost white noise, like that of the exemple above, and a 1/f noise are present. The 1/f spectrum $S(\omega)= S_o/|\omega|$ will merge in the wide band $P/(1+\omega^2\tau^2)$ one when $\omega_{1/f}=S_o/P$, if this frequency is $\omega_{1/f} << 1/\tau$.. The frequency $\nu_{1/f}=\frac{\omega_{1/f}}{2\pi}$ is usually called the 1/f corner frequency. In this case the variance of the integral is the sum of the two terms $\pi S_o[\ln(\tau_2/T)-.577]$ and P/T. Thus there is no advantage in integrating beyond the point where the second term becomes of the order of the first contribution. This gives a $T_{max}\approx 1/\nu_{1/f}$.

Estimation of the mean square deviation R(0)

Suppose we want to estimate the value of $<\mathbf{x}^2>=R(0)$, where $\mathbf{x}(t)$ is a zero mean random process. Clearly $\mathbf{x}^2(t)$ will fluctuate around its mean value $<\mathbf{x}^2>$ so that a single value of $\mathbf{x}^2(t)$ for any value of t is an estimate of $<\mathbf{x}^2>$. To evaluate how precise is this estimate we need to know we need to estimate the mean fluctuation of $\mathbf{x}^2(t)$ itself. Thus we need to know

$$\sigma_{\mathbf{x}2} =\sqrt{<\mathbf{x}^4> - <\mathbf{x}^2>^2} \qquad 134$$

To calculate this quantity we need some more information. Let then assume that, as it is often the case, the process we are discussing is a normal one. Now for four zero mean normal random variable **x,y,z,w** the following theorem holds

$$<x,y,z,w>=<x,y><z,w>+<x,z><y,w>+<x,w><y,z> \qquad 135$$

then

$$\sigma_{x^2}=\sqrt{2}<x^2> \qquad 136$$

Eq. 136 states that the relative error with which $<x^2>$ is measured is $\sqrt{2}$.

To improve the precision let average over the time and use eq. 129. Now, in the case of the process $x^2(t)$, the mean value is R(0) and the autocorrelation is

$$R_{x^2}(\tau)=<x(t+\tau)x(t+\tau)x(t)x(t)>=R(0)^2+2R^2(\tau) \qquad 137$$

where we have used eq. 135. Thus the variance of

$$\tilde{x}^2=\frac{1}{T}\int_0^T x^2(t)dt \qquad 138$$

is

$$\sigma^2=\frac{2}{T}\int_{-T}^{T} R^2(\tau)\left(1-\frac{|\tau|}{T}\right)d\tau \qquad 139.$$

If the autocorrelation is the exponentially decaying one of the previous section than $\sigma^2=4R^2(0)\frac{\tau^*}{2T}\left[1+\frac{\tau^*}{2T}\left(e^{-(2T/\tau^*)}-1\right)\right]$. If $T>>\tau^*$ then $\sigma=2R(0)\sqrt{\frac{\tau^*}{2T}}$ with a relative error $\sigma/R(0)=\sqrt{\frac{2\tau^*}{T}}$ which is reduced in respect to the non averaged (T→0) case by the factor $\sqrt{\frac{\tau^*}{T}}$ Notice that, if $T >> \tau^*$ then $\widetilde{x^2}$ can be considered as a linear combination of an high number of independent random variables and its distribution will be well approximated again by a normal one. As a consequence on can conclude that $R(0) \approx \widetilde{x^2} \cdot \left[1 \pm \sqrt{\frac{2\tau^*}{T}}\right]$ with 0.68% probability according to the standard gaussian error formula.

Power spectrum estimation.

Consider the scheme in Fig. 11

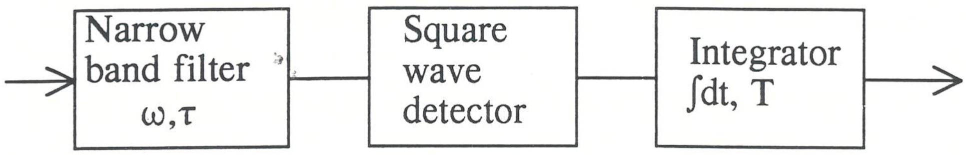

Fig 11

as a narrow band filter we can consider that of example 3.5. The square wave detector just performs the square of the input so that the mean value of the detector output is, according to eq. 93 $\langle x^2 \rangle = S_i(\omega_o)/2\tau$. Here $S_i(\omega)$ is the spectral density of the zero mean process $i(t)$ which is the input to the entire system, ω_o is the center frequency of the filter and τ is its time constant . We have also assumed that $S_i(\omega)$ varies slowly inside the band $\omega_o \pm 1/\tau$ so that it can be considered constant within this band.

The autocorrelation of the narrowband filter output is given by:

$$R(t)=\frac{S_i(\omega_o)}{2\tau} e^{-|t|/2\tau}\left[\cos(\omega_1 t)+\frac{1}{2\omega_1\tau}\sin(\omega_1|t|)\right] \qquad 140$$

and the autocovariance of the square wave detector

$$C_{x^2}(t)=2R^2(t)= \qquad 141$$

$$=\left(\frac{S_i(\omega_o)}{2\tau}\right)^2 e^{-|t|/\tau}\left[\left(1+\frac{1}{(2\omega_1\tau)^2}\right)+\left(1-\frac{1}{(2\omega_1\tau)^2}\right)\cos(2\omega_1 t)+\frac{1}{\omega_1\tau}\sin(2\omega_1|t|)\right]$$

Thus the output of the square wave detectors fluctuates around $S_i(\omega_o)/2\tau$, and is then an estimate of the power spectrum of $i(t)$ at frequency ω_o, but its relative uncertainty, $\sqrt{C_x2(0)}/R(0)$ is, as in the previous section, equal to $\sqrt{2}$.

Again, in order to reduce the uncertainty, the output of the square wave detector is integrated on a time T. To estimate the resulting uncertainty one has to substitute the correlation in eq. 141 in the first half of eq. 139 . If $T \gg \omega_1$ the oscillating terms in eq. 141 average away and the first term gives a result identical to that in eq. 139 provided that τ is substituted to τ^*. The final root mean square fluctuation of the integrator output , for $T \gg \tau$, is then $\sigma = \dfrac{S_i(\omega_o)}{\sqrt{2\tau T}}$ and the relative uncertainty on $S_i(\omega)$, $\dfrac{\sigma_S}{S} = \sigma / [S_i(\omega_o)/2\tau]$ is

$$\frac{\sigma_S}{S} = \sqrt{\frac{2\tau}{T}} \qquad 142$$

a fundamental result in spectral analysis. Notice that the square modulus of the frequency response of the narrow band filters reduces to 1/2 of its maximum amplitude at the two frequencies $\omega_\pm = \omega_o \pm 1/2\tau$ so that the bandwidth of the filter is just $\Delta\omega = 1/2\tau$. Thus eq. 142 can be restated as

$$\frac{\sigma_S}{S} = \sqrt{\frac{1}{\Delta\omega T}} \qquad 143$$

Notice that the integration <u>before</u> the square wave detector sets the bandwidth of the measurement. Increasing τ decreases the precision of the mesurement. It is only the the time constant T <u>after</u> the square wave detector that sets the final precision. Notice also that obviously $T \gg \tau$ so that high resolution measurements need high integration times.

Exemple 4.1. A "signal" consist in a random noise with a narrow line spectrum of amplitude S_o centered at the frequency ω_o and with linewidth $\delta\omega$. The signal is buried in a white noise of density P. If the spectrum of the signal has to be measured with a resolution $\Delta\omega = \delta\omega/10$ and with a 10% relative precision, the needed integration time can be calculated from

$$\frac{\sigma S}{S_o} = \frac{S_o+P}{S_o\sqrt{\frac{\delta\omega T}{10}}} = \frac{1+(P/S_o)}{\sqrt{\frac{\delta\omega T}{10}}} \qquad 144$$

If $P/S_o>>1$, noise dominated case, then $T=1000(P/S_o)^2/\delta\omega$.

A spectrum analyzer can thus in principle be built by a battery of narrowband filters followed by proper square wave detector and integrator stages. In practice this is done only for high frequencies, $\nu >$ MHz, where, however, the spectrum of the signal is shifted in frequency, by multiplying it by a proper sinusoidal carrier, so to bring it in the frequency region where the filters are available. The implications of this technique, called heterodyne analysis, will be briefly discussed later on.

For low frequency signals the most diffused technique is to digitize the random process, by sampling it at fixed time intevals, and by calculating a Fast Fourier Transform of the resulting samples. I will now discuss this method and its limitations.

Consider the Discrete Fourier Transform (DFT) $\mathbf{x}(k)$ of the N samples $\mathbf{x}(n)$ of a zero mean process taken at fixed intervals of time nT: $\mathbf{x}(k)=\frac{1}{N}\sum_{n=0}^{N-1}\mathbf{x}(n)e^{-i(2\pi/N)kn}$. I will show that $\mathbf{S}(k)=TN|\mathbf{x}(k)|^2$ is an estimate of the spectral density of the process at the frequency $\omega=k2\pi/(NT)$. In fact

$$<\mathbf{S}(k)>=\frac{1}{N}\sum_{n,m=0}^{N-1}<\mathbf{x}(n)\mathbf{x}(m)>e^{-i(2\pi/N)k(n-m)}= \qquad 145$$

$$=\frac{T}{N}\sum_{n,m=0}^{N-1}R[(n-m)T]e^{-i(2\pi/N)k(n-m)}.$$

Now the autocorrelation R(t) can be expressed as a function of the power spectrum to give

$$<S(k)>=\frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)\left|\frac{T}{N}\sum_{n=0}^{N-1} e^{-in[\omega T-k(2\pi/N)]}\right|^2 d\omega.=$$

146

$$=\frac{T}{2\pi}\int_{-\infty}^{\infty} S(\omega)\,\left|f(\omega T-k2\pi/N)\right|^2 d\omega$$

where I have defined $f(x)=\frac{1}{\sqrt{N}}\frac{1-e^{iNx}}{1-e^{ix}}$. The function f(x) is periodic of period 2π so that eq. 146 can be recast as

$$(k)>=<S \quad \frac{T}{2\pi}\int_{-\pi/T}^{\pi/T}\sum_{n=-\infty}^{\infty} S(\omega+n2\pi/T)\left|f(\omega T-k2\pi/N)\right|^2 d\omega \qquad 147$$

(k) is then an estimate of S(Sω=k2π/NT) but suffers of two sistematic errors: first there is a clear aliasing error due to the contribution of all the frequencies at $\omega+n2\pi T$. The error is suppressed if $S(\omega)=0$ for $\omega\geq\pi/T$. This is the sampling theorem applied to the power spectrum. In order to reduce the aliasing error it is necessary to use a low pass filter before sampling the signal. This can be accomplished by high order low pass filters (8 poles or higher) with roll-off frequency less than one half the sampling frequency. The high phase shifts introduced by the filter are unimportant as any phase information is lost in taking the square modulus. Notice that, because S(k)=S(N-k), only the first N/2 coefficients, up to $\omega=k2\pi/(NT)=2\pi N/(2NT)=\pi/T$, i.e. the filter frequency, have an independent meaning.

The second systematic error is due to the fact that the estimate involves the convolution of the spectrum with a function which has a central lobe of height N and width $\approx 1/N$ around $\omega=k2\pi/(NT)$. The resolution of the spectrum is then $\Delta\omega\approx 2\pi/NT=2\pi/T_{tot}$ with T_{tot} the duration of the measurement. To each coefficient S(k) also the nearby lobes give a contribution of the order $\approx 1/(2\pi n)^2$ with $n=\pm1,\pm2$ etc, the number of the lobe. To reduce the contribution of the side lobes the data can be multiplyied by a proper sequence of weights w(n). The function f(x) has to be substituted in this case by $f'(x)=\sum_{n=0}^{N-1} w(n)e^{-inx}$.

The evaluation of the variance of the spectrum is rather complicate . I just sketch the idea and discuss the main implication of the result. One first evaluate the value of $<\mathbf{S}(k)\mathbf{S}^*(k')>$which is given after a lenghty calculation by

$<\mathbf{S}(k)\mathbf{S}^*(k')>=$

$$\frac{T^2}{N^2}\sum_{n,m,j,p=0}^{N-1} <\mathbf{x}(n)\mathbf{x}(m)\mathbf{x}(j)\mathbf{x}(k)>e^{-i[(2\pi/N)k(n-m)-(2\pi/N)k'(j-p)]}= \qquad 148$$

$$=\left|\frac{T^2}{2\pi}\int_{-\pi/T}^{\pi/T} S(\omega)f(\omega T-k2\pi/N)f^*(\omega T-k'2\pi/N)d\omega\right|^2 +$$

$$+\left|\frac{T^2}{2\pi}\int_{-\pi/T}^{\pi/T} S(\omega)f(\omega T-k2\pi/N)f^*(\omega T+k'2\pi/N)d\omega\right|^2 + <\mathbf{S}(k)><\mathbf{S}^*(k')>$$

where I have assumed that the sampling condition has been fulfilled.

Taking k=k' one gets

$$<|\mathbf{S}(k)|^2>-<|\mathbf{S}(k)|>^2=<|\mathbf{S}(k)|>^2+ \qquad 149$$

$$+\left|\frac{T^2}{2\pi}\int_{-\pi/T}^{\pi/T} S(\omega)f(\omega T-k2\pi/N)f^*(\omega T+k'2\pi/N)d\omega\right|^2 \geq<|\mathbf{S}(k)|>^2$$

Again the estimate has a relative error of order one. In particular , if the spectrum is white, the relative precision is $\sqrt{2}$. This is of no surprise as The resolution is $\Delta\omega\approx 1/(NT)$ while the duration of the measurement is $T_{tot}=NT$ so that $\Delta\omega T_{tot}\approx 1$. In order to improve the precision one has to lower the resolution. Two method are used to acheive this result. Or the sequence of data is split in a set of M subsequences of

N'=N/M data each. The M resulting **S**(k) coefficients are then evaluated and averaged. If the coefficients coming from differents subsequencies can be considered as independent, then the average will have an error reduced by $\sqrt{M}$. The duration of each data set is now N'T=NT/M and thus the resolution is reduced by a factor 1/M so that eq. 143 is still obeyed.

The second way to reduce the resolution, spectral smoothing, takes instead the average or some linear combination of nearby coefficients so that the frequency resolution is lowered. If the nearby coefficients can be considered independent, which is the case only for white noise (see eq. 148) again the precision increases as $1/\sqrt{M}$, with M the number of coefficients that have been averaged, and the frequency resolutions decreases as 1/M.

Signal Recovery

I will begin this section discussing a couple of "non-optimal" signal recovery procedures that are commonly used and that give a broad feeling of what are the main issues of the signal recovering techniques. In the second part of the section I'll discuss the optimal signal recovery theory.

Periodic signal in additive noise.

Let consider the detection scheme where a periodic signal s(t) of unknown amplitude and phase, buried in an additive noise $\mathbf{n}(t)$, is detected by means of a lock-in amplifier. The total signal at the lock-in input is

$$\mathbf{x}(t)=s(t)+\mathbf{n}(t)=a(t)\sin(\omega_o t)-b(t)\cos(\omega_o t)+\mathbf{n}(t) \qquad 150$$

Both channels of the lock-in are linear non time invariant filters. The output o(t) will then be the sum of the contribution due to the signal $o_s(t)$ and of that due to the noise $\mathbf{o}(t)$. If the "sine" channel is considered, the signal contribution to the output will have, according to eq. 53, a Fourier transform $o_s(\omega)=\frac{a(\omega)}{2} \cdot \frac{1}{1+i\omega\tau'}$ with τ' the time constant of the low pass filter.

The stochastic process $\mathbf{o}(t)$ is no more stationary. By multiplying $\mathbf{n}(t)$ by $\sin(\omega_o t)$ the mixer on the sine channel produces the non stationary process $\mathbf{n}'(t)=\mathbf{n}(t)\sin(\omega_o t)$ with autocorrelation

$$R_{\mathbf{n}'}(t,t+\tau)=\frac{1}{2}[R_{\mathbf{n}}(\tau)\cos(\omega_o\tau)-$$
$$-R_{\mathbf{n}}(\tau)\cos(\omega_o\tau)\cos(2\omega_o t)+R_{\mathbf{n}}(\tau)\sin(\omega_o\tau)\sin(2\omega_o t)] \qquad 151$$

Before we proceed let us consider a few mathematical results that we will use in the next. Non stationary processes can be Fourier analyzed as stationary ones do. The Fourier spectrum is however now a function of two frequencies as the autocorrelation R(t,t+τ) is a function of two time arguments. Let then define the two dimensional Fourier spectrum:

$$\Delta(\omega_1,\omega_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}R(t,t+\tau)e^{-i\omega_1\tau+i\omega_2 t}\,dt\,d\tau \qquad 152$$

If the process is stationary then $\Delta(\omega_1,\omega_2)=2\pi S(\omega_1)\delta(\omega_2)$, where $S(\omega)$ is the ordinary power spectrum of the process.

If a process $\mathbf{x}(t)$ with bidimensional spectrum $\Delta_{xx}(\omega_1,\omega_2)$ is the input of a time invariant system with frequency response $h(\omega)$, the cross-spectrum between the output $\mathbf{y}(t)$ and the input can be calculated, applying the convolution theorem, to be

$$\Delta_{yx}(\omega_1,\omega_2)=h(\omega_1)\Delta_{xx}(\omega_1,\omega_2) \qquad 153$$

while

$$\Delta_{yy}(\omega_1,\omega_2)=h(\omega_1)h^*(\omega_1+\omega_2)\Delta_{xx}(\omega_1,\omega_2) \qquad 154$$

Going back to our lock-in, the output of the mixer $\mathbf{n}'(t)$ will have a two dimensional spectrum:

$$\Delta_{n'n'}(\omega_1,\omega_2)= 2\pi\frac{S(\omega_1-\omega_o)+S(\omega_1+\omega_o)}{4}\cdot\delta(\omega_2) \qquad 155$$

$$- 2\pi\frac{S(\omega_1-\omega_o)\cdot\delta(\omega_2-2\omega_o)}{4}+ 2\pi\frac{S(\omega_1+\omega_o)\cdot\delta(\omega_2+2\omega_o)}{4}$$

The low pass filter output $\mathbf{o}(t)$ will have a spectrum:

$$\Delta_{oo}(\omega_1,\omega_2)= 2\pi\frac{S(\omega_1-\omega_o)+S(\omega_1+\omega_o)}{2(1+\omega_1{}^2\tau'^2)}\cdot\delta(\omega_2) \qquad 155$$

$$- 2\pi\frac{S(\omega_1-\omega_o)\cdot\delta(\omega_2-2\omega_o)}{4(1+i\omega_1\tau')[1-i(\omega_1+2\omega_o)\tau']}+ 2\pi\frac{S(\omega_1+\omega_o)\cdot\delta(\omega_2+2\omega_o)}{4(1+i\omega_1\tau')[1-i(\omega_1-2\omega_o)\tau']}$$

The first term is just an ordinary spectrum of a stationary process. Because of the low pass filter, it is significantly different from zero only for $\omega_1<<1/\tau'$. Thus the noise associated to this term is contributed by a frequency band of width $\approx 1/\tau'$ around ω_o and has a spectrum

$$S_o(\omega)=\frac{S(\omega_1-\omega_o)+S(\omega_1+\omega_o)}{4(1+\omega_1^2\tau'^2)} \qquad 156$$

The terms in the second line of eq. 155 contain a denominator which is always larger than $2\sqrt{1+4\omega_o^2\tau'^2}$ and thus, if $\omega_o^2\tau'^2 >>1$, give contributions that are much smaller than that due to the first term for $\omega_1\tau'<<1$. In conclusion, with a proper choice of τ', the noise contribution to the output can be made to closely approximate a stationary process with spectrum given by eq. 156.

If the original noise spectrum $S(\omega)$ is reasonably flat over $\omega_o \pm 1/\tau$, than a single measurement of the output at time t will have a mean value $o_s(t)$ and an r.m.s. fluctuation $\sqrt{S(\omega_o)/4\tau}$. If $s(t)=a_o \sin(\omega_o t)$ than the output will be $a_o/2 \pm$ \R(S(ω_o)/4τ). so that the error on a_o is $\sigma_a=2\sigma_o=\sqrt{S(\omega_o)/\tau}$.

The most important thing to be noticed about this result is that the uncertainty on a_o is now contributed by the noise at frequency ω_o. As I mentioned in the section on systems, a constant or slowly varying signal that has to be measured can be often transformed to the amplitude modulation of a periodic carrier signal by a proper "chopping" mechanism. As a consequence, by properly chosing the frequency of the carrier one can work in the frequency region where the noise density of the apparatus is the minimum possible one. This is of great importance to overcome the 1/f noise corner frequency.

By using a second channel with a cosine drive the other quadrature component $o'(t)=o'_s(t)+o'(t)$ can be recovered as well. The noise part of this second component will have the same spectrum as the first one and a cross correlation spectrum with that given by:

$$S_{oo'}(\omega)=\frac{S(\omega_o+\omega_1)-S(\omega_o-\omega_1)}{4i(1+\omega_1^2\tau'^2)} \qquad 156$$

If again the spectrum is reasonably flat around ω_o, then $S(\omega_o+\omega_1)\approx S(\omega_o-\omega_1)$, the cross spectrum vanishes and the two channels are independent.

The estimate of the amplitude modulation of the quasiperiodic signal s(t) depends on the noise level. Let consider, for sake of clarity, the case where also the cosine component is almost constant, $b(t)\approx b_o$. The square of the amplitude $M^2=4[o^2(t)+o'^2(t)]$ can be expressed as

$M^2=a_o^2+b_o^2+4o^2(t)+4o'^2(t)+8a_oo(t)+8b_oo'(t)$. The mean value of M^2 is then

$$<M^2>=a_o^2+b_o^2+\sigma_a^2+\sigma_b^2=M_o^2+2\sigma^2 \qquad 157$$

where we have defined $\sigma=\sigma_a=\sigma_b$ and $M_o=\sqrt{a_o^2+b_o^2}$. The sum of the squares of the two components, as an estimate of the signal square amplitude, has then a systematic error in it that comes from the noise mean square fluctuation.

The mean square fluctuation of M^2 is:

$$\sigma^2{}_{M^2}=<M^4>-<M^2>^2=2\sigma_a^4+2\sigma_b^4+4a_o^2\sigma_a^2+4b_o^2\sigma_b^2=$$

$$=4\sigma^4+4M_o^2\sigma^2 \qquad 158$$

and, as anticipated, depends on the amplitude of the signal. In particular, if we define the postdetection signal to noise ratio S/N as

$$\frac{S}{N}=\frac{M_o}{\sigma} \qquad 159$$

then

$$\sigma^2{}_{M^2}=4\sigma^4\left[1+\left(\frac{S}{N}\right)^2\right] \qquad 160$$

If $S/N<<1$, then $\sigma_{M^2}=2\sigma^2$ and the error is of the order of the systematic error on M^2. If instead $S/N>>1$ than the systematic error becomes small and $\sigma_{M^2}=2M_o\sigma$. The amplitude modulation M can be measured with an error:

$$\sigma_M\approx(1/2M_o)\sigma_{M^2}=\sigma=\sqrt{S(\omega_o)/\tau} \qquad 161$$

It is interesting to evaluate, again in the limit of $S/N>>1$, also the error on the phase ϕ of the periodic signal relative to the local oscillator signal $\sin(\omega_o t)$. Let define $\phi=\text{artan}(b_o/a_o)$. Then, from the standard error propagation formula one gets:

$$\sigma_\phi\approx\sqrt{\frac{a_o^2\,\sigma^2}{(a_o^2+b_o^2)^2}+\frac{b_o^2\,\sigma^2}{(a_o^2+b_o^2)^2}}=\frac{\sigma}{M_o}=\frac{N}{S} \qquad 162$$

that is the phase error in radians is equal to the inverse of the signal to noise ratio.

An alternative method to extract the amplitude of a quasi periodic signal is to sample it and then evaluate its discrete Fourier transform. It can be calculated from the formulas of the preceding section that the basic results are the same as those obtained by the lock-in method provided the time constant τ' is substituted by NT, the total measuring time.

Charge pulse signal in additive noise.

Consider the circuit in Fig 12.

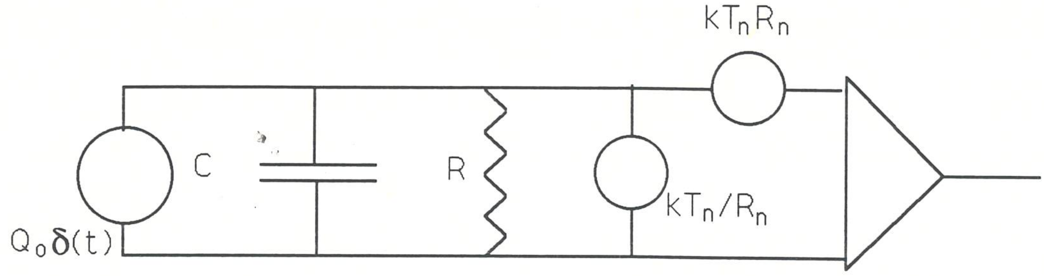

Fig 12

A current generator produces a current pulse $I(t)=Q_o\delta(t)$ that charges a capacitor of capacitance C. The voltage across the capacitor is measured by an amplifier with noise temperature T_n and noise resistance R_n. Both these parameters are assumed to be frequency independent. In parallel to the capacitor there is a resistor of resistance R. This resistor includes both the effect of the parallel losses of the capacitor and the input impedance of the amplifier. The thermal noise due to the resistor is included in the current generator spectral density $S_I=k_BT_n/R_n$.

This is an (oversimplified) model for a charge detector read by a charge amplifier. The problem I want to discuss here is how to measure the charge Q_o with minimum uncertainty. The method I will discuss is a non optimal method known as the RC-CR or zero-pole pulse shaping.

The current pulse gives a voltage signal at the input of the amplifier $V_s(t)=\frac{Q_o}{C}e^{-t/\tau_o}\Theta(t)$, with $\tau_o=RC$. In the absence of noise the charge can then be measured from the initial step rise of the amplifier output. In the presence of the noise this step will be hidden by noise fluctuations. For a truly white voltage noise the mean square root fluctuation is infinite and will cause infinite error on the step height estimation. If a low pass filter with time constant τ is used at the amplifier output the voltage noise fluctuation is reduced to $\sigma^2_V=k_BT_nR_n/\tau$ but the step rise is also lowered to $\frac{Q_o}{C}\, x^{\left(\frac{x}{1-x}\right)}$, with $x=\tau/\tau_o$, that tends to x^{-1} for large x. To compensate for this effect one can "derive" the signal using an high pass filter with frequency response $\frac{i\omega\tau'}{1+i\omega\tau'}$. The total filter frequency response will then result from this stage and the low pass and will be

$$h(\omega)=\frac{1}{1+i\omega\tau}\cdot\frac{i\omega\tau'}{1+i\omega\tau'} \qquad 163$$

The signal will have, after the filtering, a Fourier transform $V_s(\omega)=\frac{Q_o}{C}\frac{\tau_o}{1+i\omega\tau_o}h(\omega)$. Because of the high pass filter, the output will only be contributed by frequencies larger than $1/\tau'$. If the resistor R is high enough (low losses) that $\tau_o>>\tau'$ and $V_s(\omega)\approx\frac{Q_o}{C}\frac{1}{1+i\omega\tau}\cdot\frac{\tau'}{1+i\omega\tau'}$ the Fourier transform of which \F($Q_o\tau'$,C(τ'-τ)) \B\BC\((e\S\UP10(-t/τ') - $e^{-t/\tau}$) has a maximum for $t=\frac{\tau\tau'\ln(\tau'/\tau)}{\tau'-\tau}$ with value $V_{max}=\frac{Q_o}{C}\,y^{\frac{1}{1-y}}$ where $y=\tau'/\tau$.

The total voltage noise will have a spectrum

$$S(\omega)=k_BT_n\left[\frac{\tau'^2}{R_nC^2}\frac{1}{(1+\omega^2\tau'^2)(1+\omega^2\tau^2)}+\frac{R_n\omega^2\tau'^2}{(1+\omega^2\tau'^2)(1+\omega^2\tau^2)}\right]$$

164

where again we have assumed $\omega\tau_o>>1$.

The mean square fluctuation of the voltage σ^2 can be calculated with a little algebra and is given by

$$\sigma^2=k_BT_nR_n\left[\frac{\tau}{2R_n^2C^2}\,y+\frac{1}{2\tau}\right]\frac{y}{1+y} \qquad 165$$

Let again define a post detection signal to noise ratio as $S/N=V_{max}/\sigma$. Then:

$$\left(\frac{S}{N}\right)^2=\left(\frac{Q_o}{C}\right)^2\frac{1}{k_BT_nR_n}\frac{y^{\frac{1+y}{1-y}}(1+y)}{\left[\frac{\tau}{2R_n^2C^2}y+\frac{1}{2\tau}\right]} \qquad 166$$

The denominator in the right hand side of eq. 166 has a minimum when $\tau=R_nC/\sqrt{y}$ so that $(S/N)^2$ has a maximum with value

$$\left(\frac{S}{N}\right)^{*2}=\frac{Q_o^2}{C\,k_BT_n}\,y^{\frac{1+3y}{2(1-y)}}(1+y) \qquad 167$$

it is straightforward to calculate that $(S/N)^*$ as a maximum for y=1 with value

$$\left(\frac{S}{N}\right)^2_{max}=\frac{Q_o^2}{C\,k_BT_n}\,2e^{-2} \qquad 168$$

It is worth to recall that the maximum is achieved for

$$\tau=\tau'=R_nC \qquad 169$$

provided that $R>>R_n$.

Let discuss now the results in eq. 168 and 169. First notice that, defining Q_{omin} as the charge that can be measured with S/N=1, then the energy the signal with this value of Q_o would release to the capacitor is

$$E_o=\frac{Q^2_{omin}}{2C}\approx 1.85\,k_BT_n \qquad 170$$

This is not the uncertainty on the energy release due to the signal. In fact the output of the filter Q, properly corrected for all the unimportant multiplicative factors, is the sum of the signal Q_o and of the zero mean noise contribution Q_n with variance $Ck_BT_ne^2/2$. The energy can be estimated as

$$E=\frac{Q^2}{2C}=\frac{(Q_o+Q_n)^2}{2C} \qquad 171$$

and has mean value

$$<E> = \frac{Q_o^2}{2C} + \frac{k_B T_n e^2}{2}. \qquad 172$$

variance

$$\sigma^2{}_E = 2\left[\frac{k_B T_n e^2}{2}\right]^2 + 4\frac{Q_o^2}{2C}\cdot\frac{k_B T_n e^2}{2}. \qquad 173$$

and relative error

$$\frac{\sigma_E}{(Q_o^2/2C)} = \sqrt{2\left[\frac{k_B T_n e^2}{Q_o^2/C}\right]^2 + 4\left[\frac{k_B T_n e^2}{Q_o^2/C}\right]} \qquad 174$$

so that the minimum energy released that can be detected with signal to noise ratio of equal to one is

$$E_{min} \approx 2.2 k_B T_n e^2 \qquad 175$$

I anticipate that the technique being non optimal translates into the presence of the factor e^2 in eq. 175.

The second thing is worth noticing is that the assumption $R >> R_n$ means that the current noise contribution to the total voltage noise

$$S_c(\omega) = \frac{k_B T_n}{R_n} \frac{R^2}{1+\omega^2 \tau_0^2} \qquad 176$$

is, at low frequency, much larger then the voltage contribution $k_B T_n R_n$. This condition holds up to the frequency ω_{max} at which the two contributions become equal. ω_{max} is given by $\omega_{max} = (1/\tau_o)\sqrt{1+(R/R_n)^2} \approx 1/R_n C$.

Thus the idea of the filter is to take the information only at those frequencies for which the current noise, that has a spectral content equal to that of the signal, dominates on the voltage noise.

A further consideration which is worth mentioning is that the current noise contribution in eq. 166, increases with the filter time constant τ. This is due to the fact that the signal maximum is reached after a time of order τ. During this time the current noise makes the voltage drift by a root mean square quantity $\propto \sqrt{\tau}$. The minimum uncertainty is then achieved when the current term that increases with $\sqrt{\tau}$ becomes equal to the voltage noise term that decreases as $1/\sqrt{\tau}$.

A final consideration is that R results from the parallel of the loss resistor of the detector itself and of the input impedance of the amplifier which is supposed here to be real. The maximum of R is then obtained when these two resistors are equal. However, if $R \gg R_n$, this matching condition has not to be fulfilled with very high precision. We will elaborate a little more on this point later on.

The Wiener-Kolmogorov theory of optimal linear filtering.

Suppose that the ouput $\mathbf{x}(t)$ of some physical instrumentation is the sum of a signal $s(t, A_1, A_2, ...)$, that depends on time but also on a set of parameters A_i, and of a zero mean gaussian random noise $\mathbf{n}(t)$.

$$\mathbf{x}(t)=s(t, A_1, A_2, ...)+\mathbf{n}(t) \qquad 177$$

The values of the parameters are unknown and have to be estimated from the knowledge of the data $\mathbf{x}(t)$ in the time interval $T_1 \le t \le T_2$. The problem we want to discuss here can be stated as follows: is it possible to build a linear combination of the data

$$\mathbf{A}_i = \int_{T_1}^{T_2} h_i(t)\mathbf{x}(t)dt \qquad 178$$

such that

$$<\mathbf{A}_i> = A_i \qquad 179a$$

and that its variance

$$\sigma^2_{\mathbf{A}i} = \int_{T_1}^{T_2}\int_{T_1}^{T_2} h_i(t')h_i(t)<\mathbf{x}(t')\mathbf{x}(t)>dt'dt = \int_{T_1}^{T_2}\int_{T_1}^{T_2} h_i(t')h_i(t)R(t-t')dt'dt$$

179b

has the minimum possible value.

Amplitude estimation

Let start with simpler and most important case where the signal s(t) is s(t)=Af(t), with f(t) a known function of the time and A the unknown amplitude that has to be estimated. We will then build only one estimator **A** according to eq. 178.

Eq. 179a implies that

$$\int_{T_1}^{T_2} h(t)f(t)dt=1 \qquad 180$$

because the noise does not contribute to the mean value of **A**.

We now have to find the function h(t) that gives the minimum variance $\sigma^2{}_{\mathbf{A}}$ with the condition in eq. 180. This can be done using the method of the Lagrange multipliers and calculating the variation of

$$\int_{T_1}^{T_2}\int_{T_1}^{T_2} h_i(t')h_i(t)R(t-t')dt'dt + \lambda \int_{T_1}^{T_2} h(t)f(t)dt \qquad 181$$

with λ the multiplier that has to be determined a posteriori imposing the condition in eq. 180.

The variation gives

$$\int_{T_1}^{T_2} \delta h(t)\left[2\int_{T_1}^{T_2} h(t')R(t-t')dt'+\lambda f(t)\right]=0 \qquad 182$$

or

$$2\int_{T_1}^{T_2} h(t')R(t-t')dt'+\lambda f(t)=0 \text{ for } T_1 \le t \le T_2 \qquad 183$$

The minimum variance can be obtained using eq. 183 and eq. 180:

$$\sigma^2{}_{A\,min} = \int_{T_1}^{T_2}\int_{T_1}^{T_2} h_i(t')h_i(t)R(t-t')dt'dt = -\frac{\lambda}{2}\int_{T_1}^{T_2} h(t)\cdot f(t)\,dt = -\frac{\lambda}{2}$$

184

Eq. 183 can be solved in general by numerical integration or by any other standard method. I will discuss here the solution in two limiting cases which are very important in themselves and that also show the main features of this method.

White noise.

The first case is when $R(\tau)=S_o\delta(\tau)$, i.e. when the noise is white. Eq. 183 becomes in this case

$$h(t)= -\frac{\lambda}{2}\frac{f(t)}{S_o} \qquad 185$$

h(t) can then be substituted in eq. 180 giving

$$-\frac{\lambda}{2} = \left[\int_{T_1}^{T_2}\frac{f^2(t)}{S_o}\,dt\right]^{-1} \qquad 186$$

Using eq. 186 one can rewrite eq. 184 as

$$\sigma^2{}_{A\,min}= \left[\int_{T_1}^{T_2}\frac{f^2(t)}{S_o}\,dt\right]^{-1} \qquad 184a$$

and eq. 185 as:

$$h(t)=\frac{\frac{f(t)}{S_o}}{\left[\int_{T_1}^{T_2}\frac{f^2(t)}{S_o}dt\right]}=\frac{f(t)}{\left[\int_{T_1}^{T_2}f^2(t)\,dt\right]} \qquad 185a$$

Let discuss some implications of this result. Notice first that the solution in eq. 185a, within an unrelevant multiplicative factor, amounts to multiply the data by a copy of the signal itself and to integrate the result on the available time interval. This is not very different from what we did in extracting the amplitude of a periodic signal with the lock-in technique in the first section of this chapter. In fact if f(t) is a periodic signal of known phase and frequency but of unknown amplitude, eq. 184a give $\sigma^2{}_{Amin}\approx 2S_o/(T_2-T_1)$, a result very close to that obtained in that section.

A second interesting observation is that the same result can be obtained by a least square fitting of the function f(t) to the "data" x(t). In fact within the least square fitting method one assumes to have N independent data x_i and looks for the value of A such that the sum:

$$\sum_{n=1}^{N}\frac{1}{\sigma^2{}_i}\cdot[x_i-Af(t_i)]^2 \qquad 187$$

is a minimum. If the data have all the same variance σ^2 the solution is independent of σ^2 and is

$$A=\frac{\sum_{n=1}^{N}x_if(t_i)}{\sum_{n=1}^{N}f^2(t_i)} \qquad 188$$

it can be seen that eq. 185a is the limit of eq. 188 for a continous set of independent data, as it is the case if the noise is white and gaussian.

Using the definition of the signal energy E in eq. 4, one can see that

$$\sigma^2_{Amin} \geq \frac{S_o}{E} \qquad 189$$

where E is the energy of the signal f(t) and where the equality is attained only if $f(t)= 0$ for $t \leq T_1$ or $t \geq T_2$ so that all the signal energy is concentrated in the interval $T_1 \leq t \leq T_2$.

Such a filter can be obtained by some off-line or delayed processor if the arrival time of the signal is known in advance. This is the case in some particle detectors where the signal itself, due to the arrival of a particle, can be first detected by some threshold sensor that generates a trigger signal. The signal is then fed, trough a properly calibrated delay line, to a stage where it is multiplied by the time function h(t) generated by a signal syntetizer triggered by the trigger signal. The product of the signal and of h(t) is then integrated to give the value of A (Fig 13) .

The filter can also realized by a time invariant real time device. **A** will than be the output of the filter at some time t. To do that let restate eq 174 as

$$A= \int_0^T h(t')x(-t')dt' \qquad 190.$$

where now the data are known for $-T \leq t \leq 0$.

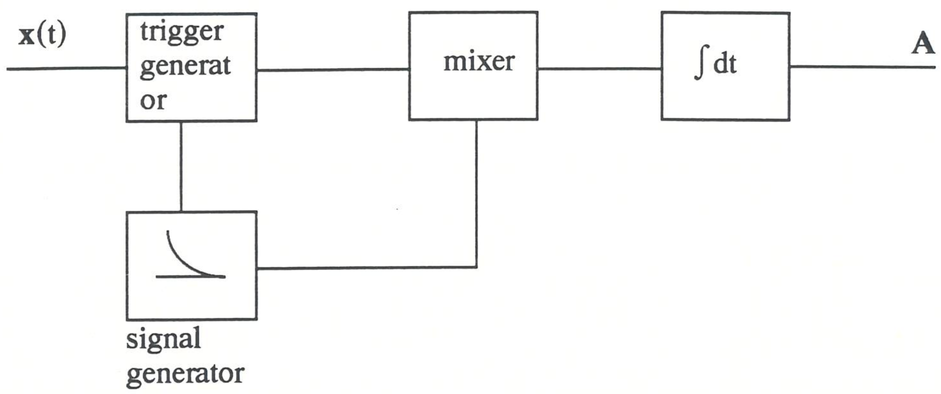

Fig 13

Nothing is changed in the results already obtained except that t has to be replaced everywhere by -t. So for instance

$$h(t)=\frac{f(-t)}{\left[\int_0^T f^2(-t')dt'\right]} \qquad 185b$$

If the filter of inpulse response h(t) is applied to the data between τ -T and τ the result will be

$$A(\tau)=\int_0^T h(t')\mathbf{x}(\tau-t')dt' \qquad 191$$

so that **A** in eq. 190 is $\mathbf{A}(\tau)$ for τ=0. $\mathbf{A}(\tau)$ is then the output of a causal filter with an impulse response consisting of a mirror image of the signal itself between -T and 0. It will result from a signal contribution $A_s(\tau)$and a noise contribution $\mathbf{A}_n(\tau)$. $A_s(\tau)$ is given by:

$$A_s(\tau)= A\frac{\left[\int_0^T f(\tau-t)f(-t)dt\right]}{\left[\int_0^T f^2(-t')dt'\right]} \qquad 192$$

If the energy of the signal in an interval of duration T is maximum between -T and 0 , then $A_s(\tau)$ reaches its maximum value A_s=A for τ=0 .

The noise contribution $\mathbf{A}_n(\tau)$ is a stationary stocastic process with zero mean, autocorrelation given by:

$$R_A(\Delta\tau)= \sigma^2{}_{Amin}\cdot\frac{\left[\int_0^T f(\Delta\tau-t)f(-t)dt\right]}{\left[\int_0^T f^2(-t')dt'\right]} \qquad 193$$

and with an r.m.s. fluctuation $\sigma_{\mathbf{A}min}$.

If the signal arrival time is not known, or if more than one signal is expected to arrive at unpredictable times, the output of the filter can be be sampled and its modulus compared with a threshold level A_{th}. If the sample exceeds the threshold one can conclude that a signal has arrived and take the sample value as an estimate of the amplitude.

To select the threshold level one can consider that in the absence of any signal, the filter output samples are gaussian zero mean random variables. If the sampling time is taken of the order of T, then different samples are also independent. The probability for the n^{th} sample $A(n)$ of being $|A(n)| \geq A_{th}$ is then $2[1\text{-Erf}(A_{th}/\sigma_{A\,min})]$[1]. If for instance $A_{th}=3\sigma_{A\,min}$ then the probability of the detector output to go above the threshold just because of the noise is ≈ .3%. The number of false alarms one can accept depends on the specific application. The lower the number the higher the threshold level and the higher also the signal to noise ratio $A/\sigma_{A\,min}$ needed to achieve a low probability of dismissal of a good signal. The probability that in the presence of a signal of amplitude A, the filter output sample is lower than the threshold is $\text{Erf}(|A|+A_{th})\text{-Erf}(|A|-A_{th})$ if $A \geq A_{th}$. Thus, for instance, if $A_{th}>\sigma_{A\,min}$ and $A=A_{th}$ the probability of a dismissal is still ≈50%. The threshold is then set up compromising between this two conflicting needs according to some cost function that depends on the specific application.

Exemple 5.1 Charge detector dominated by the voltage noise.
In the example in the preceding section of this chapter, at low frequency the current noise dominated on the voltage noise. This translates in the fact that $R>>R_n$. If the opposite is instead true, $R<< R_n$, in any frequency range the noise coincides then with the voltage noise with spectral density $k_BT_nR_n$. One can then apply the above theory to a signal $f(t)=\frac{1}{C}e^{-t/\tau_o}\Theta(t)$ and obtain that the best signal to noise ratio is obtained by multiplying the signal by itself and integrating on a time T longer than τ_o. The minimum detectable charge is, for $T \gtrapprox 3\tau_o$, $2Ck_BT_nR_n/R >> 2Ck_BT_n$.

Non-white noise and infinite time bounds.

If T_1 and T_2 are such that the signal energy is all contained in the integration interval and if $R(T_2-T_1)$ is neglegible, then in all the equations one can set $T_1=-\infty$ and $T_2 = +\infty$. Eq. 183 becomes

---

[1] The error function is here defined as $\text{Erf}(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-x^2/2}$

$$\int_{-\infty}^{\infty} h(t')R(t-t')dt' = -\frac{\lambda}{2} f(t) \text{ for } -\infty \le t \le +\infty \qquad 194$$

that can be solved using the Fourier transforms to give

$$h(\omega) = -\frac{\lambda}{2}\frac{f(\omega)}{S(\omega)} \qquad 195$$

with $S(\omega)$ the power spectrum of the noise. Using Parseval relation in eq. 180 one gets

$$-\frac{\lambda}{2} = \sigma^2{}_{\mathbf{A}min} = \left[\frac{1}{2\pi}\cdot\int_{-\infty}^{+\infty}\frac{|f(\omega)|^2}{S(\omega)}d\omega\right]^{-1} \qquad 196$$

Again **A** can be considered as the output at time zero of a non causal filter with impulse response $h'(t)= h(-t)$. This filter will then have a frequency response

$$h'(\omega) = \frac{\dfrac{f^*(\omega)}{S(\omega)}}{\left[\dfrac{1}{2\pi}\cdot\displaystyle\int_{-\infty}^{+\infty}\frac{|f(\omega)|^2}{S(\omega)}d\omega\right]} \qquad 197$$

When applied to the signal, the filter will produce an output whose Fourier transform is

$$A_s(\omega) = A\,\frac{\dfrac{|f(\omega)|^2}{S(\omega)}}{\left[\dfrac{1}{2\pi}\cdot\displaystyle\int_{-\infty}^{+\infty}\frac{|f(\omega)|^2}{S(\omega)}d\omega\right]} \qquad 198$$

while the output noise $\mathbf{A}_n(t)$ will have spectrum:

$$S_A(\omega) = \sigma^2_{A\min} \cdot \frac{\frac{|f(\omega)|^2}{S(\omega)}}{\left[\frac{1}{2\pi} \cdot \int_{-\infty}^{+\infty} \frac{|f(\omega)|^2}{S(\omega)} d\omega\right]} \qquad 199$$

It is worth noting that $\frac{|A_s(\omega)|^2}{S_A(\omega)} = \frac{A^2}{\sigma_{2A\min}}$ and is independent of frequency so that any further Wiener filter will not improve the signal to noise ratio.

Notice that both $\sigma^2_{A\min}$ and $S_A(\omega)$ do not depend on any filtering stage that acts both on the signal and on the noise.

Exemple 5.2 Charge detector

Let consider the problem of the charge detector we treated by the method of the RC-CR filter. The amplitude of the signal is here Q_o, the unknown charge. The unit amplitude signal has Fourier transform

$$f(\omega) = \frac{1}{C} \frac{\tau_o}{1 + i\omega\tau_o} \qquad 200$$

while the noise has spectrum

$$S(\omega) = kBT_n \left\{ R_n + \frac{\tau_o^2}{R_n C^2} \cdot \frac{1}{1 + \omega^2 \tau_o^2} \right\} \qquad 201$$

The variance of the estimate is

$$\sigma^2_{Q\min} = \left\{ \frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} \frac{\frac{R^2}{1+\omega^2\tau_o^2}}{kBT_n \left\{ R_n + \frac{\tau_o^2}{R_n C^2} \cdot \frac{1}{1+\omega^2\tau_o^2} \right\}} d\omega \right\}^{-1} =$$

$$= \left\{ \frac{1}{2\pi} \cdot \frac{R^2 R_n}{R^2 + R_n^2} \cdot \frac{1}{kBT_n} \int_{-\infty}^{\infty} \frac{1}{1+\omega^2\tau^2} d\omega \right\}^{-1} = \qquad 202$$

$$= 2Ck_BT_n \sqrt{\frac{R^2 + R_n^2}{R^2}}$$

where $\tau' = \tau_o \sqrt{\frac{R_n^2}{R^2 + R_n^2}}$.

The optimum filter frequency response $h(\omega)$ is then

$$h(\omega) = \sigma^2 Q_{min} \frac{\frac{1}{C}\frac{\tau_o}{1+i\omega\tau_o}}{k_BT_n\left\{R_n + \frac{\tau_o^2}{R_nC^2}\cdot\frac{1}{1+\omega^2\tau_o^2}\right\}} = \qquad 203$$

$$= \frac{CR_n}{\sqrt{R^2+R_n^2}}\frac{1-i\omega\tau_o}{1+\omega^2\tau'^2}$$

whose Fourier transform is

$$h(t) = \frac{1}{2R} e^{-|t|/\tau'}\left[1-(2\Theta(t)-1)\sqrt{\frac{R^2+R_n^2}{R_n^2}}\right]. \qquad 204$$

Let discuss these results in the two limit $R>>R_n$ and $R<< R_n$. As it can be seen from eq. 202 both the result obtained with the RC-CR filter and the one obtained instead with the optimal filter in the white noise case, are predicted by eq. 202 as two limiting cases. The filter opens a band pass up to a frequency $1/\tau'$ which is $\tau'=R_nC$ for large R_n and RC for small R_n.

The function $h(t)$ is just an esponential times a $\Theta(t)$ function if $\tau_o=\tau'$, i.e. $R<<R_n$, while it becomes dominated, approaching the opposite limit, by the odd function $[2\Theta(t)-1]e^{-|t|/\tau'}$, that performs both a derivative at time t=0 and an integration on a time of order τ'.

The minimum signal energy is

$$E_{min} = \frac{\sigma^2 Q_{min}}{2C} = k_BT_n\sqrt{\frac{R^2+R_n^2}{R^2}} \qquad 205$$

while the energy uncertainty is given by:

$$\sigma^2_E = \frac{1}{4C^2}\cdot\left(2\,\sigma^4 Q_{min} + 4\sigma^2 Q_{min} Q_o^2\right) \qquad 206$$

so that the energy innovation due to the signal that can be measured with a signal to noise ratio of one is $2.2\, Ck_BT_n\sqrt{\frac{R^2+R_n^2}{R^2}}$.

Notice that, as we already pointed out above, R is the parallel of the source output resistance R_s and of the amplifier input one R_a. Thus, for a given value of R_s, R reaches a maximum, and E_{min} reaches a minimum, when $R_a=R_s$. If this matching is not obtained, but still $R>>R_n$, E_{min} is still $E_{min}\approx Ck_BT_n$.

In the calculation above we included the thermal noise due to R_s in the amplifier current noise generator. To better understand the role of this source of noise let write the total current noise as

$$S_I = \frac{2k_BT}{R_s} + \frac{k_BT_{no}}{R_{no}} \qquad 207$$

and the voltage noise density as

$$S_V = k_BT_{no}R_{no} \qquad 208$$

T_{no} and R_{no} are then the "bare" densities related to the amplifier itself.

The total noise temperature is then

$$T_n = \frac{\sqrt{S_VS_I}}{k_B} = T_{no}\sqrt{1+\frac{2R_{no}T}{R_sT_{no}}} \qquad 209$$

while the noise resistance is

$$R_n = \sqrt{\frac{S_V}{S_I}} = \frac{R_{no}}{\sqrt{1+\frac{2R_{no}T}{R_sT_{no}}}} \qquad 210$$

If again one assumes $R_s= R_a$, one can express the minimum energy as

$$E_{min} = k_BT_{no}\sqrt{1+\frac{2R_{no}T}{R_sT_{no}}+\frac{R_{no}^2}{R_s^2}} \qquad 211$$

In order that the $E_{min} \rightarrow k_BT_{no}$ one needs both $R_{no} << R_s$ and $TR_{no} << T_nR_s/2$. However notice that the temperature of the source can be greater than the noise temperature of the amplifier by the factor R_s/R_{no} and still it will not affect the sensitivity of the charge detection.

Multiple components signal

The theory above can be generalized to the case where the data are given by :

$$\mathbf{x}(t) = \sum_{n=1}^{N} A_n s_n(t) + \mathbf{n}(t) \qquad 212$$

where the signals $s_n(t)$ are known functions of the time and the amplitudes A_n have to be estimated.

One then builds the N estimators

$$\mathbf{A}_n = \int_{-\infty}^{\infty} h_n(t)\mathbf{x}(t)dt \qquad 213$$

and varies $h_n(t)$ until $\mathbf{A}_n$ reaches its minimum variance subject to the N(N-1)/2 conditions

$$\int_{-\infty}^{\infty} h_n(t)s_m(t)dt = \delta_{nm} \qquad 214$$

with δ_{nm} the Kröneker delta. One ends with the N equations

$$2\int_{-\infty}^{\infty} h_n(t')R(t-t')dt' + \sum_{m=1}^{N} \lambda_{nm}s_m(t) = 0 \qquad 215$$

that can be solved by Fourier transforming to give:

$$h_n(\omega) = \sum_{m=1}^{N} C_{nm} \frac{s_m(\omega)}{S(\omega)} \qquad 216$$

and

$$<\mathbf{A}_n\mathbf{A}_m> = C_{nm} \qquad 217$$

with the matrix C_{nm} being the inverse of the matrix

$$-\frac{\lambda_{nm}}{2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{s_n(\omega)s_m{}^*(\omega)}{S(\omega)} d\omega . \qquad 218$$

The variance of the estimator A_n is given, according to eq. 217, by $\sigma^2_{A_n} = C_{nn}$. One can easily check that the formulas above reduce to those of the previous section if N=1.

Example 5.3 The harmonic oscillator as a pulse detector.

Consider the circuit of fig. 14

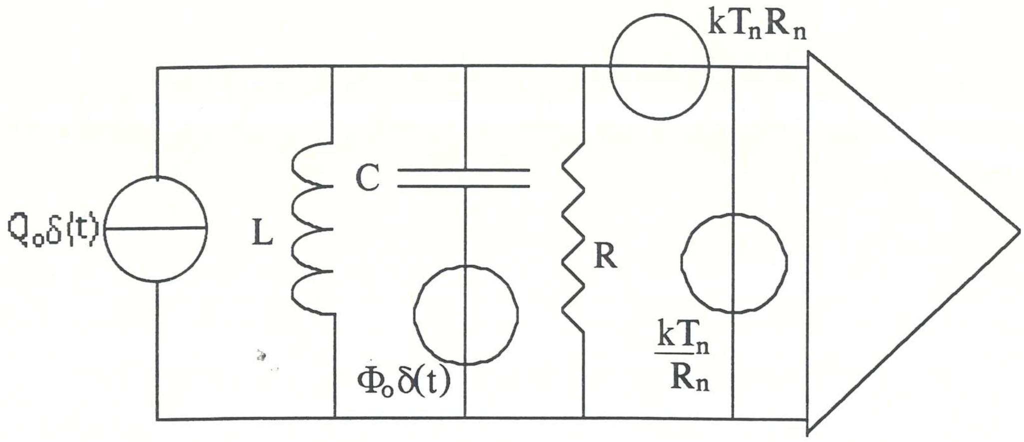

Fig 14. An harmonic oscillator as a pulse detector.

A current generator delivers a current pulse $I(t) = Q_o\delta(t)$ and a voltage generator feed the circuit with a voltage pulse $V(t) = \Phi_o\delta(t)$. The current generator will charge the capacitor with a charge Q_o while the voltage generator will give a magnetic flux step of amplitude Φ_o to the inductor. In the absence of the voltage generator and for low enough dissipation, the voltage across the amplifier input will start to oscillate as $s_1(t) \approx \frac{Q_o}{C}\cos(\omega_o t)$ with $\omega_o = 1/\sqrt{LC}$. In the absence of the current generator, the current in the inductor will oscillate as $I(t) \approx \frac{\Phi_o}{L}\cos(\omega_o t)$ while the voltage will now oscillate as $s_2(t) \approx \omega_o\Phi_o\sin(\omega_o t)$. The two signals $s_1(t)$ and $s_2(t)$ have then phases that differ by $\frac{\pi}{2}$. Let now apply the theory above to the measurement of Q_o and Φ_o. The voltage signals, as seen by the amplifier, have, for unit Q_o and Φ_o, Fourier transforms given by:

$$s_1(\omega) = \frac{i\omega L\omega_o^2}{\omega_o^2 - \omega^2 + i\omega/\tau} \qquad 219a$$

$$s_2(\omega) = \frac{\omega_o^2}{\omega_o^2 - \omega^2 + i\omega/\tau} \qquad 219b$$

with τ=RC.

The noise has a spectral density

$$S(\omega) = k_B T_n R_n\left[1 + \frac{1}{R_n^2C^2}\cdot\frac{\omega^2}{(\omega_o^2-\omega^2)^2+\omega^2/\tau^2}\right] \qquad 220$$

The elements of the matrix $-\frac{\lambda_{nm}}{2}$ are

$$-\frac{\lambda_{11}}{2} = \frac{1}{k_BT_nR_n}\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\omega^2L^2\omega_o^4}{(\omega_o^2-\omega^2)^2+\omega^2/\tau_1^2}\,d\omega = \frac{1}{2Ck_BT_n}\sqrt{\frac{R^2}{R^2+R_n^2}} \qquad 221a$$

$$-\frac{\lambda_{22}}{2} = \frac{1}{k_BT_nR_n}\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\omega_o^4}{(\omega_o^2-\omega^2)^2+\omega^2/\tau_1^2}\,d\omega = \frac{1}{2Lk_BT_n}\sqrt{\frac{R^2}{R^2+R_n^2}} \qquad 221b$$

$$-\frac{\lambda_{12}}{2} = \left(-\frac{\lambda_{21}}{2}\right)^* = \frac{1}{k_BT_nR_n}\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{i\omega L\omega_o^4}{(\omega_o^2-\omega^2)^2+\omega^2/\tau_1^2}\,d\omega = 0 \qquad 221c$$

where I have defined $\frac{1}{\tau_1^2} = \frac{1}{C^2}\left(\frac{1}{R^2}+\frac{1}{R_n^2}\right)$.

The matrix is then diagonal and its inverse has as diagonal elements the inverse of the diagonal elements. The two coefficients A_1 and A_2 are independent and have errors

$$\sigma^2_Q = 2Ck_BT_n\sqrt{\frac{R^2+R_n^2}{R^2}} \;;\; \sigma^2_\Phi = 2Lk_BT_n\sqrt{\frac{R^2+R_n^2}{R^2}} \qquad 222$$

so that the minimum signal energy is now

$$E_{min} = \frac{\sigma^2_Q}{2C} + \frac{\sigma^2_\Phi}{2L} = 2k_BT_n\sqrt{\frac{R^2+R_n^2}{R^2}} \qquad 223$$

The filter frequency functions can be obtained easily as

$$h_1(\omega) = \sigma^2_Q\frac{\dfrac{i\omega L\omega_o^2}{\omega_o^2-\omega^2+i\omega/\tau}}{k_BT_n R_n\left[1+\dfrac{1}{R_n^2C^2}\cdot\dfrac{\omega^2}{(\omega_o^2-\omega^2)^2+\omega^2/\tau^2}\right]} =$$

$$= 2\sqrt{\frac{R^2+R_n^2}{R^2R_n^2}}\ \frac{i\omega(\omega_o^2-\omega^2-i\omega/\tau)}{(\omega_o^2-\omega^2)^2+\omega^2/\tau_1^2} \qquad 224a$$

and

$$h_2(\omega) = \sigma^2_\Phi \frac{\dfrac{\omega_o^2}{\omega_o^2 - \omega^2 + i\omega/\tau}}{k_BT_n\ R_n\left[1+\dfrac{1}{R_n^2C^2}\cdot\dfrac{\omega^2}{(\omega_o^2-\omega^2)^2+\omega^2/\tau_1^2}\right]} =$$

$$= \frac{2}{C}\sqrt{\frac{R^2+R_n^2}{R^2R_n^2}}\ \frac{(\omega_o^2-\omega^2-i\omega/\tau)}{(\omega_o^2-\omega^2)^2+\omega^2/\tau_1^2} \qquad 224b$$

To get a feeling of what the filter response looks like in the time domain one can Fourier transform $h_2(t)$ and consider that, aside the multiplicative factors, $h_1(t)$ is just the time derivative of $h_2(t)$. The Fourier transform of $h_2(t)$ is

$$h_2(t)= \frac{2}{C}\sqrt{\frac{R^2+R_n^2}{R^2R_n^2}}\left[\omega_o^2 g(t) + \frac{d^2g(t)}{dt} - \frac{1}{\tau}\frac{dg(t)}{dt}\right] \qquad 225$$

with

$$g(t)= \frac{\tau_1}{2\omega_o^2}\cdot e^{-|t|/2\tau_1}\left[\cos(\omega_1 t) + \frac{1}{2\omega_1\tau_1}\sin(\omega_1|t|)\right] \qquad 226$$

the Fourier transform of $[(\omega_o^2-\omega^2)^2+\omega^2/\tau_1^2]^{-1}$ and $\omega_1= \sqrt{\omega_o^2-1/4\tau_1^2}$. In conclusion $h_2(t)$ is

$$h_2(t)=e^{-|t|/2\tau_1}\left[\frac{1}{\omega_1\tau}\sin(\omega_1 t)+\frac{1}{\omega_1\tau_1}\sin(\omega_1|t|)\right] \qquad 227$$

Again, as in the case of the capacitive detector, the filter function for $\tau >> \tau_1$, performs a "derivative" multiplying the data, that contain the odd function $\sin(\omega_1 t)$, by the even function $\sin(\omega_1|t|)$. The exponential with the optimum time constant τ_1 takes care of setting the proper integration time.

Because the two pulses give origin to two exponentialy damped harmonic signals with 90° phase shift, the two reduced quantities $a_1=Q_o/C$ and $a_2=\Phi_o/\sqrt{LC}$ can be considered as the amplitudes of the two phase components of the total oscillating voltage signal. The signal will have then phase $\phi = \arctan(a_2/a_1)$. The uncertainty of the phase can then be derived as ususal (see eq. 162) and is given by

$$\sigma_\phi \approx \sqrt{\frac{k_B T_n}{E}} \qquad 228$$

where we have assumed that $R>>R_n$ and we have defined $E=\frac{Q_o^2}{2C}+\frac{\Phi_o^2}{2L}$.

Estimation of the arrival time.

In the preceding sections we have always assumed that the functions $s_n(t)$ were known. It often occurs in practice that the function is known within an arbitrary shift along the time axis. The signal to be extracted from the data takes now the form of $s(t-t_o)$ where t_o, the signal arrival time, is a parameter to be determined. To discuss this case we will assume that only one signal has to be extracted from the data so that these now are given by:

$$\mathbf{x}(t)= As(t-t_o) +\mathbf{n}(t) \qquad 229$$

As a preliminary observation, consider the amplitude estimator as the output $\mathbf{A}(t)$ of a (non causal) filter measured at time zero. $A_s(\omega)$ has an even and positive Fourier transform (Eq. 198) and thus reaches it maximum[1] for t=0. The noise is instead stationary so that the r.ms. fluctuation is independent of the time. Thus a first guess of the arrival time can be obtained taking the time for which the amplitude estimator reaches is maximum.

To have an estimate of the uncertainty of this procedure, let assume that after the first guess the uncertainty in the knowledge of the arrival time is so small than the signal $s(t-t_o)$ can be expanded as

$$s(t-t_o) = s(t) - \frac{ds(t)}{dt}t_o \qquad 230$$

one can then apply the the amplitude estimation theory considering the data as given by

$$\mathbf{x}(t)= As(t)+B\frac{ds(t)}{dt}+\mathbf{n}(t) \qquad 231$$

obtain A and B and estimate t_o as $t_o=-B/A$.

The matrix $-\frac{\lambda_{nm}}{2}$ has elements:

---

1 $\int_{-\infty}^{\infty} f^2(\omega)e^{i\omega t}dt = 2\int_0^{\infty} f^2(\omega)\cos(\omega t)dt \leq 2\int_0^{\infty} f^2(\omega)dt$

$$-\frac{\lambda_{AA}}{2} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{|s(\omega)|^2}{S(\omega)}\,d\omega \qquad 232a$$

$$-\frac{\lambda_{BB}}{2} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\omega^2|s(\omega)|^2}{S(\omega)}\,d\omega \qquad 232b$$

$$-\frac{\lambda_{BA}}{2} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{i\omega|s(\omega)|^2}{S(\omega)}\,d\omega=0 \qquad 232c$$

so that **A** and **B** are independent random variables. The error on t_o can be easily obtained by propagation of errors and is

$$\sigma^2{}_t \approx \frac{\sigma^2{}_B}{A^2} + \frac{B^2}{A^2}\cdot\frac{\sigma^2{}_A}{A^2} = \frac{\sigma^2{}_A}{A^2}\cdot\frac{\sigma^2{}_B}{\sigma^2{}_A} \qquad 233$$

because t_o=0 on the average.

The error is then

$$\sigma^2{}_t \approx \frac{\sigma^2{}_A}{A^2}\,\frac{\int_{-\infty}^{\infty}|s(\omega)|^2 d\omega}{\int_{-\infty}^{\infty}\omega^2|s(\omega)|^2 d\omega} \qquad 234$$

and is inversely proportional to the signal to noise ratio.

For a narrow band signal or noise, where one can assume $\int_{-\infty}^{\infty}\omega^2|s(\omega)|^2 d\omega \approx \omega^2{}_o \int_{-\infty}^{\infty}|s(\omega)|^2 d\omega$, with ω_o the center frequency of the signal. One gets then $\sigma_t \approx 1/\omega_o \cdot \sigma_A/|A|$, or a phase uncertainty $\sigma_\phi = \omega_o\sigma_t = \sigma_A/|A|$ as obtained for the pulse driven oscillator in the preceding section.

Notice however that **A**(t), for a narrow band signal, will present an oscillating pattern with period ω_0 . This is due to the fact that the filter multiplies two oscillating signals shifted by t. If the signal to noise ratio is not too high, many maxima of this oscillating pattern will have comparable amplitudes and the identification of the true maximum is affected by uncertainty. The discussion on this point ,and on the signal to noise ratio needed to suppress the peak ambiguity, goes beyond the scope of these lectures and can be found in the literature.

If the noise is white eq. 234 gives

$$\sigma^2_t \approx \frac{\sigma^2_A}{A^2} \frac{\int_{-\infty}^{\infty} s(t)^2 dt}{\int_{-\infty}^{\infty} \left[\frac{ds(t)}{dt}\right]^2 dt} \qquad 235$$

that, for instance, for an exponential signal $e^{-t/\tau}\,\Theta(t)$ gives $\sigma^2_t \approx \frac{\sigma^2_A}{A^2}\,\tau^2$.

Quantum limits

In this section I will discuss the limits put by quantum mechanics on some of the results derived earlier. Specifically I will discuss the quantum limit on the noise temperature of an amplifier and the quantum limit of the theory of the thermodynamic fluctuations.

Quantum limit on the noise of an amplifier.

Consider the circuit in Fig. 14 of the preceding section without the two signal generators. The amplifier has high enough gain so that any ouput signal can be considered a classical physical variable that can be measured without any significant uncertainty.

The amplifier output can be considered as a measurement of the istantaneous value of the charge on the capacitor as the input voltage is V(t)= Q(t)/C.

If the effect of the resistor can be neglected, $R >> \sqrt{L/C}$, the dynamic of the oscillating circuit at the amplifier input can be obtained using Hamilton equations[1] with the Hamiltonian:

$$H(Q,\Phi) = \frac{Q^2}{2C} + \frac{\Phi^2}{2L} \qquad 236$$

where Φ is the magnetic flux across the inductor. As a consequence Q and Φ are canonically conjugate variables that in, the quantum limit, have to obey the uncertainty principle

$$\Delta Q \Delta \Phi \geq \frac{\hbar}{2} \qquad 237$$

In the absence of any filter, the amplifier output will have infinite uncertainty and so will have the charge. If an integrator is used, that integrates the amplifier output on a time τ, the uncertainty on the capacitor integrated voltage will be

$$\sigma^2{}_V = \frac{k_B T_n R_n}{\tau} \qquad 238$$

and thus the error in estimating the charge is

---

[1] $\frac{dQ}{dt} = \frac{\partial H}{\partial \Phi} = \frac{\Phi}{L}$; $\frac{d\Phi}{dt} = -\frac{\partial H}{\partial \Phi} = -\frac{Q}{C} = -V$

$$\sigma^2_Q = \frac{C^2 k_B T_n R_n}{\tau} \qquad 239$$

Notice that that eq. 238 gives the uncertainty on the total charge. This total charge includes also that due to the current noise. After the measurement has taken place the state of the system is let with an uncertainty on the charge $\Delta Q = \sigma_Q$

During the integration time, due to the noise current, the flux in the inductor drifts by a mean square quantity:

$$\Delta^2\Phi = <[\Phi(t+\tau) - \Phi(t)]^2> = 2[R_\Phi(0) - R_\Phi(\tau)] \qquad 240$$

The flux noise autocorrelation $R_\Phi(\tau)$ can be calculated from its noise spectrum:

$$S_\Phi(\omega) = \frac{k_B T_n}{R_n} \frac{L^2 \omega_o^4}{(\omega_o^2 - \omega^2)^2 + \omega^2/\tau'^2} \qquad 241$$

with $\tau' = RC$ and is

$$R_\Phi(\tau) = \frac{k_B T_n R}{2R_n} L\, e^{-|t|/2\tau'} \left[\cos(\omega_1 \tau) + \frac{1}{2\omega_1 \tau'} \sin(\omega_1 |\tau|)\right] \qquad 242$$

with $\omega_1 \sqrt{\omega_o^2 - 1/4\tau^2}$, that for $\tau << \tau', 1/\omega_1$ becomes

$$R_\Phi(\tau) = \frac{k_B T_n R}{2R_n} L \left(1 - \frac{|\tau|}{\tau'}\right) \qquad 243$$

As a consequence $\Delta^2\Phi$ is given by

$$\Delta^2\Phi = \frac{k_B T_n R}{R_n} L \frac{|\tau|}{\tau'} \qquad 244$$

Whatever the precision with which the flux was measured at the beginning of the integration time, at the end the uncertainty on the flux value will be $\Delta\Phi$.

The product $\Delta^2\Phi\Delta^2 Q$ is then given by

$$\Delta^2\Phi\Delta^2 Q = \frac{k_B T_n R}{R_n} L \frac{\tau}{\tau'} \cdot \frac{C^2 k_B T_n R_n}{\tau} = \left(\frac{k_B T_n}{\omega_o}\right)^2 \qquad 245$$

Imposing the uncertainty principle in eq. 245 implies:

$$k_B T_n \geq \frac{\hbar\, \omega_o}{2} \qquad 246$$

The noise energy $2k_BT_n$, which is also the minimum signal energy that can be detected, has to be larger than the energy of one quantum in the input oscillator.

For this reason it is customary, for very low noise amplifiers, to express the noise energy at a given frequency in number of quanta or in multiple of $\hbar$. Notice that, under this respect , the same noise temperature corresponds to quite different number of quanta in different frequency ranges. In the audio frequency range, say at 1 KHz, the minimum noise temperature is $T_n \approx 5 \cdot 10^{-8}$ K while at 10 GHz $T_n \approx 0.5$ K and is not very far from the present limit of amplifiers.

The result above has been derived assuming the gain G of the amplifier to be infinite. It can be shown that if this is not the case, then the right hand side of eq. 246 has to be multiplied by 1-1/G.

Quantum non demolition and stroboscopic measurements.

The result obtained in the preceding paragraph has an important consequence for measurements that use an harmonic oscillator to detect weak signals and that use amplifiers with noise energies near to the quantum limit. This is the case, for instance, for resonant gravitational wave detectors where the SQUID amplifiers are not very far from the quantum limit.

Despite the quantum behaviour of the detecting electronics, the signal to be measured, like the gravitational force, is, for this kind of applications, fully classical (large number of field quanta). Thus the sensitivity limit set by the uncertainty principle of the oscillator seems to be more an accident that a fundamental problem.

Let discuss in some more detail the meaning of this limit. If we monitor the voltage with high precision, low R_n, the current noise generator will change the flux by a large quantity. As a consequence also the voltage, $V=\frac{d\Phi}{dt}$ will change in an unpredictable way contributing to the total uncertainty on the amplitude of any signal driving the oscillator. Whatever R_n, the minimum signal energy, as derived in the preceding section, cannot be less than $\hbar\omega$. The true reason for this is that the flux and the voltage are related quantities and any perturbation on one of them reflects on the other. The relation that links the voltage to the flux, on the other side, is just the Hamilton equation for the flux itself. Thus, in conclusion, it is the fact that neither the flux nor the charge are conserved quantities in the harmonic oscillator that couples them together and prevents to monitor one of them, and with arbitrary precision.

A possible way out has been suggested in the past and is the subject of experimental study by many laboratories. The idea is known as the quantum non demolition technique and bears close relation with the subject of state squeezing in laser and microwave systems. I will briefly discuss here the basic principle of one version of the idea that is known as stroboscopic measurement.

The basic idea is that, if you want to measure a classical force with an arbitrary small error, you have to monitor some observable which, in the absence of the force, is a conserved quantity of the physical system you use as a detector. Only in this case the observable will have a time evolution that will not mix with other variables of the system and will be only affected by the incoming signal.

In an harmonic oscillator neither the coordinate nor the momentum are conserved quantities. The amplitudes of the two quadrature components of the oscillation are instead both conserved quantities. In the absence of any dissipation these quadrature components, A_1 and A_2, are defined classically by the following equations:

$$\frac{Q(t)}{\sqrt{C}} = A_1 \sin(\omega_o t) + A_2 \cos(\omega_o t) \qquad 247a$$

$$\frac{\Phi(t)}{\sqrt{L}} = A_1 \cos(\omega_o t) - A_2 \sin(\omega_o t) \qquad 247b$$

so that, while the flux and the charge oscillate, A_1 and A_2 remain constant.

Consider now, in the quatum limit, Q, Φ , A_1 and A_2 to be operators. It is easy to check that the commutator $[A_1, A_2]$ is :

$$[A_1 A_2] = \omega_o^2 [\Phi Q] \qquad 248$$

so that also A_1 and A_2 have to obey the uncertainty principle

$$\Delta A_1 \Delta A_2 \geq \frac{\hbar\omega_o}{2} \qquad 249$$

I will try to show that however, if one can manage to excite with the back action noise of the amplifier only say A_2, then A_1 can in principle be monitored with infinite precision.

Consider the circuit in Fig. 15.

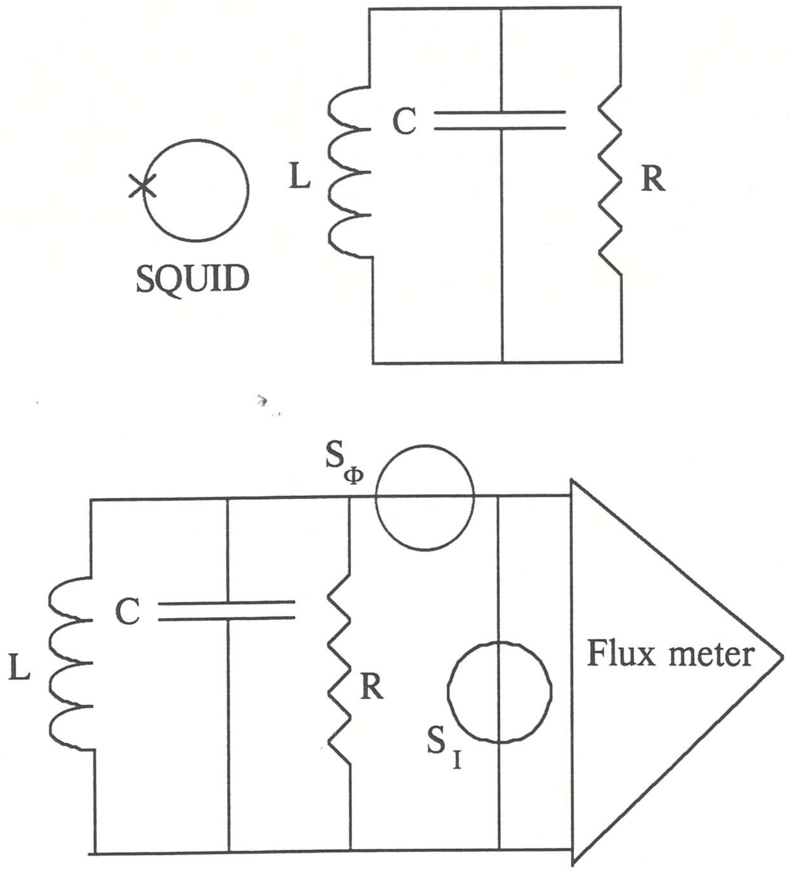

Fig 15 A SQUID coupled to an harmonic oscillator and its equivalent circuit

For the purpose of what I have to discuss here, the SQUID is just a magnetic flux detector that can detect the magnetic flux inside the inductor of the RLC circuit. The reson to refer here to such a kind of amplifier is that, in practice, in the audio frequency range most studies on quantum non demolition measurement use SQUIDs.

The SQUID has a voltage ouput which is white in the frequency range where the oscillator works. This means that the flux noise generator in Fig. 15 has a white spectrum $S_\Phi(\omega)= S_\Phi$. In addition, the SQUID applies a flux noise to the inductor whose effect can be represented by the ususal current noise generator that we will assume of white spectrum too $S_I(\omega)=S_I$. Notice that because the flux noise is white, the associated voltage noise has spectrum $S_V(\omega)=\omega^2 S_\Phi$ so that the noise temperature $T_n = \frac{\sqrt{S_V S_I}}{k_B}$ $= \omega\frac{\sqrt{S_\Phi S_I}}{k_B}$ results to be proportional to ω in agreement with the request of eq. 245. It is then natural to define $S_\Phi = N\frac{\hbar}{2}L_n$ and $S_I = \frac{N\hbar}{2L_n}$ so that at, at the quantum limit N=1.

Suppose now that, by some sort of electric switch, one can uncouple the SQUID, together with its current noise generator, from the circuit. Let switch on the coupling for a time $\delta t << \frac{2\pi}{\omega_o}$ starting at time t=0. According to equation 247b the flux measured by the SQUID, $\Phi(t)$, will be a short pulse of amplitude $\Phi_o = A_1\sqrt{L}$. To evaluate the precision with which the height of the pulse can be evaluated, one can consider that this is a signal in white noise so that the minimum uncertainty on Φ_o is $\sigma^2_\Phi = \frac{N\frac{\hbar}{2}L_n}{E}$ with E the energy of the signal of unit amplitude. This one is a unit height "box" of duration δt and thus its energy is $E=\delta t$. As a consequence the amplitude A_1 is measured with a precision:

$$\sigma^2_{A_1} = \frac{N\frac{\hbar}{2}L_n}{L\delta t} \qquad 250$$

During the measuring time the current noise, which is now coupled to the oscillator, makes both the flux and the charge to change. This variations can be calculated from

$$Q(\delta t) = \int_0^{\delta t} h_1(t')I_n(T-t')dt' \qquad 251a$$

$$\Phi(\delta t) = \int_0^{\delta t} h_2(t')I_n(T-t')dt' \qquad 251b$$

where $I_n(t)$ is the noise current and $h_1(t)$ and $h_2(t)$ are

$$h_1(t)= e^{-t/2\tau} [\cos(\omega_1 t) - \frac{1}{2\tau\omega_1} \sin (\omega_1 t)] \qquad 252a$$

$$h_2(t)=\frac{1}{C\omega_1} e^{-t/2\tau} \sin(\omega_1 t) \qquad 252b$$

The quantities $Q(\delta t)$ and $\Phi(\delta t)$ are zero mean random variable with variances

$$<Q(\delta t)^2>= \int_0^{\delta t}\int_0^{\delta t} h_1(t)h_1(t')<I_n(\delta t-t)I_n(\delta t-t')>dtdt' =$$

$$=\frac{N\hbar}{2L_n} \int_0^{\delta t} h_1^2(t)dt \approx \frac{N\hbar}{2L_n} \delta t \qquad 253a$$

$$<\Phi(\delta t)^2>=\frac{N\hbar}{2L_n} \frac{1}{3C^2} \delta t^3 \qquad 253b$$

and covariance

$$<Q(\delta t)\Phi(\delta t)>=\frac{N\hbar}{2L_n} \frac{1}{2C} \delta t^2 \qquad 253c$$

From the inverse of eqs. 247 one can obtain the variation of the two phases after the time δt as

$$\delta A_1=\frac{\Phi(\delta t)}{\sqrt{L}} + \frac{Q(\delta t)}{\sqrt{C}} \omega_o \delta t \qquad 254a$$

$$\delta A_2=\frac{Q(\delta t)}{\sqrt{C}} - \frac{\Phi(\delta t)}{\sqrt{L}} \omega_o \delta t \qquad 254b$$

the mean square fluctuation of which are:

$$\Delta^2 A_1=\frac{<\Phi(\delta t)^2>}{L} + \frac{<Q(\delta t)^2>}{C} \omega_o^2 \delta t^2 +2<Q(\delta t)\Phi(\delta t)>\omega_o^2 \delta t \approx$$

$$\approx \frac{N\hbar}{2L_n} \frac{7\omega_o^2}{3C} \delta t^3 \qquad 256a$$

$$\Delta^2 A_2 = \frac{<\Phi(\delta t)^2>}{L} \omega_o^2 \delta t^2 + \frac{<Q(\delta t)^2>}{C} - 2<Q(\delta t)\Phi(\delta t)>\omega^2_o \delta t \approx$$

$$\approx \frac{N\hbar}{2L_n} \frac{1}{C} \delta t \left[1 - \omega_o^2 \delta t^2 + \frac{\omega_o^4 \delta t^4}{3}\right] \qquad 256b$$

so that, to first order in δt, only A_2 is affected by the measurement.

Notice that the product $\sigma_{A_1} \Delta A_2 = N\frac{\hbar\omega_o}{2}$ in agreement with eq. 242 for N=1.

If now the switch is closed, again for a short time δt, for $t = n\pi/\omega_o$, with n an integer, and if the sign of the coupling during each switch-on can be reversed, the amplitude of A_1 can be monitored with a precision that increases with $1/\sqrt{M}$, with M the number of repetitions of the measurement, while the random drive on A_2 increases its mean square amplitude with $\sqrt{M}$.

It is straightforward to check, from example 5.3, that A_1 is also excited by a flux pulse arriving at time $t= n\pi/\omega_o$ while A_2 is excited by a charge pulse arriving at the same time. So the two phases are sensitive to classical "force" signals and one of them can be monitored, in principle, with arbitrary precision.

Thermodynamic fluctuations at the quantum limit

The results obtained for the thermal noise in electrical or mechanical networks are special cases of a more general theory due to Callen and Welton[1]. The theory not only achieve a generalization to a large class of physical system but also works out the case where $k_B T << \hbar\omega$ where the results obtained above are not valid anymore. Here I will try to briefly outline this more general theory.

Suppose that a physical system is described by an Hamiltonian

$$H(p,q)=H_o(p,q)+V(t)Q(p,q) \qquad 257$$

---

[1] H. B. Callen and A. Welton Phys Rev **83**, 34 (1951)

wher p,q indicate the set of coordinate and momenta of the system. V(t) is some classical "force" that couples to the system by multiplying the operator Q(p,q). The system could be for an electrical linear device and Q(q,p) would be the charge, related to the current by $I=\frac{dQ}{dt}$, while V(t) would be in this case the voltage due to the external sources.

If $V(t)=V_o\sin(\omega t)$, and if V(t)Q(p,q) can be considered a small perturbation to the hamiltonian $H_o(p,q)$, then we can use the perturbation theory to calculate the probability per unit time w_n that the system, initially prepared in an eigenstate $|E_n>$ of the unperturbed hamiltonian $H_o(p,q)$, will undergo a transition to another state $|E_m>$. From standard formulas

$$w_n=\frac{\pi V_0^2}{2\hbar^2}\sum_m |Q_{nm}|^2[\delta(\omega+\omega_{nm})+\delta(\omega-\omega_{nm})] \qquad 258$$

where $Q_{nm}=<E_n|Q|E_m>$ and $\hbar\omega_{nm}=E_n-E_m$.

Each time the system undergoes a transition it exchange an energy $\hbar\omega_{nm}$ so that the total power dissipated is

$$P_n=\frac{\pi V_0^2}{2\hbar}\,\omega\sum_m |Q_{nm}|^2[\delta(\omega+\omega_{nm}) - \delta(\omega-\omega_{nm})] \qquad 259$$

at thermodynamic equilibrium the total power dissipated has to be averaged on the thermal population of the states:

$$P=\sum_n P_n e^{F- E_n/k_BT} =$$

$$=\frac{\pi V_0^2}{2\hbar}\,\omega\sum_{n,m} |Q_{nm}|^2 e^{F- E_n/k_BT}[\delta(\omega+\omega_{nm}) - \delta(\omega-\omega_{nm})] \qquad 260$$

If the states are dense in energy eq. 260 can be integrated with an energy density $\rho(E)$:

$$P=\frac{\pi V_0^2}{2}\,\omega\cdot \qquad 261$$

$$\cdot \int_0^\infty dE\,\rho(E)e^{F-E/k_BT}\{|<E+\hbar\omega|Q|E>|^2\rho(E+\hbar\omega)-|<E-\hbar\omega|Q|E>|^2\rho(E-\hbar\omega)\}$$

Now for a system with an hamiltonian like that in eq. 257, at least if the effect of the perturbation is considered to first order, the time evolution of Q(p,q) will involve a linear equation with some inpulse response

$$Q(t)=\int_0^\infty h(t')V(t-t')dt' \qquad 262$$

where h(t) can be calculated from eq. 257. If V(t) is the periodic signal above, the mean power absorbed over one period will be given by

$$P=\frac{1}{2}\,\omega h''(\omega)V_o^2 \qquad 263$$

with h"(ω) the imaginary part of the Fourier transform of h(t).

On the other hand, in order eq 262 is obeyed in the classical limit, one needs it holds for the mean value of the operator Q(p,q). In addition the mean energy dissipation under the effect of the perturbation Q(p,q)V(t) is $\frac{d<E>}{dt} = <\frac{\partial E}{\partial t}>=<Q>\frac{dV}{dt}$ with < > indicating the thermodynamic average. As a consequence one can identify the power dissipation in eq. 261 with that in eq. 263 and obtain

$$h''(\omega)= \pi \cdot \qquad 264$$

$$\cdot \int_0^\infty dE\,\rho(E)e^{F-E/k_BT}\{|<E+\hbar\omega|Q|E>|^2\rho(E+\hbar\omega)-|<E-\hbar\omega|Q|E>|^2\rho(E-\hbar\omega)\}$$

Let now calculate the fluctuations of Q(p,q). Let assume that, in the absence of the force V(t), the mean value of Q is zero. To calculate the variance of Q is then sufficient to calculate the mean value of the square of Q. Let first calculate it in a defined quantum state

$$<E_n|Q^2|E_n>= \sum_m <E_n|Q|E_m><E_m|Q|E_n>=\sum_m |Q_{nm}|^2 \qquad 265$$

and then take the thermodinamic average:

$$<Q^2>=\sum_{n,m}|Q_{nm}|^2 e^{F-E_n/k_BT} \qquad 266$$

or for dense states

$$<Q^2>=\hbar\cdot\int_0^\infty\int_0^\infty dE d\omega \rho(E) e^{F-E/k_BT}\{|<E+\hbar\omega|Q|E>|^2\rho(E+\hbar\omega)$$
$$+|<E-\hbar\omega|Q|E>|^2\rho(E-\hbar\omega)\} \qquad 267$$

The remarkable thing about eq. 264 and eq. 267 is that they involve the same couple of energy integral: on one side the dissipation is due to transitions between energy states in a $\approx k_BT$ range around the ground state so that the square moduli of the Q_{nm}'s between states in that energy range are involved. On the other side the fluctuations are due to the quantum fluctuations within each state, that involve again the Q_{nm}'s, again thermally averaged.

Now, because $<E-\hbar\omega|Q|E>=0$ if $E < 0$, one of the two integrals can be rewritten as

$$\int_0^\infty dE\rho(E)e^{F-E/k_BT}|<E-\hbar\omega|Q|E>|^2\rho(E-\hbar\omega)=$$
$$=\int_0^\infty dE\rho(E)e^{F-(E+\hbar\omega)/k_BT}|<E+\hbar\omega|Q|E>|^2\rho(E+\hbar\omega)= \qquad 268$$
$$=e^{-\hbar\omega/k_BT}\int_0^\infty dE\rho(E)e^{F-(E/k_BT)}|<E+\hbar\omega|Q|E>|^2\rho(E+\hbar\omega)$$

so that we can write eq 264 and 267 as

$$h''(\omega) = \pi(1-e^{-\hbar\omega/k_BT})\int_0^\infty dE\rho(E)e^{F-(E/k_BT)}|<E+\hbar\omega|Q|E>|^2\rho(E+\hbar\omega)$$

269

and

$$<Q^2> = \hbar\int_0^\infty\int_0^\infty dEd\omega(1+e^{-\hbar\omega/k_BT})\rho(E)e^{F-(E/k_BT)}|<E+\hbar\omega|Q|E>|^2\rho(E+\hbar\omega)) \quad 270$$

or, in conclusion

$$<Q^2> = \frac{\hbar}{\pi}\int_0^\infty \frac{1+e^{-\hbar\omega/k_BT}}{1-e^{-\hbar\omega/k_BT}} h''(\omega)\, d\omega \quad 271$$

The function

$$E_{ho}(\omega,T) = \frac{\hbar\omega}{2}\frac{1+e^{-\hbar\omega/k_BT}}{1-e^{-\hbar\omega/k_BT}} = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\hbar\omega/k_BT}-1} \quad 272$$

is the mean energy of an harmonic oscillator of frequency ω and temperature T so that eq. 271 becomes:

$$<Q^2> = \frac{2}{\pi}\int_0^\infty E_{ho}(\omega,T)\frac{h''(\omega)}{\omega}\, d\omega \quad 273$$

that for $\hbar\omega << k_BT$ becomes:

$$<Q^2>=\frac{2}{\pi}\int_0^\infty k_B T \frac{h''(\omega)}{\omega} d\omega \qquad 274$$

To have a connection with the theory of, for instance the Nyquist noise in a linear electrical device think, as already stated, to Q as the charge and to V as the voltage in a passive linear device. Then $i\omega h(\omega)=Y(\omega)$ becomes the admittance of the device. Then, in the thermodynamic limit

$$<I^2>=\frac{2}{\pi}\int_0^\infty k_B T\ \omega^2 \frac{h(\omega)}{\omega} d\omega=\frac{2}{\pi}\int_0^\infty k_B T\ \ Y'(\omega)\ d\omega \qquad 275$$

which is the Nyquist formula.

Notice that the general result of eq. 273 shows that the thermal approximation breaks down, as already stated, when $\hbar\omega \approx k_B T$ where the mean thermal energy of the oscillator has to be expressed by the proper quantum formula. At room temperature the transition frequency is $\approx 7\ 10^{12}$ Hz, a microwave frequency, and becomes $\approx$ 1 MHz at 50 μK.

Though one could think of $E_{ho}(\omega,T)h''(\omega)/\omega$ as to a spectral density, this identification is correct only in the classical limit. In the quantum regime an autocorrelation is an ill-defined quantity as operators that refer to different times do not in general commute. As a consequence only the integral formula of eq. 273 has to be used.

Example 6.1 Radiation damping.

This exemple is taken from the paper of Callen and Welton. Consider an electrical dipole of charge q elastically bound to some equilibrium position. Any acceleration a(t) of the dipole will cause it to radiate energy and experience a radiation damping force

$$F=\frac{q^2}{6\pi\varepsilon_o c^3}\frac{da}{dt} \qquad 276$$

with obvious meaning of the symbols.

If an external force is driving the dipole the frequency response that links the dipole coordinate $x(\omega)$ to the force $f(\omega)$ is given by

$$x(\omega)=\frac{f(\omega)}{m}\frac{1}{\omega_o^2-\omega^2+i\omega^3\dfrac{q^2}{6\pi\varepsilon_o c^3}}=h(\omega)f(\omega) \qquad 277$$

the power dissipated by a sinusoidal force at frequency ω_o will be

$$P=\frac{f_o^2}{2}\,\omega_o h''(\omega_o) \qquad 278$$

It is interesting, more than to calculate the fluctuations of x, to refer these fluctuations to a fluctuating force with a "spectrum" $S_f(\omega)=S_a(\omega)/|h(\omega)|^2$ · The "spectrum" of this force is then

$$S_f(\omega)= 2E_{ha}(\omega,T)\frac{h''(\omega)}{\omega|h(\omega)|^2} = \frac{q^2}{6\pi\varepsilon_o c^3}\,\omega^2 E_{ha}(T,\omega) \qquad 279$$

if the force is expressed as an equivalent driving electic field E=f/q this field will have a mean energy density per unit frequency $U(\omega)=\varepsilon_o<S_E(\omega)>$ given by

$$U(\omega)= \frac{1}{2\pi c^3}\omega^2 E_{ha}(\omega,T) \qquad 280$$

where I have assumed that there is a fluctuating force, and then a fluctuating electric field, along eachone of the three space directions. Eq. 280 is just the Plank formula for the black body radiation including the zero point energy of the radiation field.