Chapter 1

Statistical Properties of a Random Process

1.1. Definitions

1.1.1. *Random variable*

A random variable is a quantity whose instantaneous value cannot be predicted. Knowledge of the values of the variable before time t does not make it possible to deduce the value at the time t from it.

Example: the Brownian movement of a particle.

If a vibration was perfectly random, its analysis would be impossible. The points that define the signal would have an amplitude that varied in a completely unpredictable way. Thankfully, in practice, it is possible to associate with all the points that characterize the signal a probability law which will enable a statistical analysis [AND 11].

The principal characteristic of a random vibration is to simultaneously excite all the frequencies of a structure [TUS 67]. In contrast to sinusoidal functions, random vibrations are made up of a continuous range of frequencies, the amplitude of the signal and its phase varying with respect to time in a random fashion [TIP 77] [TUS 79]. Thus, the random vibrations are also called *noise*.

Random functions are sometimes defined as a continuous distribution of sinusoids of all frequencies whose amplitudes and phases vary randomly with time [CUR 64], [CUR 88].

1.1.2. *Random process*

Let us consider, as an example, the acceleration recorded at a given point on the dial of a truck traveling on a good road between two cities A and B. For a journey, the recorded acceleration obeys the definition of a random variable. The vibration characterized by this acceleration is said to be *random* or *stochastic.*

Complexity of the analysis

Even in the most simple hypothesis where a vehicle runs at a constant speed on a straight road in the same state, each vibration measure ${}^{i}\ell(t)$ at one point of the vehicle is different from the other. An infinity of measures to completely characterize the trip should be completed *a priori*.

We define as a *random process* or *stochastic process* the ensemble of the time functions $\left\{{}^{i}\ell(t)\right\}$ for t included between $-\infty$ and $+\infty$, this ensemble being able to be defined by statistical properties [JAM 47].

By their very nature, the study of vibrations would be intensive if we did not have the tools to limit the complete process analysis, made up of a large number of signals according to time, with a very long duration, to that of a very restricted number of samples of reasonable duration. Fortunately, random movements are not erratic in the common sense, but instead follow well-defined statistical laws. The study of statistical process properties, with averages in particular, will enable the simplification of the analysis from two very useful notions for this objective: stationarity and ergodicity.

1.2. Random vibration in real environments

By its nature, the real vibratory environment is random [BEN 61a]. These vibrations are encountered:

– on road vehicles (irregularities of the roads);

– on aircraft (noise of the engines, aerodynamic turbulent flow around the wings and fuselage, creating non-stationary pressures, etc.) [PRE 56a];

– on ships (engine, swell, etc.);

– on missiles. The majority of vibrations encountered by military equipment, and in particular by the internal components of guided missiles, are random with respect to time and have a continuous spectrum [MOR 55]: the gas jet emitted with a large velocity creates important turbulences resulting in acoustic noise which attacks the

skin of the missile until its velocity exceeds Mach 1 approximately (or until it leaves the Earth's atmosphere) [ELD 61], [RUB 64], [TUS 79];

– in mechanical assemblies (ball bearings, gears, etc.).

1.3. Random vibration in laboratory tests

Tests using random vibrations first appeared around 1955 as a result of the inability of sine tests to correctly excite equipment exhibiting several resonances [DUB 59], [TUS 73]. The tendency in standards is thus to replace the old swept sine tests which excite resonances one after the other by a random vibration whose effects are nearer to those of the real environment.

Random vibration tests are also used in a much more marginal way:

– to identify the structures (research of the resonance frequencies and measurement of Q factors), their advantage being that of shorter test duration;

– to simulate the effects of shocks containing high frequencies and which are difficult to replace by shocks of simple form.

1.4. Methods of random vibration analysis

The first mathematical analysis of random vibrations was carried out by A. Einstein [EIN 05], [PAE 11] for the study of Brownian motion of particles in a liquid medium. The theoretical works of S.O. Rice [RIC 39], [RIC 44] on the mathematical analysis of random noise and the distribution of the peaks of random signals quickly became the basis for different studies such as the effect of turbulence on airplanes [DIE 56], [FUN 53], [FUN 55], action of swell on marine structures [CAR 56], [LON 52] or that of wind on civil engineering structures [DAV 64], then for the calculation of the response of structures to the noise of propulsion engines in airplanes and missile thrusters [CRA 58], [CRA 63].

Taking into account their randomness and their frequency contents, these vibrations can be studied only using statistical methods applied to the signals with respect to time or using curves plotted in the frequency domain (spectra).

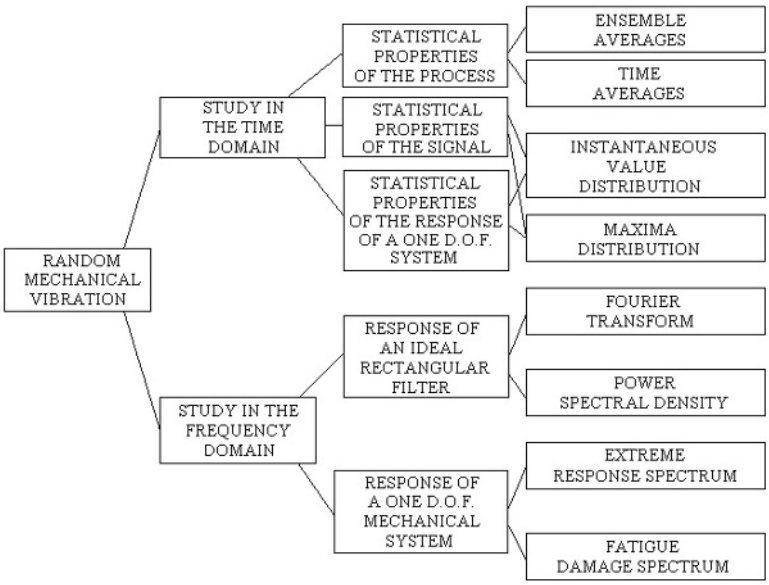


Figure 1.1. *Analysis possibilities for random vibration*

We can schematically distinguish four ways of approaching the analysis of random vibrations [CUR 64], [RAP 69]:

– analysis of the ensemble statistical properties of the process;

– methods of correlation;

– spectral analysis;

– analysis of statistical properties of the signal with respect to time.

The block diagram (Figure 1.1) summarizes the main possibilities which will be considered in turn in what follows.

The parameters most frequently used in practice are:

– the rms (root mean square) value of the signal and, if it is the case, its variation as a function of time;

– the distribution of instantaneous accelerations of the signal with respect to time;

– the PSD.

1.5. Distribution of instantaneous values

1.5.1. *Probability density*

One of the objectives of the analysis of a random process is to determine the probability of finding extreme or peak values, or of determining the percentage of time that a random variable (acceleration, displacement, etc.) exceeds a given value [RUD 75]. Figure 1.2 shows a sample of a random signal with respect to time defined over duration T.

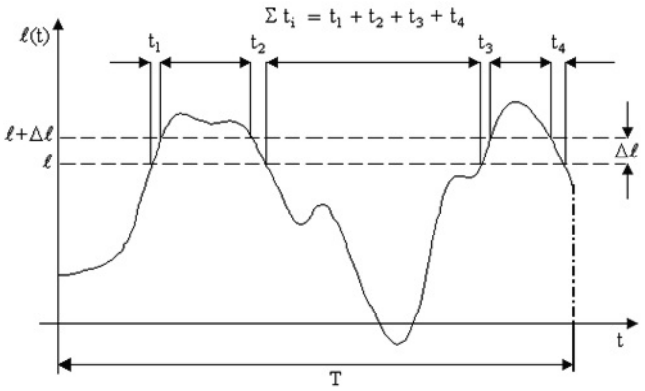


Figure 1.2. *Sample of random signal*

The probability that this function $\ell(t)$ is in the interval ℓ ($\ell + \Delta\ell$ being equal to the percentage of time during which it has values in this interval) is expressed mathematically as:

$$\text{prob}\left[\ell < \ell(t) < \ell + \Delta\ell\right] = \sum_i \frac{t_i}{T} \qquad [1.1]$$

If this interval $\Delta\ell$ is small, a density function probability $p(\ell)$ is defined by:

$$\text{prob}\left[\ell < \ell(t) < \ell + \Delta\ell\right] = p(\ell)\,\Delta\ell \qquad [1.2]$$

where:

$$p(\ell) = \frac{1}{T} \frac{\sum t_i}{\Delta \ell} \qquad [1.3]$$

To precisely define $p(\ell)$, it is necessary to consider very small intervals $\Delta\ell$ and of very long duration T, so that, mathematically, the probability density function is defined by:

$$p(\ell) = \lim_{\Delta \ell \to 0} \left[\lim_{T \to \infty} \left(\frac{1}{T} \frac{\sum t_i}{\Delta \ell} \right) \right] \qquad [1.4]$$

1.5.2. *Distribution function*

Owing to the fact that $p(\ell)$ was given for the field of values of $\ell(t)$, the probability that the signal is inside the limits $a < \ell(t) < b$ is obtained by integration from [1.2]:

$$\text{prob}[a < \ell(t) < b] = \int_a^b p(\ell)\, d\ell \qquad [1.5]$$

Since the probability that $\ell(t)$ within the limits $-\infty, +\infty$ is equal to 1 (absolutely certain event), it follows that

$$\int_{-\infty}^{+\infty} p(\ell)\, d\ell = 1 \qquad [1.6]$$

and the probability that ℓ exceeds a given level L is simply

$$\text{prob}[L \leq \ell(t)] = 1 - \int_L^{\infty} p(\ell)\, d\ell \qquad [1.7]$$

There are electronic equipment and calculation programs that make it possible to determine either *the distribution function* or *the probability density function* of the instantaneous values of a real random signal $\ell(t)$.

Among the mathematical laws representing the usual probability densities, we distinguish two that are particularly important in the field of random vibrations: Gauss's law (or Normal law) and Rayleigh's law.

1.6. Gaussian random process

A *Gaussian random process* $\ell(t)$ is such that the ensemble of the instantaneous values of $\ell(t)$ obeys a law of the form:

$$p[\ell(t)] = \frac{1}{s\sqrt{2\pi}} \exp\left\{-\frac{[\ell(t) - m]^2}{2s^2}\right\} \qquad [1.8]$$

where m and s are constants. The utility of the Gaussian law lies in the central limit theorem, which establishes that the sum of independent random variables follows a roughly Gaussian distribution whatever the basic distribution.

This is the case for many physical phenomena, for quantities which result from a large number of independent and comparable fluctuating sources, and in particular the majority of vibratory random signals encountered in the real environment [BAN 78], [CRE 56], [PRE 56a].

A Gaussian process is fully determined by knowledge of the mean value m (generally zero in the case of vibratory phenomena) and of the standard deviation s.

Moreover, it is shown that:

– if the excitation is a Gaussian process, the response of a linear time-invariant system is also a Gaussian process [CRA 83], [DER 80];

– the vibration, in part excited at resonance, tends to be Gaussian.

For a strongly resonant system subjected to broadband excitation, the central limit theorem makes it possible to establish that the response tends to be Gaussian even if the input is not. This applies when the excitation is not a white noise, provided that it is a broadband process covering the resonance peak [NEW 75] (provided that the probability density of the instantaneous values of the excitation does not have too significant an asymmetry [MAZ 54] and that the structure is not very strongly damped [BAN 78], [MOR 55]).

In many practical cases, we are thus led to conclude that the vibration is stationary and Gaussian, which simplifies the problem of calculation of the response of a mechanical system (Chapter 9) and, consequently, the simplification of fatigue

damage calculations, the possibility of simply evaluating the probability of exceeding a given value, etc.

The reduced variable $t = \frac{\ell(t) - m}{s}$ is sometimes used. The distribution function makes it possible to calculate the probability for the amplitude to be lower than a given value. For a Gaussian distribution, it is equal to:

$$F(L) = P(\ell < L) = \frac{1}{s\sqrt{2\pi}} \int_{-\infty}^{L} e^{-\frac{1}{2}\left(\frac{\ell - m}{s}\right)^2} d\ell \qquad [1.9]$$

It can also be written as:

$$F(T) = \frac{1}{2}\left[1 + E_1\left(\frac{T}{\sqrt{2}}\right)\right] \qquad [1.10]$$

if we say $T = \frac{L - m}{s}$ and if E_1 is the error function defined by:

$$E_1(x) = \frac{2}{\sqrt{\pi}} \int_0^L e^{-t^2} dt \qquad [1.11]$$

The interest of the relation between the distribution function and error function lies in the calculation possibility from a development in series, thus enabling us to avoid making an integration (Appendix, section A4.1).

Figure 1.3 provides an image of the way to estimate the probability density and distribution function of a signal $\ell(t)$. Signal amplitudes are divided into small $\Delta\ell$ intervals. In each one, we count the number of times the signal is located in each one of these intervals for all its duration. This number is transferred to a curve based on the average amplitude of interval $\Delta\ell$. The resulting diagram is no more than the histogram of instantaneous values of signal. When the amplitude interval becomes very small, this histogram tends toward the probability density of instantaneous values, which is generally close to a Gaussian distribution.

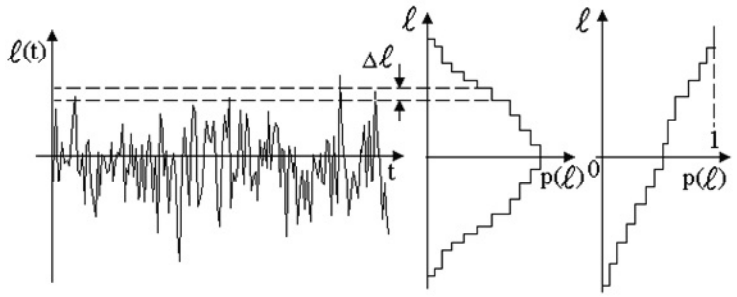


Figure 1.3. *Distribution of the instantaneous values of the signal*

Example 1.1.

Let us take a time history vibratory signal having a constant PSD equal to 1 $(m/s^2)^2$/Hz between 0 and 2,000 Hz (Figure 1.4.).

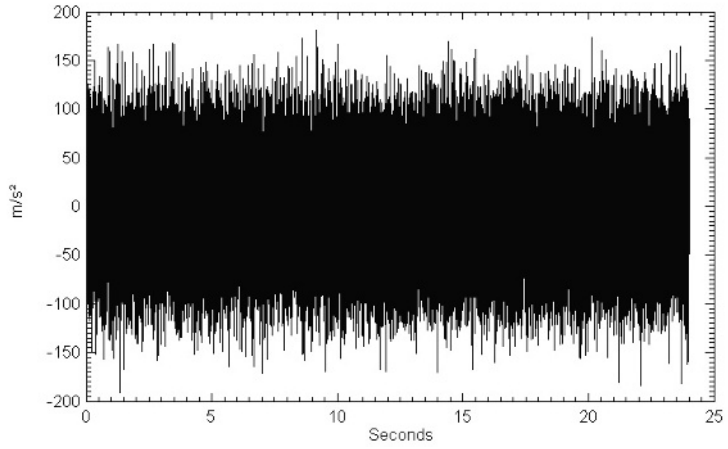


Figure 1.4. *Random vibration signal according to time (96,000 points, rms value: 44.72 m/s²)*

The histogram of the instantaneaous values, representing the number of values counted in a given amplitude class, is close to a Gaussian distribution, the difference being due to statistic dispersion; the result is all the closer as the duration of the signal sample analyzed is greater (24 seconds in this case). We

verify that the skewness (0.005) (see section 1.9.5) and kurtosis (3.0) (see section 1.9.6) are respectively close to zero and three (Figure 1.5).

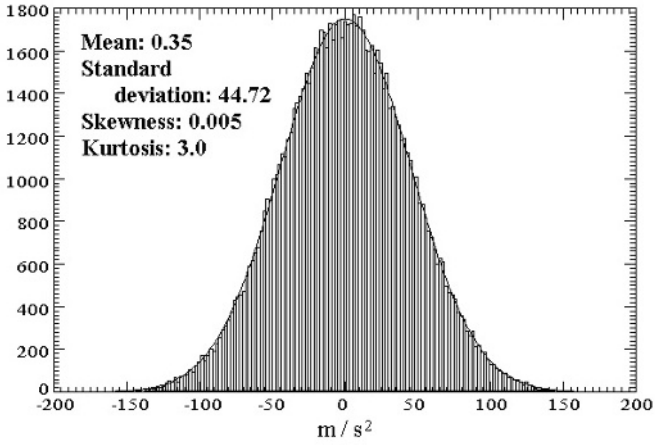


Figure 1.5. *Histogram of instantaneous values of signal of Figure 1.4 compared to the density of probability of a Gaussian distribution with the same mean and standard deviation*

Most vibrations (stationary phase) measured in the real environment are Gaussian.

NOTE.– *A signal with a varying rms value over time (and thus non-stationary) cannot be Gaussian. The standard deviation of instantaneous values (or rms value) is then seen as a constant in the definition of probability density of a Gaussian distribution.*

Example 1.2.– Signal presenting a transitory phase between two stationary phases

Consider a random vibration made up of three phases with the same duration (8 s), one stationary with an rms value equal to 44.72 m/s2, followed by a transitory phase during which the rms value decreases, in a linear way, until it reaches a second stationary phase with an rms value of 11.18 m/s2 (Figure 1.6).

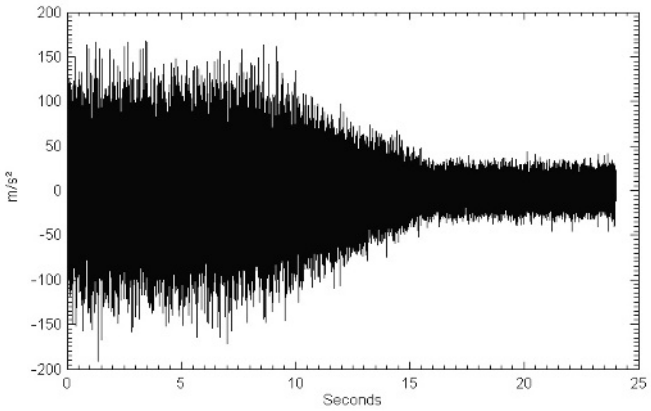


Figure 1.6. *Signal presenting a transitory phase between two stationary phases*

The instantaneous values histogram of this signal is traced in Figure 1.7, superposed to the probability density of a Gaussian distribution where the standard deviation is equal to the rms value of the complete signal (zero mean). This density is far from the histogram.

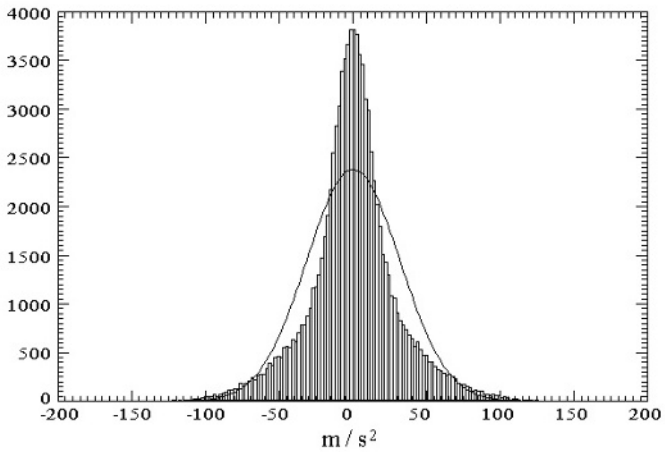


Figure 1.7. *Histogram of instantaneous values of signal of Figure 1.6 and Gaussian probability density calculated from the global rms value of this same signal (31.62 m/s²)*

1.7. Rayleigh distribution

Rayleigh distribution, of which the probability has the form

$$p(\ell) = \frac{\ell}{s^2} e^{-\frac{\ell^2}{2 s^2}} \qquad [1.12]$$

$(\ell \geq 0)$ is also an important law in the field of vibration for the representation of:

– variations in the instantaneous value of the envelope of a narrowband Gaussian random process;

– peak distribution in a narrowband Gaussian process.

Because of its very nature, the study of vibration would be very difficult if we did not have tools enabling the limitation of analysis of the complete process, which comprises a great number of signals varying with time and of very great duration, using a very restricted number of samples of reasonable duration. The study of statistical properties of the process will make it possible to define two very useful concepts with this objective in mind: stationarity and ergodicity.

1.8. Ensemble averages: through the process

1.8.1. *n order average*

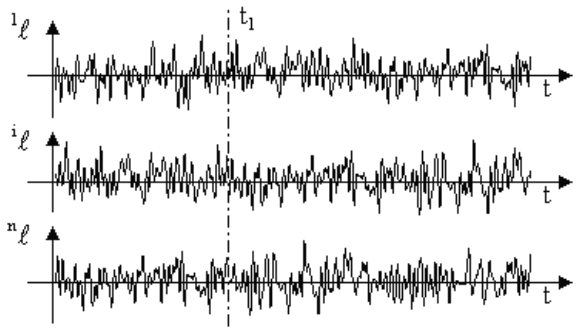


Figure 1.8. *"Through the process" study*

Let us consider N recordings of a random phenomenon varying with time $^i\ell(t)$ $[i \in (1, N)]$ for t varying from 0 to T (Figure 1.8). The ensemble of the curves $^i\ell(t)$

constitutes the process $\left\{ {}^{i}\ell(t) \right\}$. A first possibility may consist of studying the distribution of the values of ℓ for $t = t_1$ given [JAM 47].

If we have (N) records of the phenomenon, we can calculate, for a given t_1, the mean [BEN 62], [BEN 63], [DAV 58], [JEN 68]:

$$\overline{\ell(t)} = \frac{{}^{1}\ell(t_1) + {}^{2}\ell(t_1) + \cdots + {}^{n}\ell(t_1)}{N} \qquad [1.13]$$

If the values ${}^{i}\ell(t)$ belong to an infinite discrete ensemble, the *moment of order n* is defined by:

$$E\left[\ell^{n}(t_1)\right] = \lim_{N\to\infty} \sum_{i=1}^{N} \frac{{}^{i}\ell^{n}(t_1)}{N} \qquad [1.14]$$

($E[\ ]$ = mathematical expectation). By considering the ensemble of the samples at the moment t_1, the statistical nature of $\ell(t_1)$ can be specified by its probability density [LEL 73]:

$$p\left[\ell(t_1)\right] = \lim_{\Delta\ell\to 0} \frac{\text{Prob}\left[\ell \le \ell(t_1) \le \ell + \Delta\ell\right]}{\Delta\ell} \qquad [1.15]$$

and by the moments of the distribution:

$$E\left[\ell^{n}(t_1)\right] = \int_{-\infty}^{\infty} \ell^{n}(t_1)\, p\left[\ell(t_1)\right] d\ell(t_1) \qquad [1.16]$$

if the density $p\left[\ell(t_1)\right]$ exists and is continuous (or the distribution function). The moment of order 1 is the *mean* or *expected value*; the moment of order 2 is the *quadratic mean*.

For two random variables

The joint probability density is written:

$$p(\ell_1, t_1; \ell_2, t_2) = \lim_{\substack{\Delta\ell_1\to 0 \\ \Delta\ell_2\to 0}} \frac{\text{Prob}\left[\ell_1 \le \ell(t_1) \le \ell_1 + \Delta\ell_1;\ \ell_2 \le \ell(t_2) \le \ell_2 + \Delta\ell_2\right]}{\Delta\ell_1\, \Delta\ell_2} \qquad [1.17]$$

and joint moments:

$$E\left[\ell(t_1)\ \ell(t_2)\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} {}^{i}\ell(t_1)\ {}^{j}\ell(t_2)\ p\left[\ell(t_1), \ell(t_2)\right] d\ell(t_1)\ d\ell(t_2) \qquad [1.18]$$

1.8.2. *Centered moments*

The central moment of order n (with regard to the mean) is the quantity:

$$E\left\{\left[\ell(t_1) - m\right]^n\right\} = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N}\left[{}^{i}\ell(t_1) - m\right]^n \qquad [1.19]$$

in the case of a discrete ensemble and, for $p(\ell)$ continuous:

$$E\left\{\left[\ell(t_1) - m\right]^n\right\} = \int_{-\infty}^{\infty}\left[\ell(t_1) - m\right]^n\ p\left[\ell(t_1)\right] d\ell(t_1) \qquad [1.20]$$

1.8.3. *Variance*

The *variance* is the centered moment of order 2

$$s^2_{\ell(t_1)} = E\left\{\left[\ell(t_1) - m\right]^2\right\} \qquad [1.21]$$

By definition:

$$s^2_{\ell(t_1)} = \int_{-\infty}^{\infty}\left[\ell(t_1) - m\right]^2\ p\left[\ell(t_1)\right] d\ell(t_1) \qquad [1.22]$$

$$s^2_{\ell(t_1)} = \int_{-\infty}^{\infty} \ell^2(t_1)\ p\left[\ell(t_1)\right] d\ell(t_1)$$

$$-2\ m \underbrace{\int_{-\infty}^{\infty} \ell(t_1)\ p\left[\ell(t_1)\right] d\ell(t_1)}_{m} + m^2 \underbrace{\int_{-\infty}^{\infty} p\left[\ell(t_1)\right] d\ell(t_1)}_{1}$$

$$s^2_{\ell(t_1)} = E\left\{\left[\ell(t_1)\right]^2\right\} - 2\ m^2 + m^2$$

$$s^2_{\ell(t_1)} = E\left\{\left[\ell(t_1)\right]^2\right\} - m^2 \qquad [1.23]$$

1.8.4. *Standard deviation*

The quantity $s_{\ell(t_1)}$ is called the *standard deviation*. If the mean is zero,

$$s^2_{\ell(t_1)} = E\left\{\left[\ell(t_1)\right]^2\right\} \qquad [1.24]$$

When the mean m is known, an absolutely unbiased estimator of s^2 is $\sum \frac{\left({}^i\ell - m\right)^2}{N}$. When m is unknown, the estimator of s^2 is $\sum \frac{\left({}^i\ell - m'\right)^2}{N-1}$ where $m' = \frac{1}{N}\sum {}^i\ell$.

Example 1.3.

Let us consider 5 samples of a random vibration $\ell(t)$ and the values of ℓ at a given time $t = t_1$ (Figure 1.9).

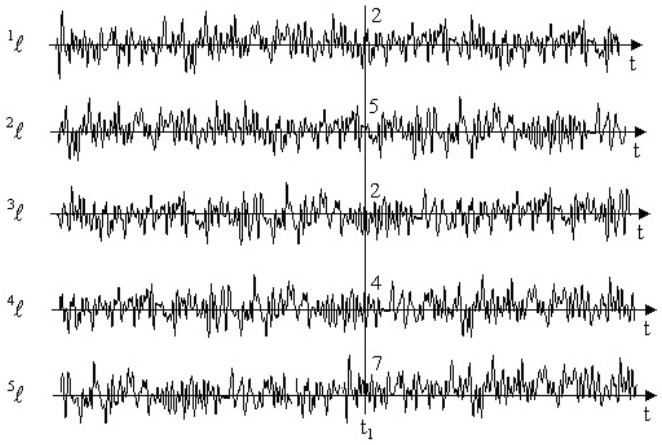


Figure 1.9. *Example of a stochastic process*

If the exact mean m is known ($m = 4.2$ m/s^2 for example), the variance is estimated from:

$$s^2 = \frac{(2-4.2)^2 + (5-4.2)^2 + (2-4.2)^2 + (4-4.2)^2 + (7-4.2)^2}{5} \ (\text{m/s}^2)^2$$

$$s^2 = \frac{18.2}{5} = 3.64 \ (\text{m/s}^2)^2$$

If the mean m is unknown, it can be evaluated from

$$m' = \frac{1}{N} \sum_i {}^i\ell(t_1) = \frac{2+5+2+4+7}{5} = \frac{20}{5} = 4 \ \text{m/s}^2$$

$$s^2 = \frac{18}{4} = 4.50 \ (\text{m/s}^2)^2$$

1.8.5. *Autocorrelation function*

Given a random process ${}^i\ell(t)$, the *autocorrelation function* is the function defined, in the discrete case, by:

$$R(t_1, t_1+\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_i {}^i\ell(t_1)\ {}^i\ell(t_1+\tau) \qquad [1.25]$$

$$R(t_1, t_1+\tau) = E\left[x(t_1).x(t_1+\tau)\right] \qquad [1.26]$$

or, for a continuous process, by:

$$R(\tau) = \int_{-\infty}^{\infty} x(t_1)\ x(t_1+\tau)\ p\left[x(t_1)\right] dx(t_1) \qquad [1.27]$$

1.8.6. *Cross-correlation function*

Given the two processes $\{\ell(t)\}$ and $\{u(t)\}$ (for example, the excitation and the response of a mechanical system), the *cross-correlation function* is the function:

$$R_{\ell u}(t_1, t_1 + \tau) = E\left[\ell(t_1) \,.\, u(t_1 + \tau)\right] \qquad [1.28]$$

or

$$R(\tau) = \lim_{N \to \infty} \frac{1}{N} \sum_{i}{}^{i} \ell(t_1) \,.\, {}^{i}u(t_1 + \tau) \qquad [1.29]$$

The *correlation* is a number measuring the degree of resemblance or similarity between two functions of the same parameter (time generally) [BOD 72].

1.8.7. *Autocovariance*

Autocovariance is the quantity:

$$C(t_1, t_1 + \tau) = E\left\{\left[\ell(t_1) - \overline{\ell(t_1)}\right]\left[\ell(t_1 + \tau) - \overline{\ell(t_1 + \tau)}\right]\right\} \qquad [1.30]$$

$$C(t_1, t_1 + \tau) = R(t_1, t_1 + \tau) - \overline{\ell(t_1)}\ \overline{\ell(t_1 + \tau)} \qquad [1.31]$$

$C(t_1, t_1 + \tau) = R(t_1, t_1 + \tau)$ if the mean values are zero.

We have in addition:

$$R(t_1, t_2) = R(t_2, t_1) \qquad [1.32]$$

1.8.8. *Covariance*

We define *covariance* as the quantity:

$$C_{\ell u} = E\left\{\left[\ell(t_1) - \overline{\ell(t_1)}\right]\left[u(t_1 + \tau) - \overline{u(t_1 + \tau)}\right]\right\} \qquad [1.33]$$

1.8.9. *Stationarity*

A phenomenon is *strictly stationary* if every moment of all orders and all the correlations are invariable with time t_1 [CRA 67], [JAM 47], [MIX 69], [PRE 90] [RAP 69], [STE 67].

The phenomenon is *wide-sense (*or *weakly) stationary* if only the mean, the mean square value and the autocorrelation are independent of time t_1 [BEN 58], [BEN 61b], [SVE 80].

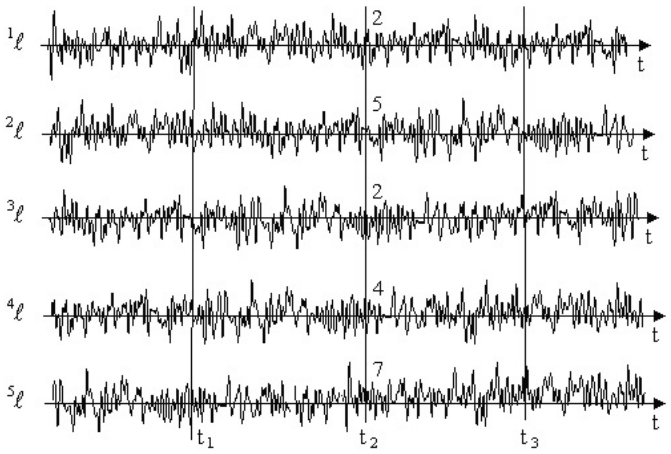


Figure 1.10. *Calculation of ensemble averages at different times*

Interest of stationarity

Since statistical process properties do not evolve over time in a stationary process, it is not necessary to record the signals for a long period of time. This time, however, must be long enough to subsequently enable a significant frequency analysis. In Chapter 4, we will see which rule must be respected.

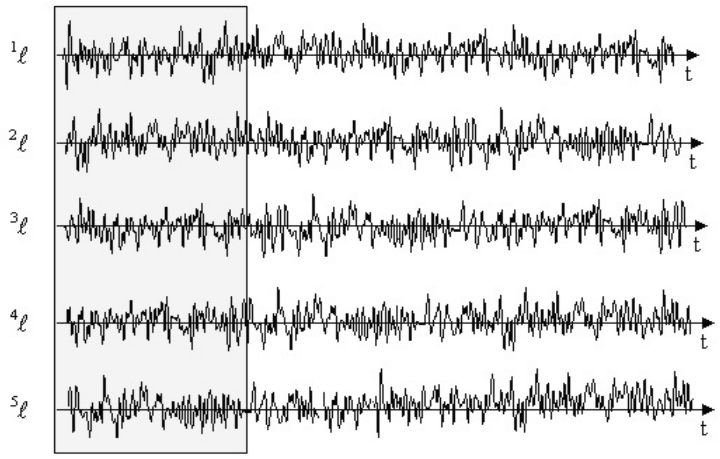


Figure 1.11. *Calculation of ensemble averages on a short duration if the process is stationary*

Example 1.4.

The following cases are completely unrealistic and are only illustrated for learning purposes.

CASE 1.–*The process is described by 5 measurements of a constant value signal equal to 1.*

Since all averages are equal regardless of moment t used for calculation, the process represented in Figure 1.12 is stationary.

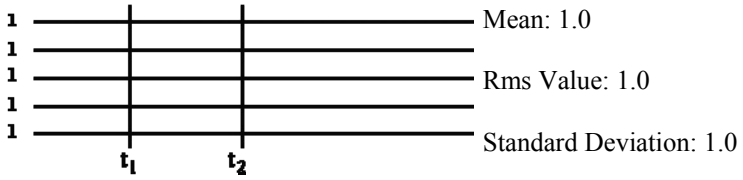


Figure 1.12. *Example of a stationary process*

CASE 2.– *The process is described by 5 measurements of signals with different constant values*

Means are independent of moment t chosen: the process is stationary (Figure 1.13).

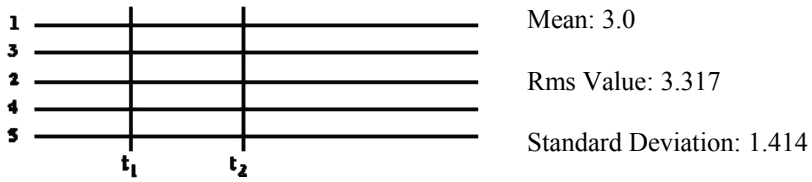


Figure 1.13. *Example of a stationary process*

CASE 3.– *The process is described by 4 signals with a constant value equal to 1 and by a sine curve*

The means in the process in Figure 1.14 vary according to the moment of calculation: the process is not stationary.

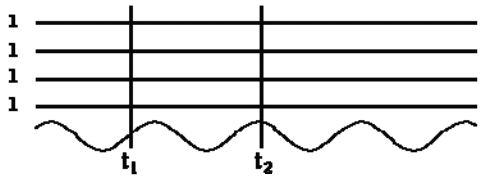


Figure 1.14. *Example of a non-stationary process*

Autostationarity

If only one recording of the phenomenon $\ell(t)$ is available, we sometimes define the autostationarity of the signal by studying the stationarity with n samples taken at various moments of the recording, by regarding them as samples obtained independently during n measurements (Figure 1.15).

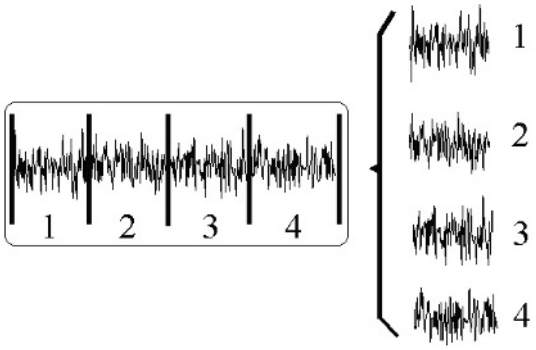


Figure 1.15. *Study of autostationarity*

We can also define strong autostationarity and weak autostationarity.

For a stationary process, the autocorrelation function is written:

$$R(\tau) = E\left[\ell(0)\ \ell(\tau)\right]$$

$$R(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} {}^{i}\ell(0)\ {}^{i}\ell(\tau) \qquad [1.34]$$

NOTES.–

Based on this assumption, we have:

$$R(-\tau) = E\{\ell(0)\,\ell(-\tau)\}$$

$$R(-\tau) = E\{\ell(\tau)\,\ell(0)\}$$

$$R(-\tau) = R(\tau) \qquad [1.35]$$

(R is an even function of τ) [PRE 90].

$$R(0) = E\{\ell(0)\,\ell(0)\} = E\{\ell^2(t)\} \qquad [1.36]$$

$R(0)$ *is the ensemble mean square value at the arbitrary time t.*

– $R(0) \geq |R(\tau)|$

We have

$$E\left\{\left[\ell(0) \pm \ell(\tau)\right]^2\right\} \geq 0$$

yielding

$$E\{\ell^2(0)\} \pm 2\,E\{\ell(0)\,\ell(\tau)\} + E\{\ell^2(\tau)\} \geq 0$$

$$R(0) \pm 2\,R(\tau) + R(0) \geq 0$$

and

$$R(0) \geq |R(\tau)| \qquad [1.37]$$

As for the cross-correlation function, it becomes, for a stationary process,

$$R_{\ell u}(\tau) = E\{\ell(0)\ \ell(\tau)\} = \lim_{N\to\infty} \frac{\sum_{i=1}^{N} {}^{i}\ell(0)\ {}^{i}\ell(\tau)}{N} \qquad [1.38]$$

Properties

1. $$R_{\ell u}(-\tau) = R_{u\ell}(\tau) \qquad [1.39]$$

Indeed

$$R_{\ell u}(-\tau) = E\{\ell(0)\ u(-\tau)\}$$

$$R_{\ell u}(-\tau) = E\{\ell(\tau)\ u(0)\}$$

$$R_{\ell u}(-\tau) = E\{u(0)\ \ell(\tau)\}$$

$$R_{\ell u}(-\tau) = R_{u\ell}(\tau)$$

2. Whatever τ

$$R_{\ell u}(\tau) \le \sqrt{R_{\ell}(0)\ R_{u}(0)} \qquad [1.40]$$

Cyclostationary process

A random process $\{x(t)\}$ is *cyclostationary* if its random properties repeat periodically over the course of time. At order 2, this means that the correlation

$$R(\tau, t) = E\left[\ell(t)\,\ell(t-\tau)\right]$$

is periodic in t : $R(\tau, t) = R(\tau, t+T)$.

An important example is that of a signal modulated by a linear modulation. If T is the period, the average power of the process is defined by

$$P_{moy} = \lim_{T\to\infty} \frac{1}{2T} E\left[\int_{-T}^{T} |x(t)|^2\ dt\right] \qquad [1.41]$$

As the process is cyclostationary, $E\left[\left|x(t)\right|^2\right] = R(0,t)$ is periodic of period T, and thus

$$P_{moy} = \frac{1}{T}\int_{(T)} E\left[\left|x(t)\right|^2\right] dt \qquad [1.42]$$

where (T) designates any interval of period T. The average power is thus obtained by averaging $E\left[\left|x(t)\right|^2\right]$ over a period.

1.9. Temporal averages: along the process

1.9.1. *Mean*

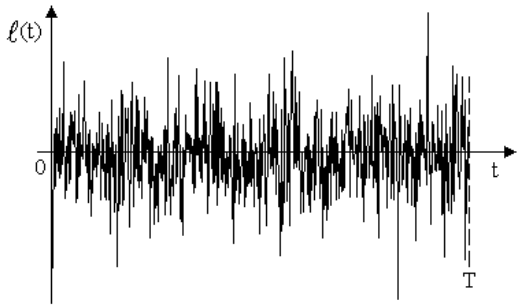


Figure 1.16. *Sample of a random signal*

Let us consider a sample $\ell(t)$ of duration T of a recording. It can be interesting to study the statistical properties of the instantaneous values of the function $\ell(t)$. The first possibility is to consider the temporal mean of the instantaneous values of the recording.

We have:

$$\overline{\ell(t)} = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \ell(t)\, dt \qquad [1.43]$$

if this limit exists. This limit may very well not exist for some or all the samples and, if it exists, it may depend on the selected sample $\ell(t)$; however, it does not depend on time[1].

For practical reasons, we in fact calculate the mean value of the signal $\ell(t)$ over one finite duration T:

$$\overline{\ell(t)} = \frac{1}{T}\int_0^T \ell(t)\ dt \qquad [1.44]$$

The mean value is related to the difference between the positive and negative areas ranging between the curve $\ell(t)$ and the time axis [GRE 81].

The mean m of a centered signal is zero, so this parameter cannot be used by itself to correctly evaluate the severity of the excitation.

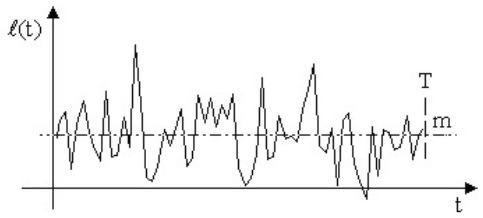


Figure 1.17. *Random vibration with non-zero mean*

The mean value is equal to the absolute value of the parallel shift of the Ot axis necessary to cancel out this difference. A signal $\ell(t)$ of mean m can be written:

$$\ell(t) = m + \ell^*(t) \qquad [1.45]$$

where $\ell^*(t)$ is a centered signal. This mean value is generally a static component which can be due to the weight of the structure, to the maneuverings of an aircraft,

1. We also define $\overline{x(t)}$ from:

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)\ dt \quad \text{or} \quad \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t)\ dt$$

to the thrust of a missile in phase propulsion, etc. In practice, we often consider this mean to be zero.

1.9.2. *Quadratic mean – rms value*

The vibration $\ell(t)$ generally results in an oscillation of the mechanical system around its equilibrium position, so that the arithmetic mean of the instantaneous values can be zero if the positive and negative values are compensated. The arithmetic mean represents the signal poorly [RAP 69], [STE 67]. Therefore, it is sometimes preferred to calculate the mean value of the absolute value of the signal

$$\overline{|\ell(t)|} = \frac{1}{T}\int_0^T |\ell(t)|\, dt \qquad [1.46]$$

and, much more generally, by analogy with the measurement of the rms value of an electrical quantity, the *quadratic mean* (or *mean square value*) of the instantaneous values of the signal of which the square root is the *rms value*.

The rms value $\ell_{rms} = \sqrt{\overline{\ell^2(t)}}$ is the simplest statistical characteristic to obtain. It is also most significant since it provides an order of magnitude of the intensity of the random variable.

If we can analyze the curve $\ell(t)$ by dividing the sample of duration T into N intervals of duration Δt_i $\left(i \in [1, N]\right)$, and if ℓ_i is the value of the variable during the time interval Δt_i, the mean quadratic value is written:

$$\overline{\ell^2} = \frac{\ell_1^2\ \Delta t_1 + \cdots + \ell_i^2\ \Delta t_i + \cdots + \ell_N^2\ \Delta t_N}{T} \qquad [1.47]$$

with $T = \sum_{i=1}^{N} \Delta t_i$. If the intervals of time are equal to (Δt) and if N is the number of points characterizing the signal, $T = N\ \Delta t$ and:

$$\ell_{rms} = \frac{1}{N}\sum_i \ell_i^2 \qquad [1.48]$$

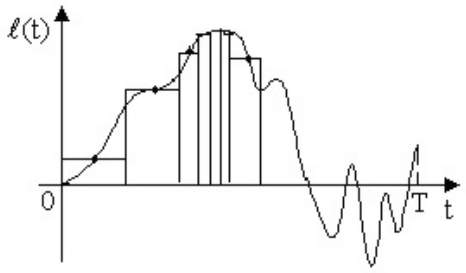


Figure 1.18. *Approximation of the signal*

If all Δt_i tend towards zero and if $N \to \infty$, the quadratic mean is defined by [BEN 63]:

$$\overline{\ell^2(t)} = \frac{1}{T}\int_0^T \ell^2(t)\, dt \qquad [1.49]$$

(or by $\frac{1}{2T}\int_{-T}^{T} \ell^2(t)\, dt$).

Two signals having very different frequency contents, corresponding to very dissimilar temporal forms, can have the same mean quadratic value. In this expression, the rms value takes into account the totality of the frequencies of the signal.

Example 1.5.

1. Consider a signal is defined by 11 points (Figure 1.19).

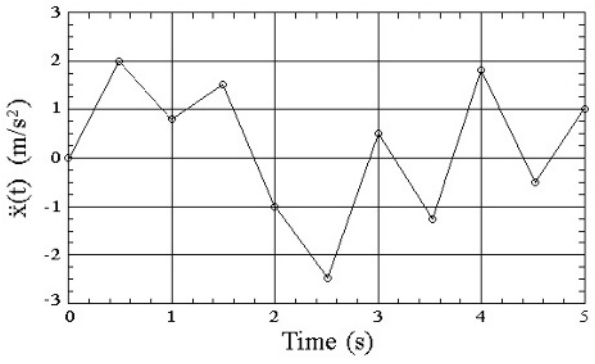


Figure 1.19. *Signal of acceleration*

The use of relation [1.48] leads to:

$$\ddot{x}_{rms} = \sqrt{\frac{0^2 + 2^2 + 0.8^2 + 1.5^2 + (-1)^2 + (-2.5)^2 + 0.5^2 + (-1.25)^2 + 1.8^2 + (-0.5)^2 + 1^2}{11}}$$

or: $\ddot{x}_{rms} \approx 136$.

2. Let us consider a sinusoid $\ddot{x}(t) = \ddot{x}_m \sin(\Omega t + \varphi)$

$$\overline{\ddot{x}(t)} = 0$$

$$s^2 = \overline{\ddot{x}^2(t)} = \frac{\ddot{x}_m^2}{2}$$

$$\overline{|\ddot{x}(t)|} = \frac{2}{\pi} \ddot{x}_m \left(= \frac{2}{\pi}\sqrt{2}\ s \approx 0.9\ s \right)$$

(for a normal distribution, $\overline{|\ddot{x}(t)|} \approx 0.798\ s$).

1.9.3. *Moments of order n*

As in the preceding section, we also define:

– moments of an order higher than 2; the *moment of order n* is expressed:

$$E\left\{\ell^n(t)\right\} = \overline{\ell^n(t)} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \ell^n(t)\ dt \qquad [1.50]$$

– *centered moments*: measured vibratory signals usually have a zero mean value. We then call the signal *centered*. When that is not the case, the rms value is still calculated as in section 1.9.2. We can also look at centered moments of order n defined by:

$$\mu_n = E\left\{\left[\ell(t) - \overline{\ell(t)}\right]^n\right\} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \left[\ell(t) - \overline{\ell(t)}\right]^n\ dt \qquad [1.51]$$

For a signal made up of N points of mean $\bar{\ell}$:

$$\mu_n = \frac{1}{N}\sum_{i=1}^{N}\left(\ell_i - \bar{\ell}\right)^n \qquad [1.52]$$

1.9.4. *Variance – standard deviation*

The rms value is calculated from the mean quadratic value of the instantaneous values of the signal. The centered moment of order 2 is the *variance,* denoted by s_ℓ^2:

$$s_\ell^2 = E\left[\left(\ell - \bar{\ell}\right)^2\right] = \overline{\ell^2(t)} - \overline{\ell(t)}^2 \qquad [1.53]$$

s_ℓ is called the *standard deviation*. This parameter characterizes the dispersion of the signal around its mean.

For a signal defined by N points:

$$s_\ell^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\ell_i - \bar{\ell}\right)^2 \qquad [1.54]$$

If the mean m is zero, the standard deviation s is equal to the rms value of the signal $\ell(t)$.

NOTE.– *On the assumption of zero mean, we can however note a difference between the standard deviation and the rms value when the latter is calculated starting from the PSD, which does not necessarily cover all of the frequency contents of the signal, in particular beyond 2,000 Hz (a value often selected as the upper limit of the analysis band). The rms value is then lower than the standard deviation. The comparison of the two values makes it possible to evaluate the importance of the neglected range*.

Non biased estimator

Relationship [1.54] gives a biased measurement of the variance. The bias can be corrected using the relation:

$$s_\ell^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(\ell_i - \bar{\ell}\right)^2 \qquad [1.55]$$

Example 1.6.

Using the values of Example 1.5, the mean m value is equal to:

$$m = \sqrt{\frac{0+2+0.8+1.5-1-2.5+0.5-1.25+1.8-0.5+1}{11}}$$

$$m \approx 0.214$$

The variance is given by:

$$V = \sqrt{\frac{(0-0.214)^2+(2-0.214)^2+(0.8-0.214)^2+\cdots+(-0.5-0.214)^2+(1-0.214)^2}{11}}$$

$$V \approx 1.813$$

Standard deviation: $s = \sqrt{V} \approx 1.346$. We can verify that $\ddot{x}_{rms}^2 = s^2 + m^2$.

Non-biased standard deviation: ≈ 1.412.

1.9.5. *Skewness*

The centered moment of order 3, denoted by μ_3, is sometimes reduced by division by s_ℓ^3:

$$\mu_3' = \frac{E\left\{\left[\ell(t) - \overline{\ell(t)}\right]^3\right\}}{s_\ell^3} \qquad [1.56]$$

We can show [GMU 68] that μ_3' is characteristic of the symmetry of the probability density law $p(\ell)$ with regard to the mean $\overline{\ell(t)}$; for this reason, μ_3' is sometimes called *skewness*.

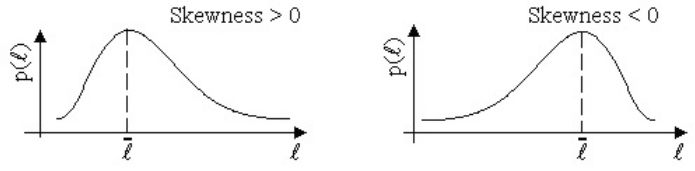


Figure 1.20. *Probability densities with non-zero skewness*

For a signal that is made up of N points:

$$\mu'_3 = \frac{\sum_{i=1}^{N} \left(\ell_i - \bar{\ell} \right)^3}{N \, s_\ell^3} \qquad [1.57]$$

$\mu'_3 = 0$ characterizes a normal process.

For $\mu'_3 > 0$, the probability density curve presents a peak towards the left and for $\mu'_3 < 0$, the peak of the curve is shifted towards the right.

Non-biased estimator of skewness

Definition [1.57] is a biased measurement of the asymmetry of the population. A non-biased estimator of the asymmetry is given by the relation:

$$\mu'_3 = \frac{N}{(N-1)(N-2)} \frac{\sum_{i=1}^{N} \left(\ell_i - \bar{\ell} \right)^3}{s_\ell^3} \qquad [1.58]$$

where $\bar{\ell}$ and s_ℓ are the non-biased estimators of the average and the standard deviation.

1.9.6. *Kurtosis*

The centered moment of order 4, reduced by division by s_ℓ^4, is also sometimes considered, as it makes it possible to estimate the flatness of the probability density curve. This is often called *kurtosis* [GUE 80].

$$\mu'_4 = \frac{E\left\{ \left[\ell(t) - \overline{\ell(t)} \right]^4 \right\}}{s_\ell^4} \qquad [1.59]$$

For a signal made up of N points:

$$\mu'_4 = \frac{\sum_{i=1}^{N} \left(\ell_i - \bar{\ell} \right)^4}{N \, s_\ell^4} \qquad [1.60]$$

The kurtosis characterizes the relative importance of the major distribution values in relations to the values close to zero:

$\mu'_4 = 3$ for a normal process.

$\mu'_4 < 3$ characteristic of a truncated signal or existence of a sinusoidal component ($\mu'_4 = 1.5$ for a pure sine).

$\mu'_4 > 3$ presence of peaks of high value (more than in the normal case).

We sometimes use a normalized expression of the kurtosis in the form

$$\gamma_4 = \frac{\sum_{i=1}^{N} \left(\ell_i - \bar{\ell} \right)^4}{N \, s_\ell^4} - 3 \qquad [1.61]$$

The quantity γ_4 is called "excess kurtosis". γ_4 is positive if the probability density is more acute than that of a normal law and γ_4 is negative if it is flatter.

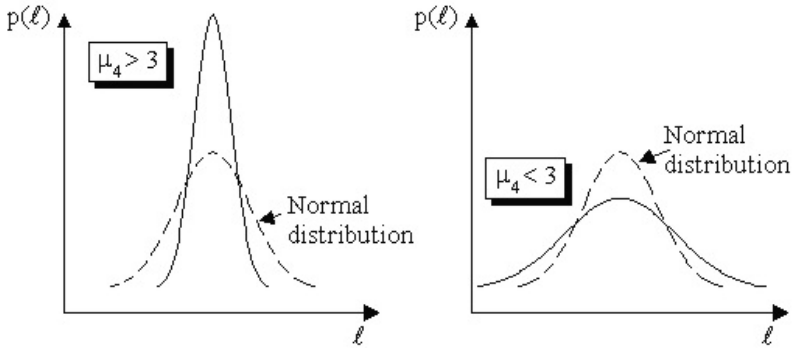


Figure 1.21. *Kurtosis influence on probability density*

Non-biased estimator of kurtosis

Definition [1.60] is a biased measurement of the flattening of the population. A non-biased estimator of this parameter is given by the relation:

$$\mu'_4 = \frac{(N+1)\,N}{(N-1)\,(N-2)\,(N-3)} \frac{\sum_{i=1}^{N}\left(\ell_i - \bar{\ell}\right)^4}{s_\ell^4} + 3\,\frac{(N-2)\,(N-3) - (N-1)^2}{(N-2)\,(N-3)} \qquad [1.62]$$

where $\bar{\ell}$ and s_ℓ are the non-biased estimators of the average and the standard deviation.

The normalized form of the kurtosis can be estimated without bias by:

$$\mu'_4 = \frac{(N+1)\,N}{(N-1)\,(N-2)\,(N-3)} \frac{\sum_{i=1}^{N}\left(\ell_i - \bar{\ell}\right)^4}{s_\ell^4} - 3\,\frac{(N-1)^2}{(N-2)\,(N-3)} \qquad [1.63]$$

Example 1.7.

With the information from Example 1.5, we obtain:

$$As = \frac{1}{(1.346)^3}\,\frac{(0-0.214)^3 + (2-0.214)^3 + \cdots + (-0.5-0.214)^3 + (1-0.214)^3}{11}$$

$$Ap = \frac{1}{(1.346)^4}\,\frac{(0-0.214)^4 + (2-0.214)^4 + \cdots + (-0.5-0.214)^4 + (1-0.214)^4}{11}$$

yielding As ≈ – 0.475 and Ap ≈ 2.241.

Non-biased skewness: ≈ – 0.553

Non-biased kurtosis: ≈ 2.568.

The advantage of these parameters

The rms value provides information on the global severity of the vibration. This information is useful, but not quite useful enough since it does not indicate the energy distribution in the frequency domain and thus potential vibration effects based on the natural frequencies of a mechanical structure.

Since measured vibration averages are mostly zero, the standard deviation is generally equal to the rms value. Comparing these two parameters makes it possible to verify this condition and to evaluate the value of the average when it is not zero.

Skewness and kurtosis are two parameters for verifying that the analyzed signal has a Gaussian instantaneous value distribution. When that is the case (the most common), skewness is theoretically equal to zero and kurtosis to 3. Strong variations of these parameters (particularly of kurtosis) calculated from sliding means also make it possible to detect the presence of mechanical shocks and signal problems (see section 1.16).

1.9.7. *Crest Factor*

The *crest factor* (or *peak-to-average* ratio) is equal to the ratio between the largest peak of the signal (in absolute value) and its rms value.

$$c_f = \frac{\max\left[\left|\ell(t)\right|\right]}{\sigma_\ell}$$

For a sinusoid, the crest factor is equal to $\sqrt{2}$. Random signals have an undefined crest factor which is extremely large, which may be expected on rare occasions very much larger than its rms value.

Each peak is associated with a probability. It is thus always possible, in theory, to find a very large peak whose probability of occurrence is very small. In practice, if the signal is Gaussian, we can observe crest factor values in the order of 5 to 6.

1.9.8. *Temporal autocorrelation function*

We define in the time domain the autocorrelation function $R_\ell(\tau)$ of the calculated signal, for a given τ delay, of the product $\ell(t)\ \ell(t+\tau)$ [BEA 72], [BEN 58], [BEN 63], [BEN 80], [BOD 72], [JAM 47], [MAX 65], [RAC 69], [SVE 80].

Figure 1.22. *Sample of a random signal*

$$R_\ell(\tau) = E\left[\ell(t)\ \ell(t+\tau)\right] \qquad [1.64]$$

$$R_\ell(\tau) = \lim_{T\to\infty} \frac{1}{2\,T} \int_{-T}^{T} \ell(t)\ \ell(t+\tau)\ dt \qquad [1.65]$$

The result is independent of the selected signal sample i. The *delay* τ being given, we thus create, for each value of t, the product $\ell(t)$ and $\ell(t+\tau)$ and we calculate the mean of all the products thus obtained. The function $R_\ell(\tau)$ indicates the influence of the value of ℓ at time t on the value of the function ℓ at time $t+\tau$. Indeed let us consider the mean square of the variation between $\ell(t)$ and $\ell(t+\tau)$, i.e. $E\left\{\left[\ell(t)-\ell(t+\tau)\right]^2\right\}$, equal to:

$$E\left\{\left[\ell(t)-\ell(t+\tau)\right]^2\right\} = E\left[\ell^2(t)\right] + E\left[\ell^2(t+\tau)\right] - 2\ E\left[\ell(t).\ell(t+\tau)\right]$$

$$E\left\{\left[\ell(t)-\ell(t+\tau)\right]^2\right\} = 2\ R_\ell(0) - 2\ R_\ell(\tau) \qquad [1.66]$$

We note that the weaker the autocorrelation $R_\ell(\tau)$, the greater the mean square of the difference $\left[\ell(t)-\ell(t+\tau)\right]$ and, consequently, the less $\ell(t)$ and $\ell(t+\tau)$ resemble each other.

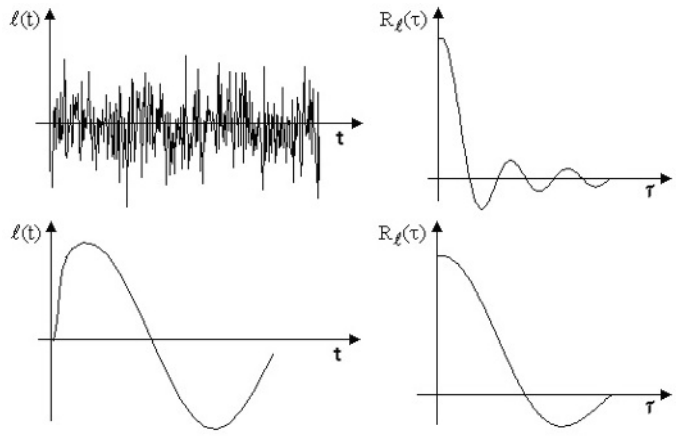


Figure 1.23. *Examples of autocorrelation functions*

The autocorrelation function measures the correlation between two values of $\ell(t)$ considered at different times t. If R_ℓ tends towards zero quickly when τ becomes large, the random signal probably fluctuates quickly and contains high frequency components.

If R_ℓ tends slowly towards zero, the changes in the random function are probably very slow [BEN 63], [BEN 80], [RAC 69].

R_ℓ is thus a measurement of the degree of random fluctuation of a signal.

Autocovariance

When the signal studied has a non-zero $\overline{\ddot{x}}$ average, we sometimes use the *autocovariance function* defined in an analog way by:

$$C_{\ddot{x}}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \left[\ddot{x}(t) - \overline{\ddot{x}}\right]\left[\ddot{x}(t+\tau) - \overline{\ddot{x}}\right] dt \qquad [1.67]$$

Autocovariance is connected to autocorrelation by:

$$C_{\ddot{x}}(\tau) = R_{\ddot{x}}(\tau) - \overline{\ddot{x}}^2 \qquad [1.68]$$

Discrete form

The autocorrelation function calculated for a sample of signal digitized with N points separated by Δt is equal, for $\tau = m\,\Delta t$, to [BEA 72]:

$$R_\ell(\tau) = \frac{1}{N-m} \sum_{i=1}^{N-m} \ell_i \cdot \ell_{i+m} \qquad [1.69]$$

Catalogs of correlograms exist allowing typological study and facilitating the identification of the parameters characteristic of a vibratory phenomenon [VIN 72]. Their use makes it possible to analyze, with some care, the composition of a signal (white noise, narrowband noise, sinusoids, etc.).

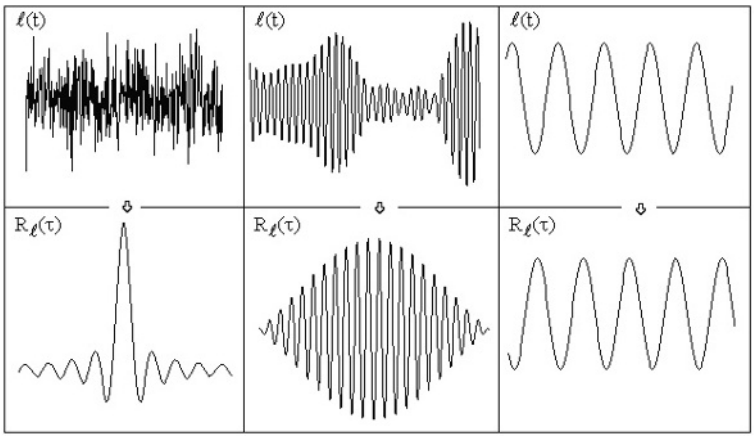


Figure 1.24. *Examples of autocorrelation functions*

Example 1.8.

Wideband noise autocorrelation

In the examples that follow, we often use a random vibration defined by a signal lasting 1 s (2,600 points), by a 31.6 m/s^2 rms value and by constant PSD equal to 1 $(m/s^2)^2/Hz$ between 1 Hz and 1,000 Hz (Figure 1.25).

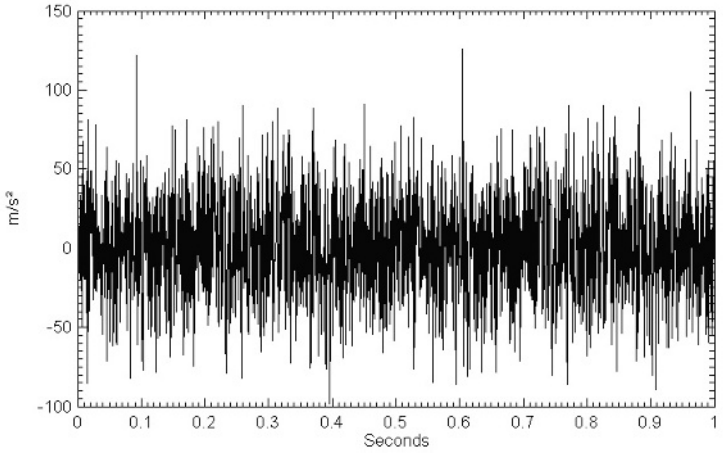


Figure 1.25. *Wideband random vibration*

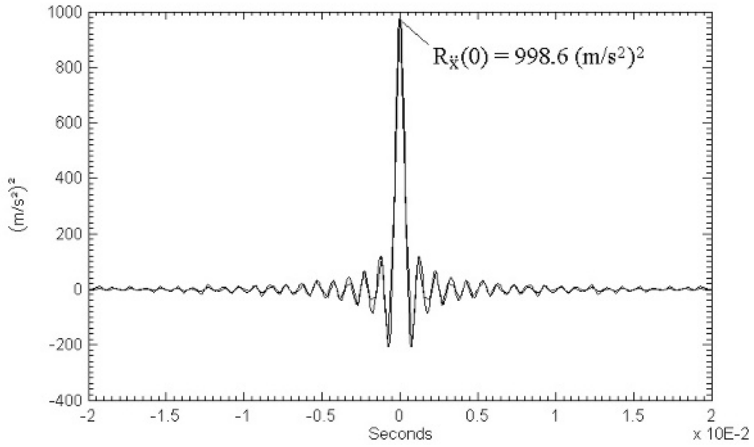


Figure 1.26. *Autocorrelation function of wideband noise between 1 Hz and 1,000 Hz, calculated from the signal according to time and its PSD*

Figure 1.26 illustrates the autocorrelation function of this vibration calculated from the signal and its PSD.

The two curves are very close, the difference being due to the number of points used for its line. This number is equal to that of the original signal in the first case, whereas it can be greater in the second case, leading to a smoother curve.

Autocorrelation shows a peak at origin with an amplitude that is equal to the square of the vibration rms value ($31.6^2 = 998.6\ (m/s^2)^2$). Since the noise is wideband, the curve quickly tends toward zero.

Example 1.9.

Narrowband noise autocorrelation

This same time history signal was used to calculate the response of a linear one-degree-of-freedom system ($f_0 = 300$ Hz, $Q = 10$) and its autocorrelation was calculated.

The response is a narrowband noise with an rms value equal to 65.55 m/s^2. The autocorrelation of the response does not tend toward zero as quickly as a

wideband noise. The peak at origin remains equal to the square of the signal's rms value.

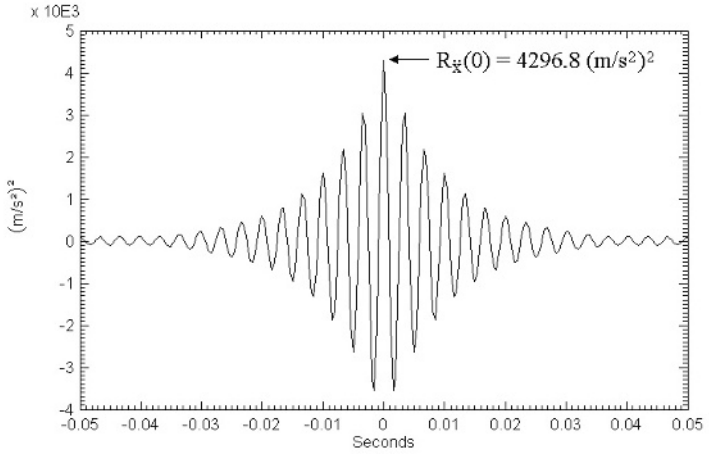


Figure 1.27. *Autocorrelation of the response of a linear one-degree-of-freedom system (300 Hz, Q = 10)*

Calculation of the autocorrelation function of a sinusoid

$$\ell(t) = \ell_m \sin(\Omega t) \qquad [1.70]$$

$$R_\ell(\tau) = \frac{1}{T}\int_0^T \ell_m \sin \Omega t \sin \Omega(t+\tau)\, dt$$

$$R_\ell(\tau) = \frac{\ell_m^2}{2}\cos \Omega\tau \qquad [1.71]$$

The correlation function of a sinusoid of amplitude ℓ_m and angular frequency Ω is a cosine of amplitude $\frac{\ell_m^2}{2}$ and pulsation Ω. The amplitude of the sinusoid can thus, conversely, be deduced from the autocorrelation function:

$$\ell_m = \sqrt{2}\left[R_\ell(\tau)\right]_{max} \qquad [1.72]$$

Example 1.10.

Autocorrelation of a sinusoid

Let us take a sinusoidal vibration with 20 m/s^2 amplitude and 300 Hz frequency. The autocorrelation of this signal (Figure 1.28) is a 200 $(m/s^2)^2$ amplitude and 300 Hz frequency sinusoid.

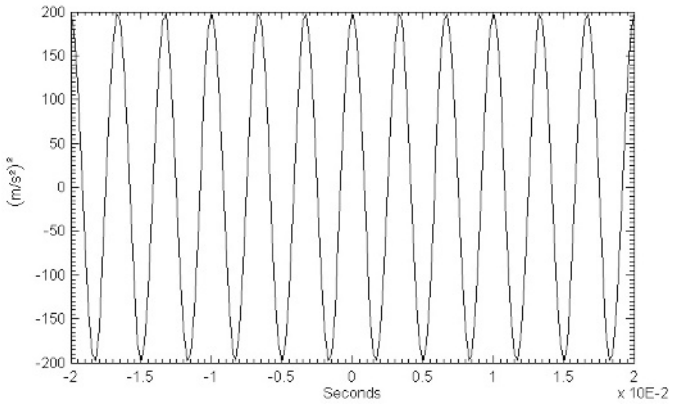


Figure 1.28. *Autocorrelation of a sinusoidal vibration*

1.9.9. *Properties of the autocorrelation function*

1. $R_\ell(0) = E\left[\ell^2(t)\right] = \overline{\ell^2(t)}$ = quadratic mean

$$R_\ell(0) = s^2 + \bar{\ell}^2 \qquad [1.73]$$

For a centered signal $\left(\bar{\ell} = 0\right)$, the ordinate at the origin of the autocorrelation function is equal to the variance of the signal.

2. The autocorrelation function is even [BEN 63], [BEN 80], [RAC 69]:

$$R_\ell(\tau) = R_\ell(-\tau) \qquad [1.74]$$

3. $\left|R_\ell(\tau)\right| < R_\ell(0) \quad \forall\ \tau$ [1.75]

If the signal is centered, $R_\ell(\tau) \to 0$ when $\tau \to \infty$. If the signal is not centered, $R_\ell(\tau) \to \bar{\ell}^2$ when $\tau \to \infty$.

4. It is shown that:

$$\frac{dR_\ell(\tau)}{d\tau} = E\left[\ell(t-\tau)\,\dot{\ell}(t)\right] \qquad [1.76]$$

$$\frac{d^2R_\ell(\tau)}{d\tau^2} = -E\left[\dot{\ell}(t)\,\dot{\ell}(t+\tau)\right] \qquad [1.77]$$

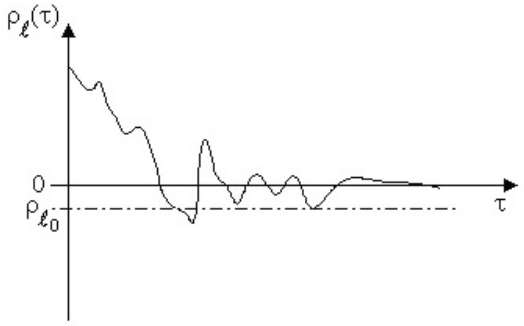


Figure 1.29. *Correlation coefficient*

NOTES.–

1. *The autocorrelation function, normalized autocorrelation function [BOD 72] or correlation coefficient is sometimes expressed in the reduced form:*

$$\rho_\ell(\tau) = \frac{R_\ell(\tau)}{R_\ell(0)} \qquad [1.78]$$

$\rho_\ell(\tau)$ *varies between –1 and +1,*

$\rho_\ell = 1$ *if the signals are identical (superimposable),*

$\rho_\ell = -1$ *if the signals are identical in absolute value and of opposite sign.*

2. *If the mean m is not zero, the correlation coefficient is given by*

$$\rho_{\ell}(\tau) = \frac{R_{\ell}(\tau) - m^2}{R_{\ell}(0)} \qquad [1.79]$$

1.9.10. *Correlation duration*

Correlation duration of a signal is the term given to the value τ_0 of τ from which the reduced autocorrelation function ρ_{ℓ} is always lower, in absolute value, than a certain value ρ_{ℓ_0}.

The correlation duration of:

– a wideband noise is weak;

– a narrowband noise is large; in extreme cases, a sinusoidal signal, which is thus deterministic, has an infinite correlation duration.

This last remark is sometimes used to detect in a signal $\ell(t)$ a sinusoidal wave $s(t) = S \sin \Omega t$ embedded in a random noise $b(t)$:

$$\ell(t) = s(t) + b(t) \qquad [1.80]$$

The autocorrelation is written:

$$R_{\ell}(\tau) = R_S(\tau) + R_b(\tau) \qquad [1.81]$$

If the signal is centered, for sufficiently large τ, $R_b(\tau)$ becomes negligible so that:

$$R_{\ell}(\tau) = R_S(\tau) = \frac{S^2}{2} \cos \Omega \tau \qquad [1.82]$$

This calculation makes it possible to detect a sinusoidal wave of low amplitude embedded in a very significant noise [SHI 70a].

Examples of application of the correlation method are as follows [MAX 69]:

– determination of the dynamic characteristics of a system;

– extraction of a periodic signal embedded in a noise;

– detection of periodic vibrations of a vibratory phenomenon;

– study of transmission of vibrations (cross-correlation between two points of a structure);

– study of turbulences;

– calculation of PSDs [FAU 69];

– more generally, applications in the field of signal processing, in particular in medicine, astrophysics, geophysics, etc. [JEN 68].

Example 1.11.

A sinusoidal vibration with a frequency of 300 Hz and rms value of 10 m/s^2 was superimposed on the random vibration in Figure 1.25 (rms value 31.6 m/s^2). Because of the low rms value of the sinusoid, drowning in random noise and almost undetectable in the global signal (Figure 1.28), the rms value of the composite signal is not much different (33.16 m/s^2).

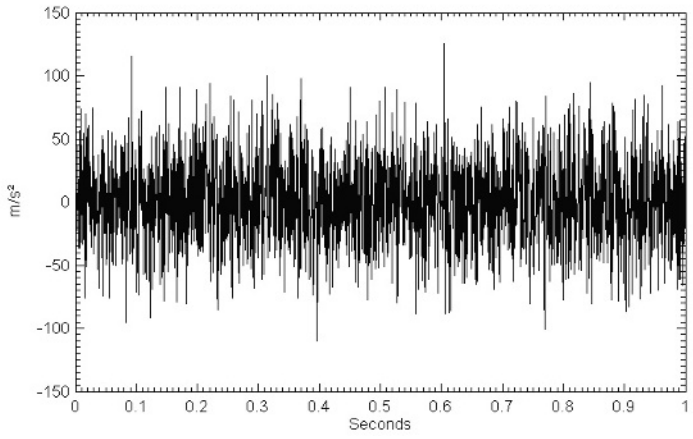


Figure 1.30. *Sine 300 Hz superimposed to random vibration in Figure 1.25*

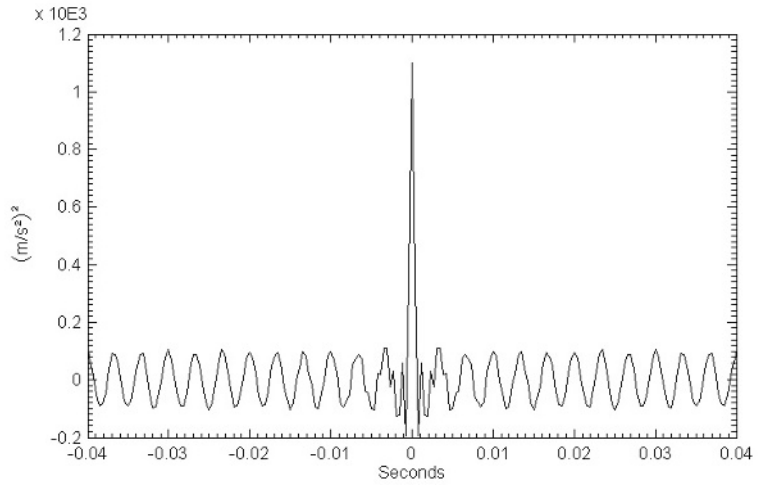


Figure 1.31. *Autocorrelation of the vibration made up of the random wideband vibration and the rms value 10 m/s^2 at 300 Hz sinusoid*

The autocorrelation function of this signal (Figure 1.31) clearly highlights the presence of this sinusoid and makes it possible to find its frequency and amplitude. The highest peak is equal to the square of the global signal rms value (random plus sine).

With this method, we can detect and characterize a low amplitude sinusoidal component. The threshold of detection, defined by the ratio of the rms values of the sine and of the PSD in the frequency interval between two consecutive points (frequency step Δf), is approximately 4.

Example 1.12.

Let us take a PSD with a 1 $(m/s^2)^2/Hz$ amplitude, calculated with a $\Delta f = 1.27$ Hz frequency step.

The calculation of the autocorrelation function can also be done from this PSD on which a 300 Hz frequency line was superimposed 0. The amplitude G of the composite PSD at this frequency is such that $G = \frac{\ddot{x}^2_{\text{rms sine}}}{\Delta f}$ ($\ddot{x}_{\text{rms sine}}$ = sinusoid rms value, i.e. 10 m/s^2) (Figure 1.32).

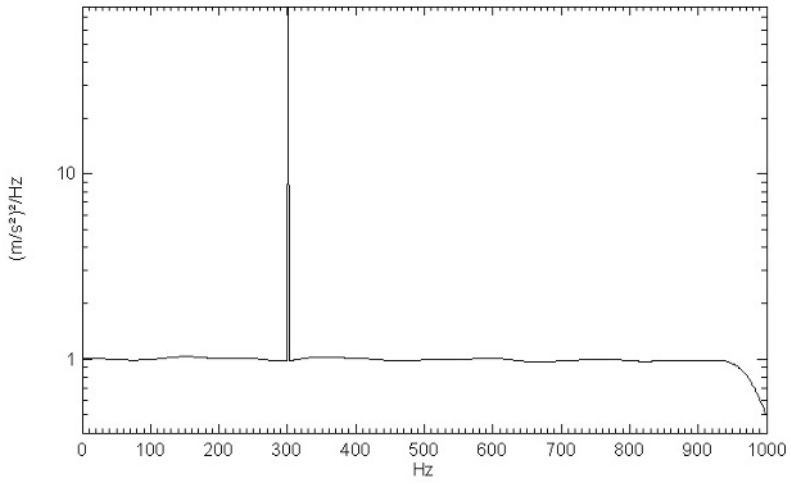


Figure 1.32. *PSD of wideband noise with 300 Hz sinusoidal line*

The sinusoid can be identified if its rms value is higher than approximately $4\sqrt{G\,\Delta f} = 4\sqrt{1.27} \approx 4.5$ m/s^2.

The resulting autocorrelation function can be superimposed over that in Figure 1.30. In order for this to occur, the line added to the PSD must be defined in a single frequency point. Otherwise, the autocorrelation function is transformed and tends toward that of a narrowband noise when this number increases.

Example 1.13.

Consider the wideband noise PSD in Figure 1.25 on which a 300 Hz, 20 m/s^2 rms value sinusoidal line is superimposed.

Figures 1.33 to 1.36 illustrate the autocorrelation function obtained when the sinusoidal line is defined consecutively by 1, 2, 4 and 9 points.

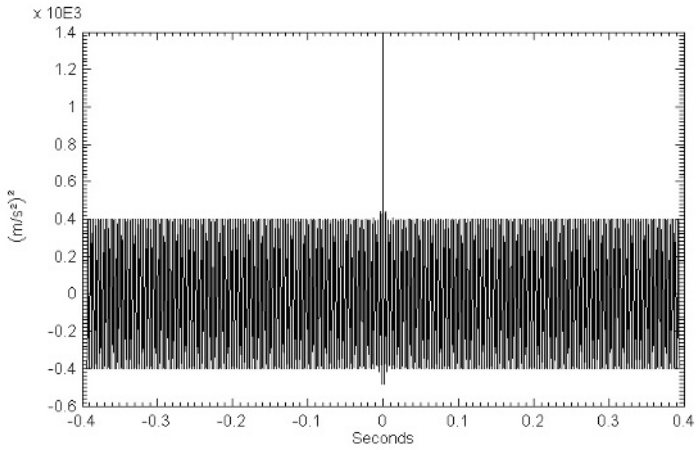


Figure 1.33. *Autocorrelation calculated from the wideband noise PSD with a sinusoidal line defined by a single point*

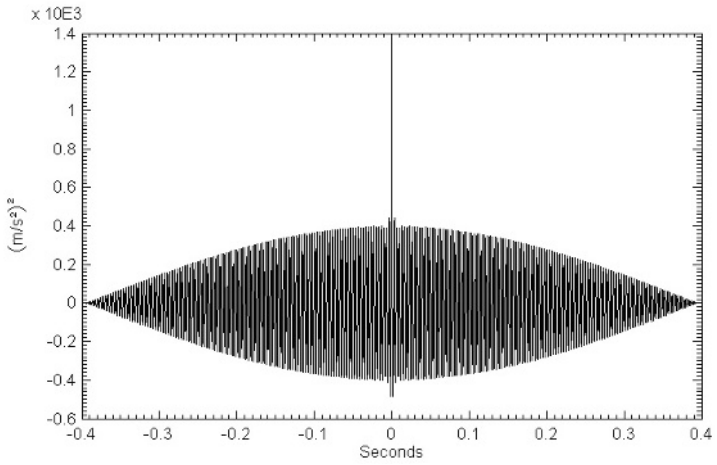


Figure 1.34. *Autocorrelation calculated from the wideband noise PSD with a sinusoidal line defined by two points*

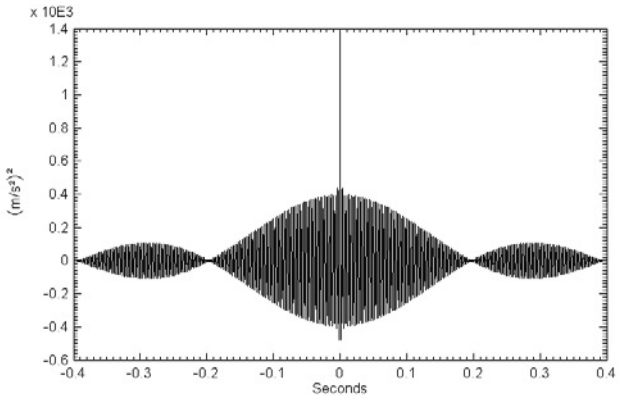


Figure 1.35. *Autocorrelation calculated from the wideband noise PSD with sinusoidal line defined by four points*

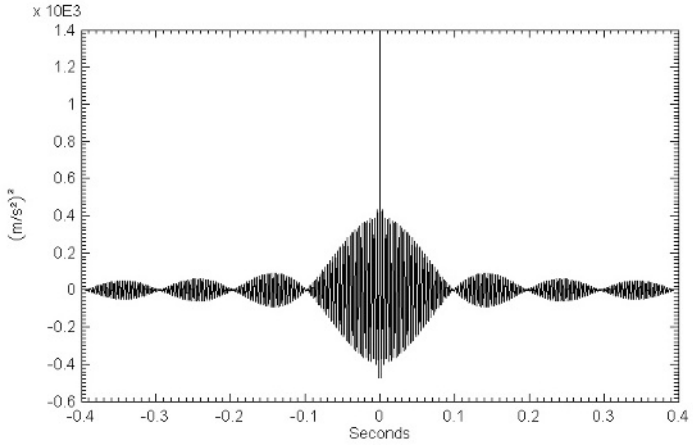


Figure 1.36. *Autocorrelation calculated from the wideband noise PSD with a sinusoidal line defined by eight points*

1.9.11. *Cross-correlation*

Let us consider two random functions $\ell(t)$ and $u(t)$; the *cross-correlation function* is defined by:

$$R_{\ell u}(\tau) = E\left[\ell(t)\ u(t+\tau)\right] = \lim_{T\to\infty} \frac{1}{2\,T} \int_{-T}^{T} \ell(t)\ u(t+\tau)\ dt \qquad [1.83]$$

The cross-correlation function makes it possible to establish the degree of resemblance between two functions of the same variable (time in general).

Discrete form [BEA 72]

If N is the number of sampled points and τ is a delay such that $\tau = m\ \Delta t$, where Δt is the temporal step between two successive points, the cross-correlation between two signals ℓ and u is given by

$$R_{\ell u}(\tau) = \frac{1}{N - m} \sum_{i=1}^{N-m} \ell_i\ u_{i+m} \qquad [1.84]$$

Covariance

The *covariance* of vibrations $\ddot{x}(t)$ and $\ddot{y}(t)$ is equal to the autocorrelation function of these centered signals. If $\overline{\ddot{x}}$ and $\overline{\ddot{y}}$ are the mean values of $\ddot{x}(t)$ and $\ddot{y}(t)$, covariance is defined by

$$C_{\ddot{x}\ddot{y}}(\tau) = \lim_{T\to\infty} \frac{1}{2\,T} \int_{-T}^{T} \left[\ddot{x}(t) - \overline{\ddot{x}}\right] \left[\ddot{y}(t+\tau) - \overline{\ddot{y}}\right] dt \qquad [1.85]$$

It is easy to show that:

$$C_{\ddot{x}\ddot{y}}(\tau) = R_{\ddot{x}\ddot{y}}(\tau) - \overline{\ddot{x}}\ \overline{\ddot{y}} \qquad [1.86]$$

If one of the mean values is zero, then the covariance is equal to the intercorrelation.

Example 1.14.

Consider a wideband random vibration and the response of a one-degree-of-freedom linear mechanical system with a natural frequency of 300 Hz and a Q factor of 10 for this noise.

Figure 1.37 shows the intercorrelation function between these two vibrations.

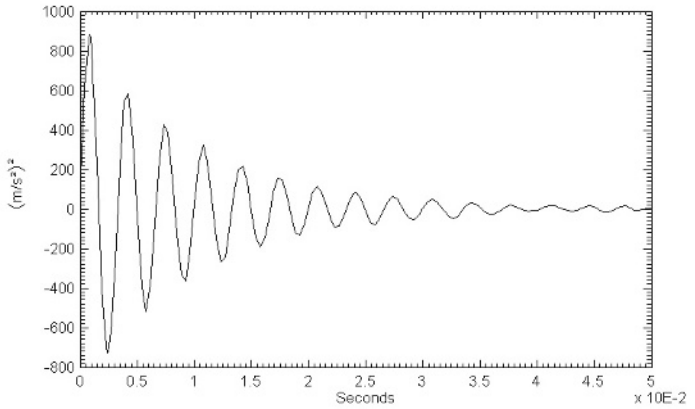


Figure 1.37. *Intercorrelation of wideband random vibration response from a linear one-degree-of-freedom system (300 Hz, Q = 10)*

Application: measure of the impulse response of a structure

If vibrations are stationary with zero mean, we can show that the intercorrelation of the vibration and response of a mechanical system is directly linked to the impulse response of the system.

Let h(t) be the impulse response of the system, $\ddot{x}(t)$ the random "excitation" vibration and $\ddot{y}(t)$ the response. This response can be calculated from the following relation:

$$\ddot{y}(t) = \int_0^t h(\alpha)\, \ddot{x}(t-\alpha)\, d\alpha \qquad [1.87]$$

The intercorrelation is equal to:

$$R_{\ddot{x}\ddot{y}}(\tau) = \int_0^{\infty} \ddot{x}(t-\tau)\, \ddot{y}(t)\, dt \qquad [1.88]$$

so, by considering [1.87]:

$$R_{\ddot{x}\ddot{y}}(\tau) = \int_0^\infty \ddot{x}(t-\tau) \int_0^\infty h(\alpha)\, \ddot{x}(t-\alpha)\, d\alpha\, dt \qquad [1.89]$$

$$R_{\ddot{x}\ddot{y}}(\tau) = \int_0^\infty \int_0^\infty h(\alpha)\, \ddot{x}(t-\tau)\, \ddot{x}(t-\alpha)\, d\alpha\, dt \qquad [1.90]$$

The "excitation" noise being wideband:

$$R_{\ddot{x}\ddot{y}}(\tau) = h(\tau)\, R_{\ddot{x}}(0) \qquad [1.91]$$

This method makes it possible to evaluate the impulse response of a system with a lower amplitude excitation than by impulsion, thus remaining in the linear domain.

Example 1.15.

The impulse response of the one-degree-of-freedom system (f_0 = 300 Hz, Q = 10) of Example 1.14 can be calculated (Figure 1.38) by dividing the intercorrelation in Figure 1.37 by $R_{\ddot{x}}(0)$, i.e. by the square of the rms value of the "excitation" wideband noise (31.6 m/s^2).

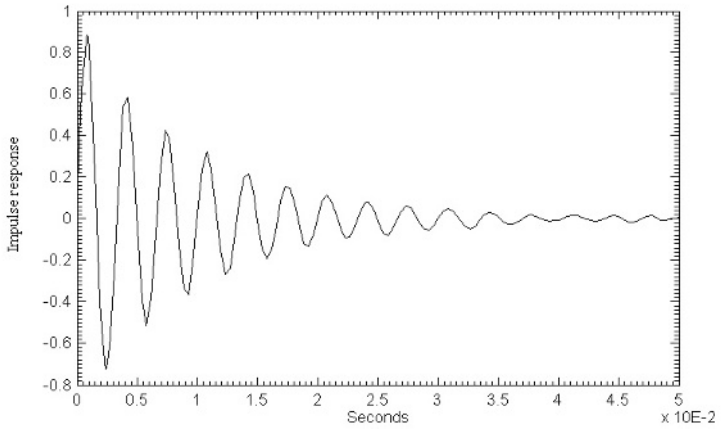


Figure 1.38. *Impulse response of the one-degree-of-freedom system (300 Hz, Q = 10) calculated from the intercorrelation function*

1.9.12. *Cross-correlation coefficient*

The *cross-correlation coefficient* $\rho_{\ell u}(\tau)$ or *normalized cross-correlation function* or *normalized covariance* is the quantity [JEN 68]

$$\rho_{\ell u}(\tau) = \frac{R_{\ell u}(\tau)}{\sqrt{R_{\ell}(0)\ R_{u}(0)}} \qquad [1.92]$$

It is shown that:

$$-1 \leq \rho_{\ell u}(\tau) \leq 1$$

If $\ell(t)$ is a random signal *input* of a system and $u(t)$ is the signal *response* at a point of this system, $\rho_{\ell u}(\tau)$ is characteristic of the degree of linear dependence of the signal u with respect to ℓ. At the limit, if $\ell(t)$ and $u(t)$ are independent, $\rho_{\ell u}(\tau) = 0$.

If the joint probability density of the random variables $\ell(t)$ and $u(t)$ is equal to $p(\ell, u)$, we can show that the cross-correlation coefficient $\rho_{\ell,u}$ can be written in the form:

$$\rho_{\ell u} = \frac{E\left[(\ell - m_{\ell})(u - m_{u})\right]}{s_{\ell}\ s_{u}} \qquad [1.93]$$

where m_{ℓ}, m_{u}, s_{ℓ} and s_{u} are respectively the mean values and the standard deviations of $\ell(t)$ and $u(t)$.

1.9.13. *Ergodicity*

A process is known as *ergodic* if all the temporal averages exist and have the same value as the corresponding ensemble averages calculated at an arbitrary given moment [BEN 58], [CRA 67], [JAM 47], [SVE 80].

Example 1.16.

In section 1.8.9, example 1.4, the process of case number 1 is ergodic and case number 2 is stationary, not ergodic. The question of ergodicity of process number 3 is irrelevant as it is not stationary.

An ergodic process is thus necessarily stationary [POU 02]. We dispose in general of only a very restricted number of records not permitting experimental evaluation of the ensemble averages. In practice, we simply calculate the temporal averages by making the assumption that the process is stationary and ergodic [ELD 61].

The concept of ergodicity is thus particularly important. Each particular realization of the random function makes it possible to consider the statistical properties of the whole ensemble of the particular realizations.

If a process is ergodic, we can limit ourselves to the frequency analysis of a short sample chosen over a single process record.

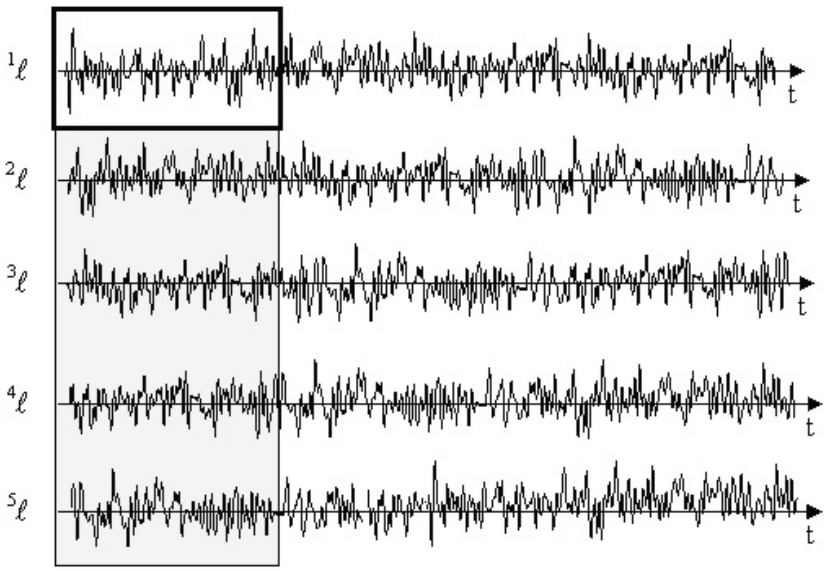


Figure 1.39. *Ergodic process: the analysis can involve a single sample*

NOTE.– *A necessary and sufficient condition such that a stationary random vibration* $\ell(t)$ *is ergodic is that its correlation function satisfies the condition [SVE 80]*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \left(1 - \frac{\tau}{T}\right) R_\ell(\tau)\, d\tau = 0 \qquad [1.94]$$

where $R_\ell(\tau)$ *is the autocorrelation function calculated from the centered variable* $\ell(t) - m$.

1.10. Significance of the statistical analysis (ensemble or temporal)

The checking of stationarity and ergodicity should in theory be carried out before any analysis of a vibratory mechanical environment, in order to ensure that the consideration of only one sample is representative of the whole process. Very often, as a result of a lack of experimental data and to save time, we make these assumptions without checking (which is regrettable) [MIX 69], [RAC 69], [SVE 80].

1.11. Stationary and pseudo-stationary signals

We saw that the signal is known as stationary if the rms value as well as the other statistical properties remain constant over long periods of time.

In the real environment, this is not the case. The rms value of the load varies in a continuous or discrete way and gives the shape of signal known as *random pseudo-stationary*. For a road vehicle, for example, variations are due to the changes in road roughness, to changes of velocity of the vehicle, to mass transfers during turns, to wind effect, etc.

The temporal random function $\ell(t)$ is known as *quasi-stationary* if it can be divided into intervals of duration T that are sufficiently long compared with the characteristic correlation time, but sufficiently short to allow treatment in each interval as if the signal were stationary. Thus, the quasi-stationary random function is a function having characteristics which vary sufficiently slowly [BOL 84].

The study of the stationarity and ergodicity is an important stage in the analysis of vibration, but it is not generally sufficient by itself; in fact, it does not make it possible to answer the most frequently encountered problems, for example the estimate of the severity of a vibration or the comparison of several stresses of this nature.

1.12. Summary chart of main definitions

	Through the process (ensemble averages)	Along the process (temporal averages)
Moment of order n	$E\left[\ell^n(t_1)\right] = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} {}^i\ell^n(t_1)$ $E\left[x^n(t_1)\right] = \int_{-\infty}^{+\infty} \ell^n(t_1)\, p\left[\ell(t_1)\right] d\ell(t_1)$	$E\left[\ell^n(t)\right] = \lim_{T\to\infty} \frac{1}{T} \int_0^T \ell^n(t)\, dt$
Centered moment of order n	$E\left\{\left[\ell(t_1) - \overline{\ell(t_1)}\right]^n\right\} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \left[{}^i\ell(t_1) - \overline{\ell(t_1)}\right]^n$ $E\left[\left[\ell(t_1) - \overline{\ell(t_1)}\right]^n\right] = \int_{-\infty}^{+\infty} \left[\ell(t_1) - \bar{\ell}\right]^n p\left[\ell(t_1)\right] d\ell(t_1)$	$E\left[\left[\ell(t) - \bar{\ell}\right]^n\right] = \lim_{T\to\infty} \frac{1}{T} \int_0^T \left[\ell(t) - \bar{\ell}\right]^n dt$
Variance	$s^2 = E\left\{\left[\ell(t_1) - \overline{\ell(t_1)}\right]^2\right\}$	$s^2 = E\left\{\left[\ell(t) - \bar{\ell}\right]^2\right\} = \overline{\ell^2} - \bar{\ell}^2$
Autocorrelation	$R_\ell(t_1, t_1+\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} {}^i\ell(t_1)\ {}^i\ell(t_1+\tau)$ $R_\ell(\tau) = \int_{-\infty}^{+\infty} \ell(t_1)\, \ell(t_1+\tau)\, p\left[\ell(t_1)\right] d\ell(t_1)$	$R_\ell(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \ell(t)\ \ell(t+\tau)\, dt$
Cross-correlation	$R_{\ell u}(t_1, t_1+\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} {}^i\ell(t_1)\ {}^iu(t_1+\tau)$ $R_{\ell u}(\tau) = \int_{-\infty}^{+\infty} \ell(t_1)\, u(t_1+\tau)\, p\left[\ell(t_1)\right] d\ell(t_1)$	$R_{\ell u}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \ell(t)\, u(t+\tau)\, dt$
	Stationarity if all the averages of order n are independent of the selected time t_1.	Ergodicity if the temporal averages are equal to the ensemble averages.

Table 1.1. *Main definitions*

Autocorrelation	$R_\ell(\tau)=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\ell(t)\,\ell(t+\tau)\,dt$		$R_\ell(\tau)=\frac{1}{N-m}\sum_{i=1}^{N-m}\ell_i\cdot\ell_{i+m}$
Normalized autocorrelation function or correlation coefficient		$\rho_\ell(\tau)=\frac{R_\ell(\tau)}{R_\ell(0)}$	
Autocovariance	$C_\ell(\tau)=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\left[\ell(t)-\bar{\ell}\right]\left[\ell(t+\tau)-\bar{\ell}\right]dt$		$C_\ell(\tau)=\frac{1}{N-m}\sum_{i=1}^{N-m}(\ell_i-\bar{\ell})(\ell_{i+m}-\bar{\ell})$
Normalized autocovariance		$\rho_\ell(\tau)=\frac{C_\ell(\tau)}{C_\ell(0)}$	
Intercorrelation function	$R_{\ell u}(\tau)=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\ell(t)\,u(t+\tau)\,dt$		$R_{\ell u}(\tau)=\frac{1}{N-m}\sum_{i=1}^{N-m}\ell_i\,u_{i+m}$
Intercorrelation coefficient or normalized intercorrelation function		$\rho_{\ell u}(\tau)=\frac{R_{\ell u}(\tau)}{\sqrt{R_\ell(0)\,R_u(0)}}$	
Covariance	$C_{\ell u}(\tau)=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\left[\ell(t)-\bar{\ell}\right]\left[u(t+\tau)-\bar{u}\right]dt$		$C_{\ell u}(\tau)=\frac{1}{N-m}\sum_{i=1}^{N-m}(\ell_i-\bar{\ell})(u_{i+m}-\bar{u})$
Normalized covariance		$\rho_{\ell u}(\tau)=\frac{C_{\ell u}(\tau)}{\sqrt{C_\ell(0)C_u(0)}}$	

Table 1.2. *Some temporal averages with their discrete form*

1.13. Sliding mean

Instead of calculating the amplitude average of a signal over all available points, we can focus on a small number of consecutive points. This "block" is shifted by one point at each calculation (Figure 1.40). Each average thus obtained is attributed to the moment corresponding to the middle point in the block. All averages determined for the signal then make it possible to draw a curve based on time called the *sliding mean*.

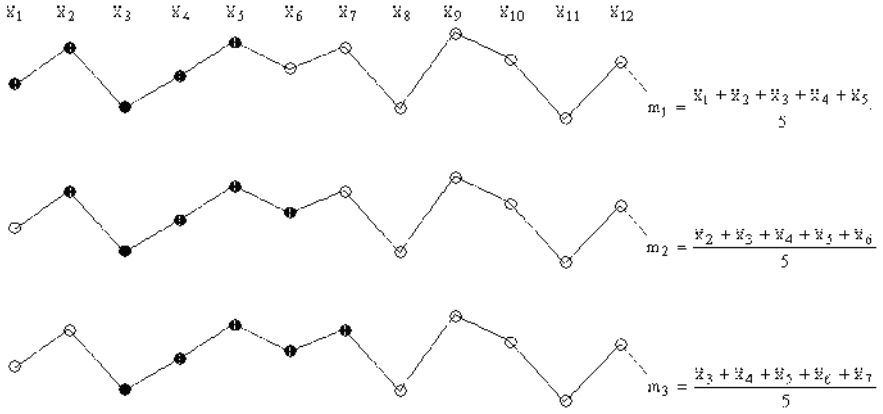


Figure 1.40. *Calculation process of the sliding mean*

There are several strategies to calculate the averages closer to the extremities. For example, we can determine the average of the first n/2 points and attribute it to the first point in the sliding mean curve, the average of the first n/2 + 1 points and the second point, etc., until we have all n points. The same methodology can be applied symmetrically for the extremity of the signal.

Similarly, this type of calculation can also be applied to the quadratic mean (for rms value, standard deviation, etc.) or a higher average (skewness, kurtosis, etc.).

The calculation speed of all these sliding means can be greatly accelerated with the use of the previous point calculation results for each point.

Example 1.17.

Figure 1.41 shows the rms value, skewness and kurtosis calculated according to time from a Gaussian signal lasting 5 seconds (sliding means over 1,000 points). We can verify that skewness and kurtosis are close to 0 and 3 respectively.

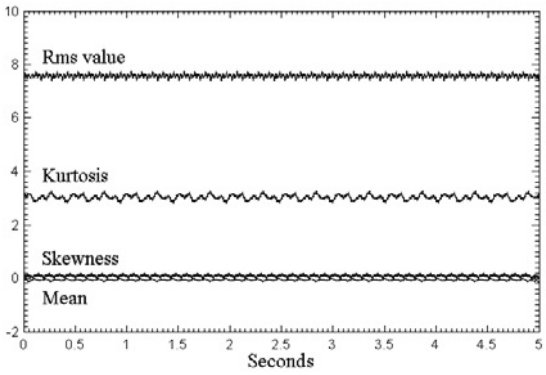


Figure 1.41. *Sliding means on 1,000 points from a Gaussian signal of 32,372 points*

Example 1.18.

Consider the random vibration taken from the platform of a truck (Figure 1.42).

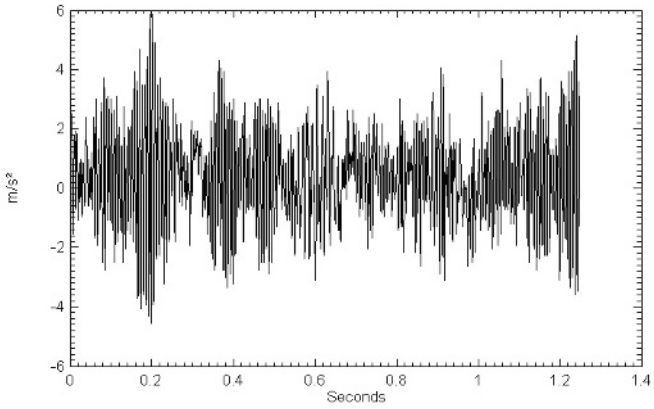


Figure 1.42. *"Truck" vibration*

Figures 1.43, 1.44 and 1.45 show these same sliding means calculated on 100, 500 and 1,000 points. The rms value, skewness and kurtosis vary with time (non-stationary vibration). The curves are smoother the higher the number of calculation points for each average.

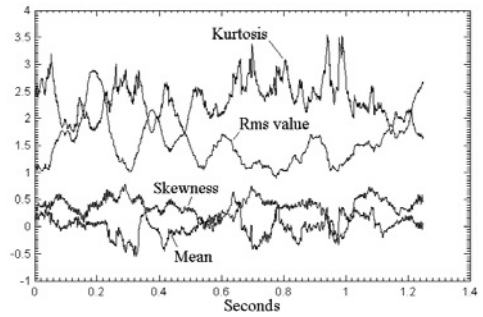


Figure 1.43. *"Truck" vibration – sliding means on 100 points*

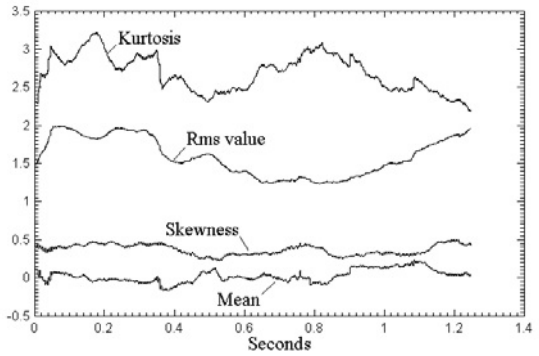


Figure 1.44. *"Truck" vibration – sliding means on 500 points*

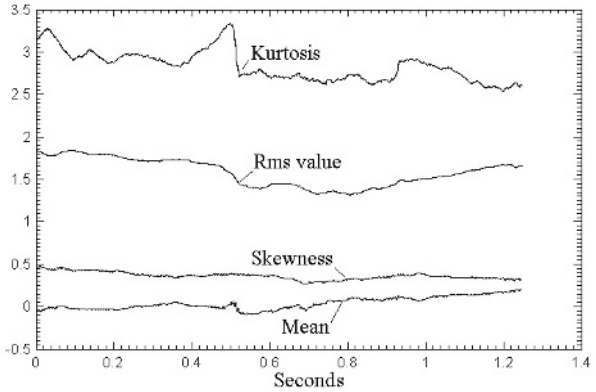


Figure 1.45. *"Truck" vibration – sliding means on 1,000 points*

1.14. Test of stationarity

The choice of analysis method of a random vibration in the frequency range is dependent on the stationary or non-stationary character of this vibration. We will see in particular that the calculation of the power spectral density only has meaning if the signal is stationary. Most often, the test simply consists of verifying that, in the time interval considered, the rms value calculated using a sliding mean varies little as a function of time.

Other more specific methods of a statistical nature have been developed, calling upon hypothesis tests.

1.14.1. *The reverse arrangements test (RAT)*

From a signal made up of N points, this method consists of [BEN 10], [BEN 65], [BRA 11], [HIM 06]:

– Cutting the signal into m samples of size equal to n points. H. Himelblau *et al.* propose to choose the integer m as a function of the signal duration (Table 1.3):

Duration	m
< 20 s	20
≥ 20 s and ≤ 100 s	Integer greater than the duration, so as to obtain samples with a duration close to 1 s
≥ 100 s	100

Table 1.3. *Choice of number of samples as a function of the signal duration*

m should be at least greater than 10, and preferably greater than 20.

– Calculating for each sample the parameter that will be used for the test: rms value in general, but another parameter could be used, such as the skewness or the kurtosis. This parameter will be denoted r_i.

– Calculating a function h such that

$$h_{ij} = \begin{cases} 1 & \text{if } r_i > r_j \\ 0 & \text{otherwise} \end{cases} \qquad [1.95]$$

where i = 1, 2, …, n - 1

$$j = i + 1, i + 2, \ldots, n$$

– Calculating $A_i = \sum_{j=i+1}^{n} h_{ij}$ [1.96]

and

$$A = \sum_{i=1}^{n-1} A_i \quad [1.97]$$

(total number of reverse arrangements).

– Choosing a significance level α as a function of the signal duration studied. H. Himelblau suggests the following values (Table 1.4):

m	α
< 20	0.10
20 to 40	0.05
> 40	0.02

Table 1.4. *Significance level α versus m*

– If parameter r is stationary, number A is a random variable having as a mean and standard deviation [HIM 06], [KEN 61]:

$$\mu_A = \frac{m\,(m-1)}{4} \quad [1.98]$$

$$\sigma_A = \sqrt{\frac{m\,(2m+5)\,(m-1)}{72}} \quad [1.99]$$

– For a chosen significance level α, the area of acceptation of the stationarity hypothesis is given by

$$A_{m;1-\alpha/2} < A < A_{m;\alpha/2} \quad [1.100]$$

where

$$A_{m;1-\alpha/2} = \mu_A - \sigma_A \, E_2^{-1}\left(\frac{1-\alpha}{2}\right) \quad [1.101]$$

$$A_{m;\alpha/2} = \mu_A + \sigma_A \, E_2^{-1}\left(\frac{1-\alpha}{2}\right) \quad [1.102]$$

$E_2^{-1}()$ being the inverse error function of $E_2(x)$ defined in Appendix A4.1.2 :

$$E_2(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} \, dt$$

NOTE.– *The inverse error function of* $E_1(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, dt$ *can also be used from*

$$A_{m;1-\alpha/2} = \mu_A - \sigma_A \sqrt{2} \, E_1^{-1}(1-\alpha) \quad [1.103]$$

$$A_{m;\alpha/2} = \mu_A + \sigma_A \sqrt{2} \, E_1^{-1}(1-\alpha) \quad [1.104]$$

There are, for example, as a reference [NEW 93], tables giving the probability for there being fewer than c arrangements of N of terms of a series whose order is random.

m	$A_{m;1-\alpha/2}$			$A_{m;\alpha/2}$		
	$1-\alpha/2 = 0.99$	$1-\alpha/2 = 0.975$	$1-\alpha/2 = 0.95$	$\alpha/2 = 0.05$	$\alpha/2 = 0.025$	$\alpha/2 = 0.01$
10	9	11	13	31	33	35
20	59	64	69	120	125	130
30	152	162	171	263	272	282
40	290	305	319	460	474	489
50	473	495	514	710	729	751
60	702	731	756	1013	1038	1067

Table 1.5. *Values of limits* $A_{m;1-\alpha/2}$ *and* $A_{m;\alpha/2}$ *as a function of the number of samples m and of the significance level* α

Example 1.19.

Let us consider the following rms values, calculated from a signal cut into $m = 20$ samples.

59.4	61.4	37.6	41.8	48.5	54.6	52.5	54.1	43.2	58.0
45.8	47.3	56.0	52.5	49.9	47.3	61.4	57.4	58.2	46.4

To a significance level $\alpha = 0.05$, Table 1.5 gives, for m = 20, the limits of the acceptance region $A_{20;\,1-0.975} = 64$ and $A_{20;\,1-0.025} = 125$. From relations [1.95] to [1.97], we obtain A = 85, value in the interval [64, 125]: the signal can be considered as stationary.

1.14.2. *The runs test*

The *runs test* (or *Wald–Wolfowitz test*) is a non-parametric statistical test that checks a randomness hypothesis for a two-valued data sequence. It is used to test the hypothesis that a series of numbers is random [BRA 68], [HOG 97], [SHE 04], [STE 39], [WAL 40].

To simplify computations, the data are first centered about or their rms (or mean or median) value.

Let us consider N observations. Each observation is classed exclusively in one of the three following categories, identified by the symbols + and – [BEN 63], [SIE 56]:

– (+) if the value is greater than the rms value;

– (–) if it is smaller than the rms value.

A *run* is defined as a sequence with the same sign ('+' or '–'). A positive run is a sequence of values greater than the rms value, and a negative run is a sequence of values less than the rms value [BEN 65], [BRA 11], [BRA 68], [YAM 87]. As an example, let us consider n = 25 observations of which n_1 are greater than the rms value and n_2 are lower (Figure 1.46).

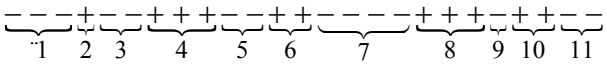


Figure 1.46. *Example of sequence of observations*

The first step is to count the number of runs in the data sequence. In this example, the n observations that lead to a total of runs $R = 11$ runs.

We can then test if the number of positive and negative runs are distributed equally in time.

W.L. Stevens [STE 39], then F.S. Swed and Eisenhart [SWE 43] have shown that the probability of the number of runs r being less than or equal to R is

$$P(r \le R) = \frac{1}{C_n^{n_1}} \sum_{r=2}^{R} f_r \qquad [1.105]$$

where, by choosing $n_1 \le n_2$,

if $r = 2k$

$$f_r = 2\, C_{n_1-1}^{k-1}\, C_{n_2-2}^{k-1} \qquad [1.106]$$

if $r = 2k-1$

$$f_r = C_{n_1-1}^{k-1}\, C_{n_2-1}^{k-2} + C_{n_1-1}^{k-2}\, C_{n_2-1}^{k-1} \qquad [1.107]$$

with $k = 1, 2, 3, \cdots, m+1$.

From these relations, it is possible to search digitally by iterations the approximate limit values of the confidence interval corresponding to a given significance level α for a given value of n.

Table 1.6 gives several values thus obtained for $\alpha = 0.01$, $\alpha = 0.05$ and $\alpha = 0.10$ for $n_1 = n_2$. The values differ very little when this is not the case. As r varies by integer values, it is not possible to find a value that corresponds exactly to the desired probability. We have retained the value of r that leads to the probability closest to the one desired.

N	$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.10$	
	Lower limit	Upper limit	Lower limit	Upper limit	Lower limit	Upper limit
12	2	11	3	10	4	9
14	3	12	4	11	4	11
16	3	14	5	12	5	12
18	4	15	5	14	6	13
20	5	16	6	15	7	14
25	7	19	8	18	9	17
30	9	22	10	21	11	20
35	10	25	12	24	13	23
40	12	29	14	27	15	26
45	14	32	16	30	17	28
50	16	35	19	32	20	31
60	21	40	23	38	24	37
70	25	46	27	44	29	42
80	29	52	32	49	33	48
90	33	58	36	55	38	53
100	38	63	41	60	42	59
120	46	75	50	71	52	69
140	55	86	54	77	61	80
160	64	97	68	93	70	91
180	73	108	77	104	79	102
200	82	119	87	114	89	112

Table 1.6. *Upper and lower limits as a function of observations N and significance level α*

There are tables giving the bounds of this interval as a function of n_1 and n_2 in [MEN 86] and [SWE 43].

The signal is considered to be stationary if R is between the two bounds calculated for n and α given.

These parameters do not assume that the positive and negative values have the same probability of occurrence, but only assume that these values are independent and identically distributed.

If the number of runs is significantly more or less than that expected, the hypothesis of a statistical independence of observations can be rejected.

Example 1.20.

Let us consider the following 20 observations:

+	+	-	-	-	+	+	+	-	+
59.4	61.4	37.6	41.8	48.5	54.6	52.5	54.1	43.2	58.0
-	-	+	+	-	-	+	+	+	-
45.8	47.3	56.0	52.5	49.9	47.3	61.4	57.4	58.2	46.4

A (+) or (-) sign was attributed to each observation according to whether it is greater or lower than the rms value of the set of observations (52.09).

We count $n_1 = 9$ values with a (–) sign and $n_2 = 11$ with a (+) sign. The number of runs R is equal to 10.

For a significance level of 0.05, the limits calculated in these conditions are equal to 6 and 15. R being equal to 10, the hypothesis of a stationary signal can be accepted with this significance level.

For a large-sample runs test (where $n_1 > 10$ and $n_2 > 10$), the number of runs R is an approximately normally distributed random variable. It is thus possible to calculate the normal random variable Z as follows [SHE 04]:

$$Z = \frac{R - \overline{R}}{\sigma_R} \qquad [1.108]$$

The mean value $\overline{R}$ and variance σ_R^2 of this distribution are [BRO 84]:

$$\overline{R} = \frac{2\, n_1\, n_2}{n} + 1 \qquad [1.109]$$

$$\sigma_R^2 = \frac{2\, n_1\, n_2 \left(2\, n_1\, n_2 - n\right)}{n^2 \left(n - 1\right)} \qquad [1.110]$$

The runs test rejects the hypothesis if $|Z| > Z_{1-\alpha/2}$ where α is the chosen significance level. It is thus necessary to verify that

$$\overline{R} - Z_{1-\alpha/2}\, \sigma_R < R < \overline{R} + Z_{\alpha/2}\, \sigma_R \qquad [1.111]$$

If there is a slowly varying periodicity in the data, with the rms value periodically increasing and dercreasing, the sequence of rms values will, of course, not be random.

The reverse arrangements test is based on rearranging the observations and it is thus not very sensitive to periodic components present in the signal. To highlight such components, which make the signal non-stationary, it is preferable to use a runs test [BRO 84], [SHE 04].

1.15 Identification of shocks and/or signal problems

The analysis of kurtosis according to time, drawn with a sliding mean, enables us to easily detect the presence of a local irregularity of the signal, whether it is due to a shock that is present in the real environment or to a signal problem that does not correspond to any physical mechanical reality (for example, a brief telemetry loss).

Kurtosis is actually very sensitive to "abnormal" signal values moving away from the distribution controlling the points.

Example 1.21.

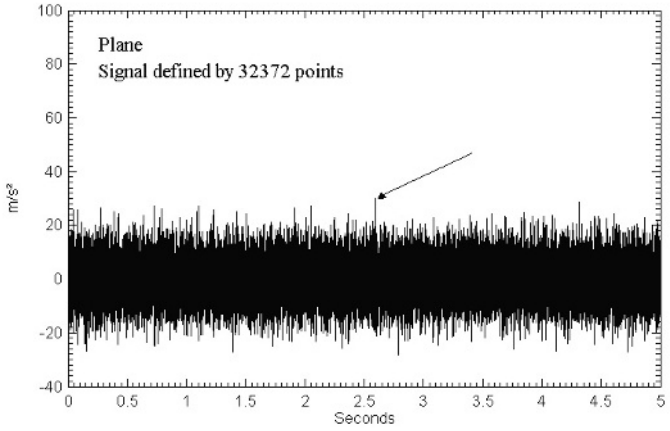


Figure 1.47. *Sample of vibratory signal measure on a plane*

In order to show the sensitivity of this parameter, consider the stationary vibratory signal measured on a plane, lasting 5 s and defined by 32,372 points (Figure 1.47). This Gaussian signal presents a few high peaks whose amplitude can be about 5 times the signal rms value in accordance with what is expected with a Gaussian distribution.

Kurtosis according to time, calculated with a sliding mean over 500 points (Figure 1.48), remains approximately constant during the measure period. We can verify that its value is close to 3 (Gaussian signal).

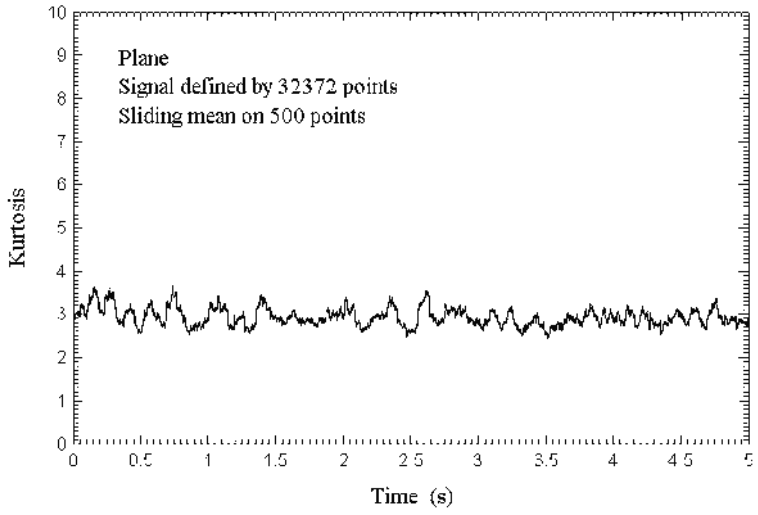


Figure 1.48. *Kurtosis versus time of vibration of Figure 1.47*

Figure 1.49 shows the signal of Figure 1.47 after multiplication of a single peak by a factor equal to 3.

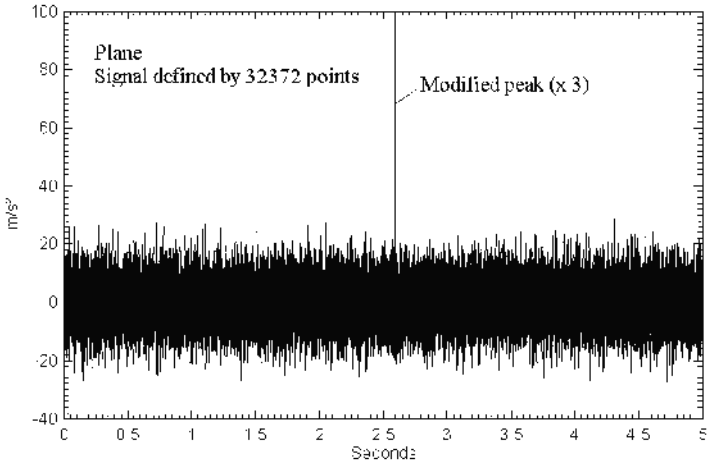


Figure 1.49. *Vibration of Figure 1.47 modified (multiplication by 3 of one peak)*

The effect on kurtosis of this modification is very important. We can observe at the peak moment a kurtosis value exceeding 40, which is thus much higher than 3 and very easily detectable (Figure 1.50).

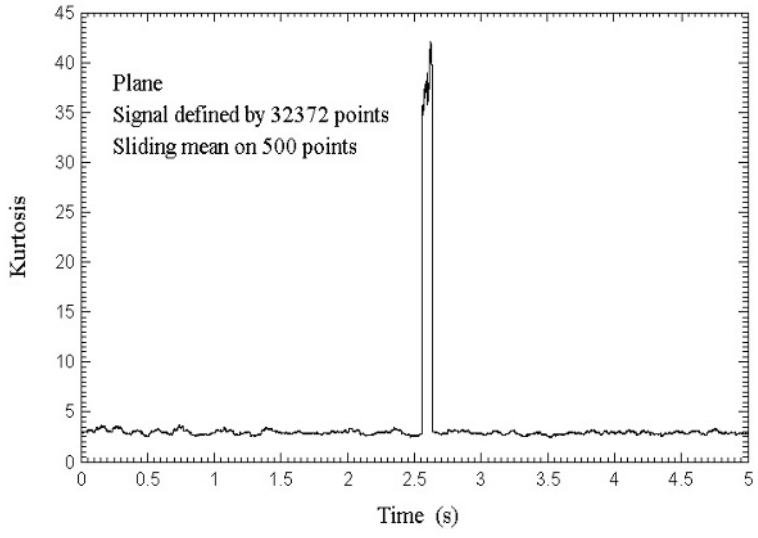


Figure 1.50. *Kurtosis of the signal modified in Figure 1.49*

If this was a mechanical shock present within the random signal, there would be more modified points and kurtosis could reach, depending on the amplitude of the shock in relation to the noise, much greater values, by a few hundred for example.

1.16. Breakdown of vibratory signal into "events": choice of signal samples

If we only have a single record, it is sufficient to ensure that it is slightly autostationary to be able to characterize it with the help of PSD. If the distribution of instantaneous instant values of samples of a slightly autostationary record follows a Gaussian distribution, the record is *strongly autostationary*.

The mean value is generally zero (it can be brought down to zero if that is not the case). If the mean square value is constant in relation to time (the rms value), we can reasonably expect that the autocorrelation will be constant.

It is sufficient to verify that the rms value does not change in relation to time.

Practical problems

Each record is made up of several different sequences (also called *events*). Take the example of transportation by place: a record will include the runway haul, take-off, climb, cruising altitude, changes of direction, air brake phases, etc. to descent and landing. The process made up with these types of records is not stationary.

If we had a large number of records, we would have as many measures of each phase and we could imagine studying their stationarity and ergodicity. However, only rarely can we have more than a few measures. To reduce the cost of the measures campaign, we only carry out some records of the phenomenon in the best of cases. The phase studies may also be too short for these statistical tests in significant conditions.

If we carry out several instrumented flights for better environment understanding, we will probably not obtain the same rms value for each sequence (if only because weather conditions can change). In order for the process to be ergodic, all the conditions would have to be strictly the same, which would necessitate the characterization of all the processes corresponding to all possible conditions.

Clearly, it never works that way. Still, a reasonable methodology consists of carrying out several measures to be able to use a small statistic for estimating the most severe conditions.

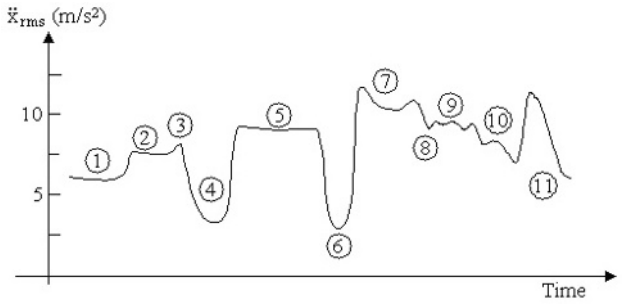
In practice, for each record, we draw the rms value in relation to time, as well as the skewness and kurtosis (sliding means on the whole signal).

The first of these curves locates temporal ranges in which the rms value does not change much and which we will consider as stationary.

Each signal is separated into sections representing a specific phase of the flight, generating a vibratory or particular shock environment, with common characteristics from one flight to another: shock linked to landing, cruising flight without turbulence, etc.

Example 1.22.

Let us take a vibratory environment on an aircraft, represented by acceleration as a function of time:



1. Taxi
2. Take-off
3. Climb
4. Cruise at high altitude
5. Maximum velocity at high altitude
6. Cruise at low altitude
7. Maximum velocity at low altitude
8. Climbing turn
9. Deceleration
10. Landing approach
11. Touchdown

Figure 1.51. *Rms acceleration recorded on an aircraft during flight [KAT 65]*

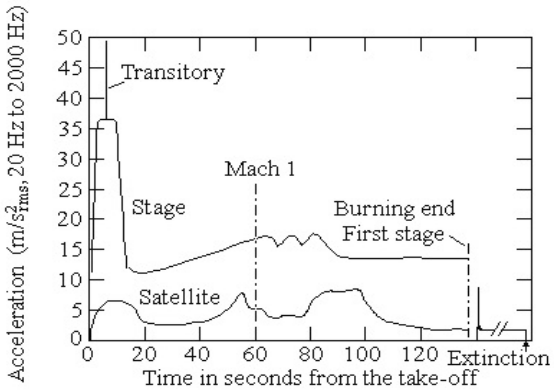


Figure 1.52. *Rms value of vibrations measured on a satellite during launch*

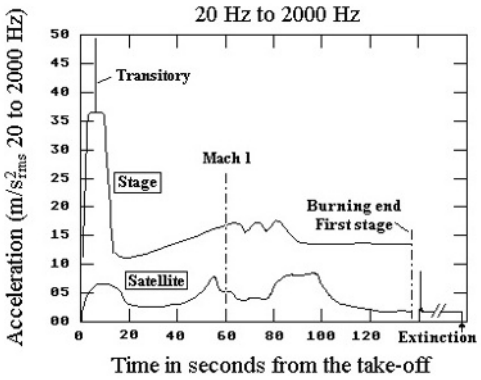


Figure 1.53. *Rms value of vibrations taken from a satellite during launch*

Skewness and kurtosis give an indication about the Gaussian character of the signal in these time intervals (close to 3 and 0 respectively in this hypothesis).

A quick and very important variation of the kurtosis generally indicated the presence of a shock. However, it is important that we consider the signal according to time to make sure it is not a problem with the signal (loss of telemetry, etc.), affecting kurtosis in a very similar way.

Example 1.23.

The chosen vibratory signal, created for this example, represents a measure that could have been carried out on a satellite launcher. It is represented in Figure 1.54.

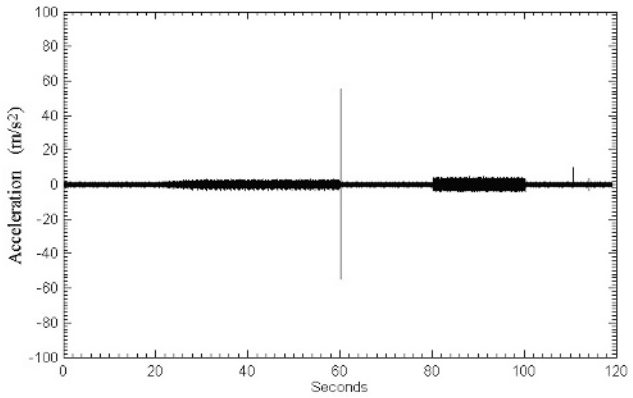


Figure 1.54. *Acceleration versus time*

The rms value, skewness and kurtosis according to time, calculated with the help of a sliding mean on 500 points, are given respectively in Figures 1.55, 1.56 and 1.57.

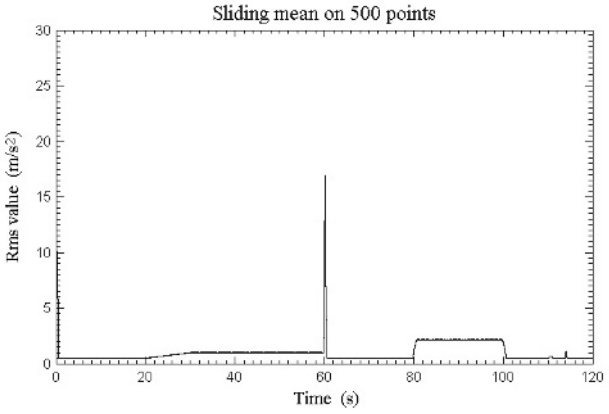


Figure 1.55. *Rms value versus time calculated from the signal in Figure 1.54*

Figure 1.55 highlights the presence of several consecutive phases:

– a shock (to analyze in a more finite way to distinguish it from a signal fault);

– a stationary phase on the first 20 s (constant rms value);

– a transitory phase lasting 10 s;

– a new stationary phase between 30 s and 60 s;

– a shock (at approximately 60 s);

– a stationary phase approximately between 60 s and 80 s;

– a very brief transitory phase;

– a stationary phase between 80 and 100 s;

– a very brief transitory phase;

– a stationary phase between 100 and 120 s approximately, during which the rms value shows 2 peaks.

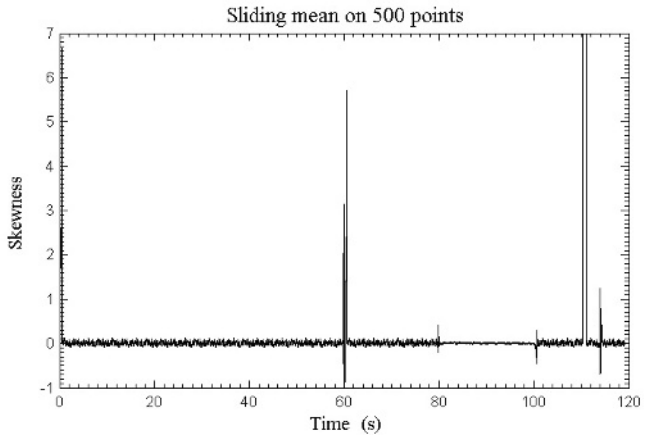


Figure 1.56. *Skewness versus time calculated from the signal in Figure 1.54*

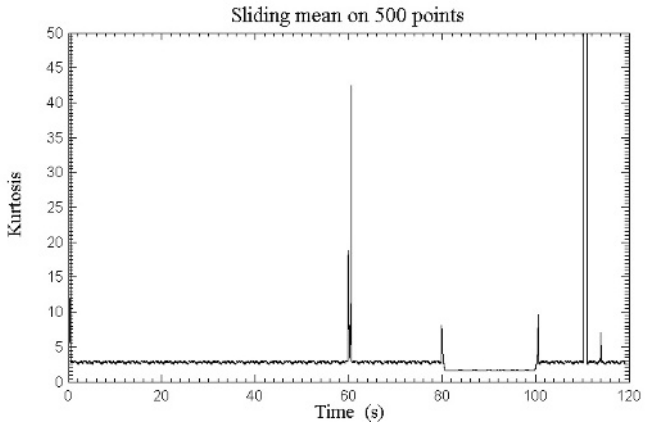


Figure 1.57. *Kurtosis versus time calculated from the signal in Figure 1.54*

Skewness (Figure 1.56) and kurtosis (Figure 1.57) show that the signal relative to stationary phases is Gaussian (skewness equal to 0 and kurtosis equal to 3), except between 80 s and 100 s (constant kurtosis lower than 3).

The shock at 60 s and the 2 peaks observed in the last stationary phase, particularly the first peak, appear clearly on the skewness and kurtosis.

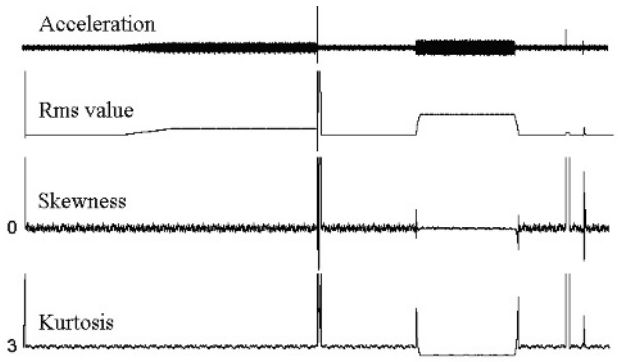


Figure 1.58. *Juxtaposition of acceleration, rms value, skewness and kurtosis according to time*

The juxtaposition of these curves (Figure 1.58) and the previous comments help us reach some conclusions. The vibratory signal is representative of a succession of events which can be described consecutively by (Figure 1.59):

– a mechanical shock (after analysis of the acceleration signal based on time), characterized by its shock response spectrum;

– a stationary random vibration, characterized by a PSD (Chapter 4), which is used to calculate an extreme response spectrum[2] (ERS) and a fatigue damage spectrum (FDS) (Volume 5);

– a non-stationary random vibration (we can use it to calculate an ERS and an FDS directly from the signal);

– a stationary random vibration;

– a shock;

– a stationary random vibration;

– between 80 s and 100 s, a periodic vibration to analyze in a more finite way to determine its components (kurtosis different from 3, and after analysis of the acceleration signal in this zone);

– a stationary phase in which we observe a signal fault (a very high peak that is extremely small and that cannot be linked to a specific phenomenon of the life profile);

– a low amplitude shock.

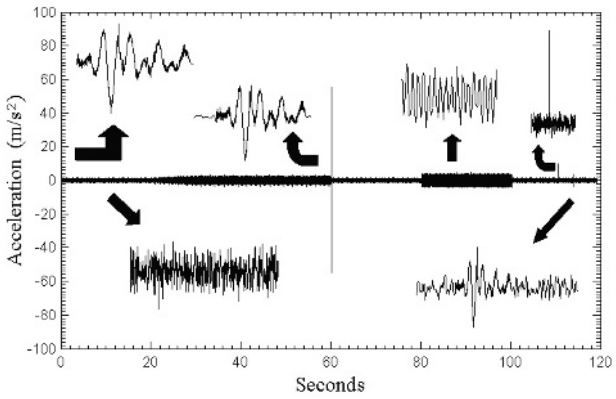


Figure 1.59. *Identification and extraction of the different events making up the vibratory signal*

If we have several logs, we should analyze each one of them.

2 Also called "maximum response spectrum" (MRS).

It is very useful to plot the variations of the rms value against time (sliding mean on n points) in order to:

– choose the time intervals over which the rms value varies little: each corresponding phase can then be characterized by a PSD;

– study the very short duration phenomena (non-stationary phenomena). The analysis for example measures the number of times that the rms value crosses a given threshold with respect to the amplitude of this threshold (rms value of the total signal or of the response of a one-degree-of-freedom mechanical system of constant Q factor, generally equal to 10, whose natural frequency varies in the useful frequency band) [KEL 61].

The variation of the rms value with time has also been used as a monitoring tool for the correct operation of rotating machinery based on a statistical study of their vibratory behavior [ALL 82].

1.17. Interpretation and taking into account of environment variation

Consider a stationary phase in a vibratory environment, for example, the cruising phase of air transport. The rms value is often different from one flight to another even though it is constant during this phase for each flight (slightly different weather conditions, etc.). This actually represents the non-ergodicity of the process.

This variation can be considered:

– by calculating the statistical PSD or, preferably, ERS and FDS (for example, average + 3 standard deviations) to describe the event with a crossing risk lower than 0.135%;

– by applying an uncertainty coefficient during the development of specifications (use of variation coefficient of environment, standard deviation and average ratio).

Example 1.24.

Let us assume that vibrations relative to air transport were measured during 6 different flights. It seems difficult to study the ensemble averages of this process with only 6 logs. Instead, each signal is analyzed to detect (from the rms value line according to time obtained with with a sliding mean) time intervals during which the signal is autostationary (slight variation of rms value).

Each of these intervals is identified for a specific event (cruising altitude phase for example) in the situation involved ("air transport") and can be characterized by calculating a PSD on part of the corresponding signal or on its whole, depending on duration. The same procedure is applied to each of the other 6 logs. At the end of these analyses, we then have 6 PSDs representative of the cruising phase (Figure 1.60).

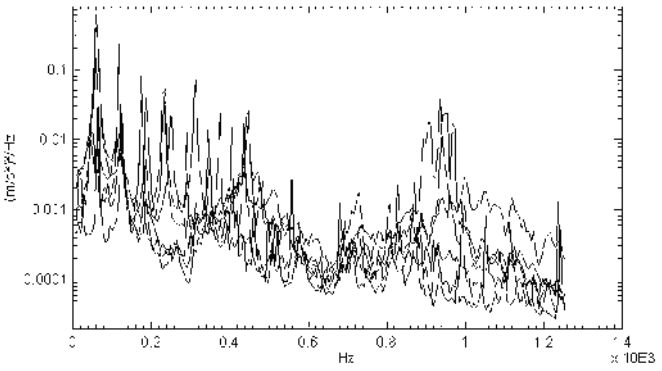


Figure 1.60. *Power spectral densities (6) characterizing the "cruising phase" random process of air transport*

The rms values of these 6 PSDs are equal to 1.5 m/s^2, 1.36 m/s^2, 1.71 m/s^2, 2.82 m/s^2, 3.16 m/s^2 and 1.65 m/s^2 respectively.

If the process was completely ergodic, all these PSDs would be extremely close with the same rms value (the differences are due to the short time of the signal samples used and the resulting statistical error).

We can observe that the PSDs are relatively dispersed, mainly in amplitude, since the dispersion is caused by the conditions of each flight (turbulences which can be important).

The random process is not really ergodic. In order to treat this general problem, we consider, from one flight to another, that the levels are random and we evaluate:

– an average spectrum (PSD) and a standard deviation spectrum (at each frequency, average and standard deviation of PSD values) if the number of available spectrums permits it (at least 4 or 5 measures);

– or, on the contrary, simply a spectrum "envelope" (greatest PSD value at each frequency).

These two spectrums can be used:

– to determine a PSD with a low probability of being exceeded (for example, average + 3 standard deviations), which can possibly be used by an engineering firm to size a mechanical part,

– or, by dividing the standard deviation by the mean at each frequency, to calculate the coefficient of variation that will characterize the dispersion of the environment involved.

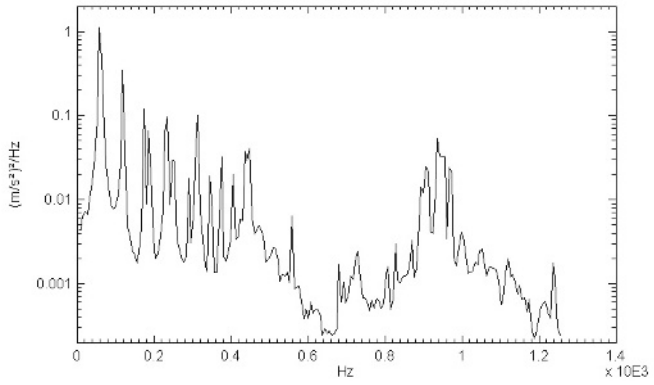


Figure 1.61. *Average PSD plus 3 standard deviations calculated from the 6 PSDs in Figure 1.60*

We will see in Volume 5 that the coefficient of variation is preferably calculated from the ERS and FDS determined with PSDs. It will be used in the calculation of an uncertainty coefficient during the development of a specification.