



HAL
open science

About periodicity and signal to noise ratio - The strength of the autocorrelation function

Nadine Martin, Corinne Mailhes

► **To cite this version:**

Nadine Martin, Corinne Mailhes. About periodicity and signal to noise ratio - The strength of the autocorrelation function. CM 2010 - MFPT 2010 - 7th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies, Jun 2010, Stratford-upon-Avon, United Kingdom. pp.n.c. hal-00449085

HAL Id: hal-00449085

<https://hal.science/hal-00449085>

Submitted on 7 Nov 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

About periodicity and signal to noise ratio - The strength of the autocorrelation function

Nadine Martin
GIPSA-lab, INPG/CNRS
Department Images Signal
Grenoble Campus
BP 46 - 961, rue de la Houille Blanche
F-38402 Saint Martin d'Hères – France
Phone 33 (0)4 76 82 62 69
Fax 33 (0)4 76 57 47 90
nadine.martin@gipsa-lab.grenoble-inp.fr

Corinne Mailhes
IRIT-TESA, INP – ENSEEIHT
2, rue Charles Camichel
B.P. 7122
31071 Toulouse Cedex 7 – France

Abstract

In condition monitoring a part of the information necessary for decision-making comes from scrutinizing a time measure or a transform of this measure. Frequency domain is commonly exploited; lag domain is not, albeit advantages of the autocorrelation function have long been known. In this paper, we dwell on the autocorrelation function in order to extract some interesting properties of the measure. We propose two indicators in order to characterize the periodicity of a signal. First is based on the non-biased autocorrelation function and indicates a fundamental periodicity rate. Second is based on the biased autocorrelation and gives a dominant-power periodicity rate. The study of the 2D-plane defined by these two indicators allows the definition of regions attached to one type of periodicity from periodic to aperiodic through almost-periodic and quasi-periodic. Combined with an estimation of the correlation support, a final decision about the periodicity of the signal is given. In case of a periodic signal, a way of estimating the global signal ratio is proposed. These new outputs are valuable for initializing more complex processing. All the algorithms proposed are fully automatic, one click use! Relevance of these indicators is shown on real-world signals, current and vibration measures mainly.

1. Introduction

In condition monitoring, a part of the information necessary for decision-making comes from scrutinizing a time measure or a transform of this measure. In this context, the frequency domain is commonly exploited. Oddly enough, it is not usual to consider the lag domain, whereas properties of this domain have long been known, specially for periodic signals⁽¹⁾. The autocorrelation or intercorrelation functions are most often used in system identification so as to extract the modes when the system is driven by a known input. For example, in the method of the random decrement vibration

signature^{(2), (3)}, the autocorrelation function is used because of its exact proportionality to the free vibration decay. Modal parameters can be estimated from cross-correlations⁽⁴⁾. The correlation dimension of a nonlinear dynamical system provides an estimation of its number of degrees of freedom⁽⁵⁾.

In the lag domain, resonant modes are not separated as in the frequency domain but noise is separated from the different harmonics. Noise is here defined as a very wide band spectral signal in which the components of interest are embedded. Investigate this domain offers the opportunity of characterizing the type of harmonicity of the signal which presents a link with the property of periodicity.

This paper focuses upon this interesting property of the autocorrelation function, its ability to characterize the periodicity property of a signal, and upon the link with the oscillating systems.

Section 2 makes an attempt to clear up the notion of periodicity relative to harmonic oscillations with an extension to random processes. Beyond a classification, the object of this section is to highlight the key role of the autocorrelation function. In section 3, two indicators are proposed in order to characterize the periodicity of a signal. First is based on the non-biased autocorrelation function and indicates a fundamental periodicity rate. Second is based on the biased autocorrelation and gives a dominant-power periodicity rate. Section 4 studies a 2D-plane defined by these two indicators, which allows the definition of regions attached to one type of periodicity from periodic to aperiodic through almost-periodic and quasi-periodic. Combined with an estimation of the correlation support, a final decision about the periodicity of the signal is given. In case of a periodic signal, a way of estimating the global signal ratio is proposed. These new outputs are valuable for initializing more complex processing. All the algorithms proposed are fully automatic, one click use! Section 5 shows relevance of these indicators on real-world signals, current and vibration measures mainly.

2. What is periodicity?

2.1 Periodic signals

Mathematically, a continuous signal $s(t)$ locally defined on the set $L^2(\mathcal{R})$ of finite energy signals is fully periodic with period T , when the signal exactly satisfies

$$s(t) = s(t + T). \quad (1)$$

For periodic signals, all the physical quantities, as amplitude and frequency, repeat at equal intervals.

As examples, a simple harmonic oscillation referred to as one sinusoidal signal

$$s(t) = A \sin(2\pi\nu t + \varphi), \quad (2)$$

with A the amplitude, $\nu=1/T$ the frequency and φ the phase, is periodic. A sum of m harmonic oscillations of period T_i is also periodic if the T_i are rationally linearly dependent. Then, all oscillating systems without energy dissipation are concerned.

Other examples of periodic signals are inharmonic oscillations, such as relaxation waves, that can be square wave or sawtooth, i.e. non-sinusoidal signals which can be decomposed in sum of sinusoidal or harmonic oscillations of pulsations multiple of a given pulsation referred to as the fundamental pulsation.

2.2 Quasi-periodic signals

Always in the case of continuous functions locally defined on the set $L^2(\mathcal{R})$ of finite energy signals, quasi-periodic signals are a generalization of periodic signals (see Figure 1). A signal $s_{qp}(t)$ is quasiperiodic⁽⁶⁾ with m periods T_1, \dots, T_m when

$$s_{qp}(t) = g \{s_1(t), s_2(t), \dots, s_m(t)\}, \quad (3)$$

where $g: \mathcal{R}^m \rightarrow \mathcal{R}$ is continuously differentiable and the m signals $s_i(t)$ are continuous periodic signals with respect to each period T_i . All the periods are required to be strictly positive and to be rationally linearly independent. This last constraint is the fundamental property of the quasi-periodic signals, which differentiates from periodic signals.

Quasi-periodic signals possess the following properties: addition and multiplication of quasi-periodic signals yield quasi-periodic ones. If $m=2$, the signal is said bi-periodic.

Therefore, an obvious example of quasi-periodic signals is the sum of m harmonic signals defined by (2) with all ratios of period T_i being irrational.

2.3 Almost-periodic and pseudo-periodic signals

In the case of continuous functions, almost-periodic signals are a generalization of quasi-periodic signals and then of periodic signals (see Figure 1). A signal $s_{ap}(t)$ is almost-periodic⁽⁷⁾ if every sequence $s_{ap}(t+T_\varepsilon)$ of translations of $s_{ap}(t)$, for all $\varepsilon>0$, has a subsequence that converges uniformly for T_ε in $(-\infty, +\infty)$, that is

$$\forall \varepsilon > 0, \exists T_\varepsilon \text{ such that } |s_{ap}(t+T_\varepsilon) - s_{ap}(t)| < \varepsilon. \quad (4)$$

There are several ways of defining classes of almost-periodic functions, based on notions of closure, of an almost-period and of translation. Each of these classes can be obtained as a closure, with respect to some metric, of the set of all finite trigonometric sums. (4) is the original definition proposed by H. Bohr (1947).

In oscillating systems, a quasi-sinusoidal oscillation is a usual denomination for oscillator variation close to a sine variation. According to definition (4) and in spite of its name, such an oscillation is not quasi-periodic but almost-periodic. That is the case of a Wien bridge oscillator.

Another example of almost-periodic signals is given by signals referred to as being pseudo-periodic signals (See Figure 1). A pseudo-periodic signal⁽⁸⁾ is defined locally on the set $L^2(\mathcal{R})$ of finite energy signals, such that all linear combinations of the translated signals have equivalent norm in $L^2(I)$, and $L^2(J)$, as soon as intervals I and J are long enough. The lower bound of I and J is referred to as the pseudo-period. Paley and Wiener has shown that a pseudo-periodic signal is an almost-periodic function⁽⁸⁾.

In other words, a pseudo-periodic signal can be decomposed into a periodic function along with a set of parameters that define the deviations of the pattern from true periodicity^{(9),(10)}. Let $s(t)$ be a real valued continuous function with compact support in $[0, T]$, a pseudo-periodic signal $s_{pp}(t)$ is

$$s_{pp}(t) = \sum_i m_i s(\alpha_i t + \beta_i) \tag{5}$$

where $s(t)$ is called the template function for $s_{pp}(t)$. The α_i are called the stretching parameters or scale frequencies, and represent the lengthening or shortening of the periods. The β_i are called the translation parameters in time, and allow nonuniform timing of the process, for instance, an acceleration or deceleration in a system. The m_i adjust the amplitude. In a musical context, the α_i correspond to the pitch of the waveform while the β_i correspond to the rhythm in which the waveforms appear. The template function plays a role analogous to that of a mother wavelet, while the stretching parameter is analogous to a scale factor. However, the template functions do not need to be orthogonal and do not form a basis, rather, they form a frame⁽¹⁰⁾.

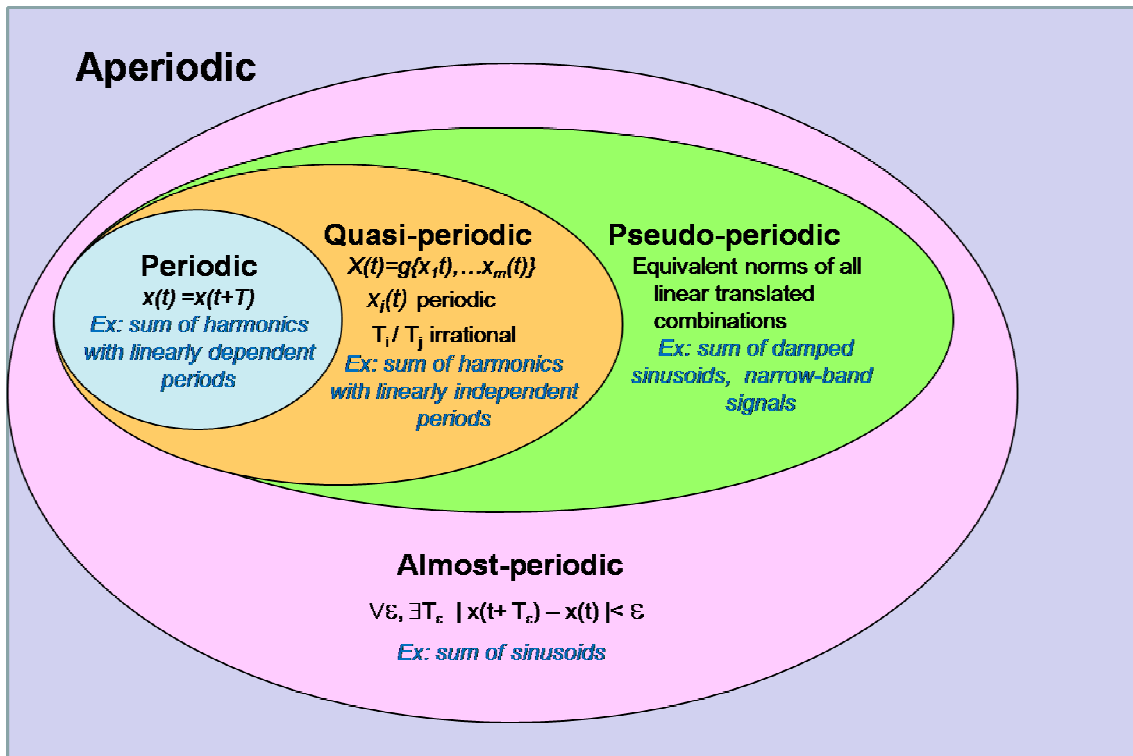


Figure 1. Periodic, quasi-periodic, almost and pseudo-periodic signals.

The most simple example is a damped harmonic oscillation also referred to as pseudo-sinusoidal oscillation,

$$s_{pp}(t) = Ae^{-t/\tau} \sin(2\pi\nu t + \varphi), \quad (6)$$

without scaling factor and time translation but with a damping parameter $1/\tau$, which modifies the amplitude of the periodic sine preventing signal $s_{pp}(t)$ to be periodic. All oscillating systems with energy dissipation are concerned.

2.4 Periodicity for stochastic processes

The previous definitions valid for deterministic signals can be extended to stationary stochastic processes but the equalities are true in a mean square sense⁽¹¹⁾. Let $x(t)$ be a stationary stochastic process. $x(t)$ is fully periodic with period T , when it satisfies

$$E\{x(t+T) - x(t)\}^2 = 0, \quad (7)$$

where T is the smallest number satisfying (7), $E\{.\}$ being the expected value. If (7) holds, then

$$R_x(\tau + T) = R_x(\tau), \quad (8)$$

for every lag τ , $\tau \neq 0$, of its autocorrelation function $R_x(\tau) = E\{x(t)x^*(t-\tau)\}$, the superscript * denoting the conjugate. Therefore, it is of interest to observe that the autocorrelation function of a periodic stochastic process is also periodic pointing out that $R_x(\tau)$ is a deterministic function. It is then straightforward to extend all previous definitions to stochastic processes by the following: definition of the periodicity in time in a mean square sense and equivalent definition than for deterministic processes with true equalities in lag domain.

As example, a laser light is a stochastic process, whose emitted frequency is dirturbed by a phase noise due to the finite spectral linewidth of the laser. This process having a narrow band spectrum relatively to the emitted frequency is pseudo-periodic.

In oscillating systems as rotating systems, vibrations contain sums of periodic signals which are not linearly dependent or rational in general. These signals are most often almost periodic. When measuring one vibration of a part of a system, the measure can be periodic. But, in many cases, even if the physical system leads to deterministic signals, a random noise often disturbed the measure, hence the interest of considering the autocorrelation function in order to define periodicity indicators.

3. Two periodicity indicators

Two indicators computed on the autocorrelation function are proposed. Besides the interesting property of the autocorrelation function face to periodicity as noticed in the previous section, another one is attractive for our purpose. When a signal is embedded

in a noise uncorrelated with the signal and with a spectral band wider than the signal of interest, in the lag domain, signal is alone for every lag greater than the noise correlation support, a helpful property for estimating the signal periodicity. In addition, estimating the autocorrelation function from a time signal requires no smoothing as it is in the dual frequency domain.

Let $r(t)$ be a stationary stochastic process, the sum of either a stochastic or deterministic signal $x(t)$ embedded into noise $n(t)$,

$$r(t) = x(t) + n(t), \quad (9)$$

$n(t)$ being white or a lightly colored. For $\tau = 0$, $R_r(\tau)$ is the sum of the power of signal $x(t)$, P_x , and of noise $n(t)$, P_n ,

$$R_r(0) = R_x(0) + R_n(0) = P_x + P_n, \quad (10)$$

If $n(t)$ is white and $x(t)$ is fully periodic with period T :

- For $\tau > 0$, $R_r(\tau)$ is fully periodic with period T

$$R_r(\tau) = R_x(\tau) = R_x(\tau + kT), k \text{ an integer.} \quad (11)$$

- For $\tau = kT$, $R_r(\tau)$ is equal to the signal power P_x

$$R_r(kT) = R_x(0) = P_x \quad (12)$$

- For $\tau \neq kT$ and $\tau \neq 0$, $R_r(\tau)$ is equal to $R_x(\tau)$ and lower to P_x

$$R_r(\tau) = R_x(\tau) < P_x \quad (13)$$

The periodicity or aperiodicity of $x(t)$ is unknown and has to be determined. Making use of the autocorrelation function leads to estimate first a fundamental periodicity of signal $x(t)$ and second a dominant-power periodicity. An estimation of the signal to noise ratio of the signal is also deduced.

3.1 Fundamental-frequency test

3.1.1 The general principle

From (12), lags multiple of the periodicity period of a signal match to the maxima of the autocorrelation function of the noisy observation. The fundamental-frequency test proposed is based upon this property and consists of detecting autocorrelation function maxima at regular intervals and with equal amplitude. The autocorrelation function being estimated, equation (12) becomes an approximation and the detection should include the statistical dispersion of the autocorrelation estimation.

The non-biased estimate of the autocorrelation function is used in order to define a ratio C_{fun} , which will be equal to 1 for a full periodic signal and 0 for an aperiodic signal. Finally, the fundamental frequency of signal $r(t)$ is estimated with the confidence of $C_{fun} \times 100\%$ and, in case of a periodic signal, a signal to noise ratio is estimated.

3.1.2 The algorithm

Let $\hat{R}_r[m]$ be the non-biased discrete estimation of $R_r(\tau)$,

$$\hat{R}_r[m] = \frac{1}{N-m} \sum_{n=m}^N r[n] r^*[n-m] \quad 0 \leq m \leq N-1, \quad (14)$$

with N the length of signal $r[n]$ sampled at F_s . In order to avoid both noise lag support of a possible non-white noise and high variance of the estimation in the high lag values, the test is applied within a lag support defined at each step of the algorithm.

The algorithm steps are:

1. Compute $\hat{R}_{r \max}$ the maximum of the autocorrelation function, within $[m_a, m_b]$ support, $\hat{R}_{r \max} = \hat{R}_r[m_{\max}]$ / $m_{\max} = \arg \max_{m \in [m_a, m_b]} \hat{R}_r[m]$.

The choice of values m_a and m_b is in 3.1.4.

2. Set $\hat{P}_{x \text{ init}}$ and $\hat{P}_{n \text{ init}}$, initial settings of P_x and P_n respectively, and check that $\hat{R}_{r \max}$ is not associated to noise,

$$\hat{P}_{x \text{ init}} = \hat{R}_{r \max} \text{ from (12) and } \hat{P}_{n \text{ init}} = \hat{R}_r(0) - \hat{P}_{x \text{ init}} \text{ from (10),}$$

$$\begin{array}{l} r[n] \text{ not aperiodic} \\ \hat{R}_{r \max} > \\ \leq \\ r[n] \text{ random noise} \end{array} c_b \times \sigma_{\max} \quad \text{with } \sigma_{\max} = \frac{\sqrt{N}}{|N - m_{\max}|} \hat{P}_{b \text{ init}}$$

The choice of value c_b is in 3.1.4.

3. Compute the lag set ξ such that autocorrelation values are close to $\hat{P}_{x \text{ init}}$, i. e; in an amplitude interval defined from the standard deviation of the estimator, and within $[m_c, m_d]$ support,

$$\xi = \left\{ m_j \in [m_c, m_d] / \hat{P}_{x \text{ init}} - c \sigma[m_j] - c_{ech} \hat{P}_{x \text{ init}} \leq \hat{R}_r(m_j) \leq \hat{P}_{x \text{ init}} + c \sigma[m_j] - c_{ech} \hat{P}_{x \text{ init}} \right\}$$

$$\text{with } \sigma[m] = \sqrt{\text{var}(\hat{R}_r[m])} \approx \frac{4N - 6m}{(N - m)^2} \hat{P}_{n \text{ init}} \hat{P}_{x \text{ init}} + \frac{N}{(N - m)^2} \hat{P}_{n \text{ init}}^2, \text{ standard deviation}$$

for all lags multiple of the signal period, given in ⁽¹³⁾ when noise is white from statistics ⁽¹²⁾ of estimator (14).

The set ξ makes sense only if card $\xi = J \geq J_{\min}$.

The choice of parameters m_c , m_d , c , c_{ech} and J_{\min} is discussed in 3.1.4.

4. Evaluate the fundamental-periodicity coefficient C_{fun} defined as the percentage of elements of ξ_d being close to the presumed signal fundamental period $1/\hat{f}_{fun}$, subject to a $\pm \text{toler}$ tolerance, ξ_d being defined as the set of distances between consecutive lags of ξ , and \hat{f}_{fun} being estimated from the median Med_{ξ_d} of ξ_d ,

$$C_{fun} = \frac{1}{J-1} \sum_{j=1}^{J-1} \delta(|d_j - Med_{\xi_d}|) \quad \text{with } \delta(u) = \begin{cases} 1 & \text{if } u < toler \ Med_{\xi_d} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{and } \xi_d = \left\{ d_j / d_j = m_{j+1} - m_j, m_j \in \xi \text{ and } 1 \leq j \leq J-1 \right\} \quad \hat{f}_{fun} = F_s / Med_{\xi_d}$$

The choice of tolerance parameter *toler* is discussed in 3.1.4.

3.1.3 A by-product: estimation of a global signal ratio

When the fundamental frequency has been estimated, the estimation of signal and noise powers, \hat{P}_x and \hat{P}_n respectively, are finally adjusted from elements of ξ ,

$$\hat{P}_x = \frac{1}{J} \sum_{j=1}^J \hat{R}_r[m_j] \quad \text{and} \quad \hat{P}_n = \hat{R}_r[0] - \hat{P}_x, \text{ with } m_j \in \xi \quad (15)$$

Hence, the signal to noise ratio *SNR* of the observed signal is estimated as

$$SNR = 10 \log(\hat{P}_x / \hat{P}_n), \quad (16)$$

this ratio having a sense only if the confidence ratio C_{fun} is close to 100%.

3.1.4 The by-default values

Fundamental-frequency test needs the setting of 9 parameters which has been chosen according to the investigation of real-world and synthetic signals in a significant data bank:

- $[m_a, m_b]$, a lag support within which $\hat{R}_{r_{max}}$ is computed: $m_a = 5\%N$; $m_b = 50\%N$; in any case, $m_a > 0$ and $m_b \leq 50\%N$.
- c_b , a tolerance factor applied to σ_{max} of $\hat{R}_{r_{max}}$: $c_b = 3$.
- $[m_c, m_d]$, a lag support within which ξ is computed: $m_c = 5\%N$; $m_d = 60\%N$; in any case, $[m_a, m_b] \subset [m_c, m_d]$.
- c , a tolerance factor applied to $\sigma[m] = \sqrt{\text{var}(\hat{R}_r[m])}$ in order to estimate ξ : $c=2$.
- c_{ech} , a tolerance factor applied to $\hat{P}_{x_{init}}$ in order to take sampling effects into account: $c_{ech}=0.05$; in any case $0.05 \leq c_{ech} \leq 0.1$.
- J_{min} , the minimal number of detected maxima : $J_{min} = 3$.
- *toler*, a tolerance factor applied to the median of ξ_d , i.e; to the presumed period, in order to compute the fundamental-periodicity coefficient: *toler* = 10%.

3.2 Dominant frequency test

3.2.1 The general principle

This second indicator, in addition to the first one, aims at estimating the periodicity of the dominant-power signal part. The theoretical principle is to determine the periodicity of the smoothed autocorrelation function, the smoothing being set so as to highlight the dominant-power harmonic. Unlike the first test, the autocorrelation function estimator should have a low variance so the biased estimator is preferred.

The autocorrelation function is filtered by a low-pass filter bank. For each filtered function, a periodicity ratio is computed; the suitable filter is the one given the first a

maximum ratio for a decreasing filter order (more and more narrow filter). A final one denoted as C_{dom} is selected such that the dominant-power frequency of signal $r(t)$ is estimated with the confidence of $C_{dom} \times 100\%$.

3.2.2 The algorithm

For sake of simplicity, the notation $\hat{R}_r[m]$ is kept while designing in this algorithm the biased estimation of the autocorrelation. Let $\hat{R}_r^k[m]$ be the filtered biased discrete

$$\text{estimation of } R_r(\tau), \quad \hat{R}_r^k[m] = FFT^{-1} \left[\frac{1}{N F_s} \left| \sum_{n=0}^{N-1} r[n] e^{-2\pi j n m / F_s} \right|^2 \times W^k(f) \right], \quad (17)$$

with N the length of signal $r[n]$ sampled at F_s , f the frequency variable, and $W^k(f)$ a filter bank defined from a window $W(f)$ raised to power k belonging to a filter power set. The choice of window $W(f)$ and of the filter power set is in 3.2.3.

The algorithm steps are:

1. Compute for each k the lag set ξ^k corresponding to minimal autocorrelation values within $[0, m'_a]$ support,

$$\xi^k = \left\{ m_j = \arg \min_{m_j \in [0, m'_a]} \hat{R}_r^k[m] \right\}.$$

The set ξ^k makes sense only if $card \xi^k = J' \geq J'_{\min}$

The choice of parameters m'_a and J'_{\min} is in 3.2.3.

2. Evaluate for each ξ^k a ratio $C_{dom}(k)$ defined as the percentage of elements of ξ_d^k being close to the presumed signal dominant-power period $1/\hat{f}_{dom}$, subject to a \pm *toler'* tolerance, ξ_d^k being defined as the set of distances between consecutive lags of ξ^k , and \hat{f}_{dom} being estimated from the mean $Mean_{\xi_d^k}$ of ξ_d^k ,

$$C_{dom}(k) = \frac{1}{J'-1} \sum_{j=1}^{J'-1} \delta \left(\left| d_j - Mean_{\xi_d^k} \right| \right) \quad \text{with } \delta(u) = \begin{cases} 1 & \text{if } u < toler' \cdot Mean_{\xi_d^k} \\ 0 & \text{otherwise} \end{cases}$$

and $\xi_d^k = \{ d_j / d_j = m_{j+1} - m_j, m_j \in \xi^k \text{ and } 1 \leq j \leq J'-1 \}$, $\hat{f}_{dom}(k) = F_s / Mean_{\xi_d^k}$

The choice of tolerance parameter *toler'* is discussed in 3.2.3.

3. Determine the dominant-power periodicity coefficient C_{dom} by selecting the maximum of $C_{dom}(k)$ at the lowest filter order k , and then the corresponding periodicity frequency,

$$C_{dom} = \max_{\min(k)} C_{dom}(k) \quad \text{and} \quad \hat{f}_{dom} = \hat{f}_{dom} \left(\arg \max_{\min(k)} C_{dom}(k) \right).$$

3.2.3 The by-default values

As for the previous one, the dominant-power frequency test needs the setting of 5 parameters which has been chosen according to the investigation of real-world and synthetic signals in a significant data bank:

1. $W(f)$, the window function for the filter bank is a Blackman-Harris 4T.
2. Filter power set = [0, 1, 2, 3, 4, 6, 8, 11, 15, 20, 30, 40, 60, 100, 200, 400];
3. $[0, m'_a]$, a lag support within which $\min \hat{R}_r^k[m]$ are computed: $m_a = 1\%, N$; it could be 100% but the computation time can then be very long.
4. J'_{\min} , the minimal number of detected maxima : $J'_{\min} = 4$.
5. $toler'$, a tolerance factor applied to the mean of ξ_d^k , i.e; to the presumed period, in order to compute the periodicity coefficient: $toler' = 10\%$;

4. A 2D plane for a final decision

4.1 Correlation tests and periodicity

Signals in condition monitoring are complex and can exhibit many type of periodicities. The aim of this paper is to assess the periodicity property of a signal with a simple and fast algorithm. The method proposed is a classification, which classes are characteristic regions in a 2D space defined from the two periodicity coefficients suggested in section 3. Figure 2 shows this space where the region definition comes from the investigation of real-world and synthetic signals in a significant data bank.

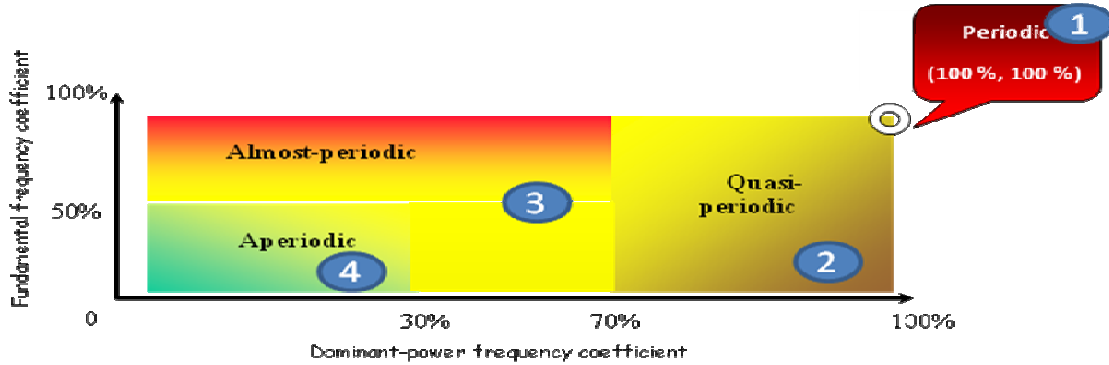


Figure 2. A first hypothesis about the periodicity of a signal.

4.2 The meaning of a correlation support

From the non-biased estimator of the autocorrelation function (see (14)), the correlation support is estimated as the maximum lag being greater than a threshold defined from the standard deviation of the estimator corresponding to a white noise only,

$$m_{cor_sup} = \max_{m \in [0, N]} \{m / \hat{R}_r[m] > C \sigma[m]\}, \quad (18)$$

if $m_{cor_sup} \geq 0,7N$ then $m_{cor_sup} = N$

with $\sigma^2[\tau] \approx \frac{N}{(N - \tau)^2} \hat{R}_r[0]^2$ as in section 3.1.2 with $\hat{P}_x = 0$, see also (15). The so-defined threshold $C \sigma[m]$ represents a limit above which the estimated autocorrelation no longer corresponds to the white noise hypothesis. The duration m_{cor_sup} corresponds

then to a rough estimate of the correlation support for the positive lags. Due to the high variance of the autocorrelation estimate, a correlation support greater than $70\%N$ is set to N . In the same way as in 3.1.4 and 3.2.3, the C by-default value has been set to 4.

Figure 3 shows along the lag axis how we have defined the signal spectral band relatively to the correlation support. Thus, five complementary intervals have been defined along this axis.



Figure 3. Correlation support and size of the signal spectral band.

4.3 The final decision

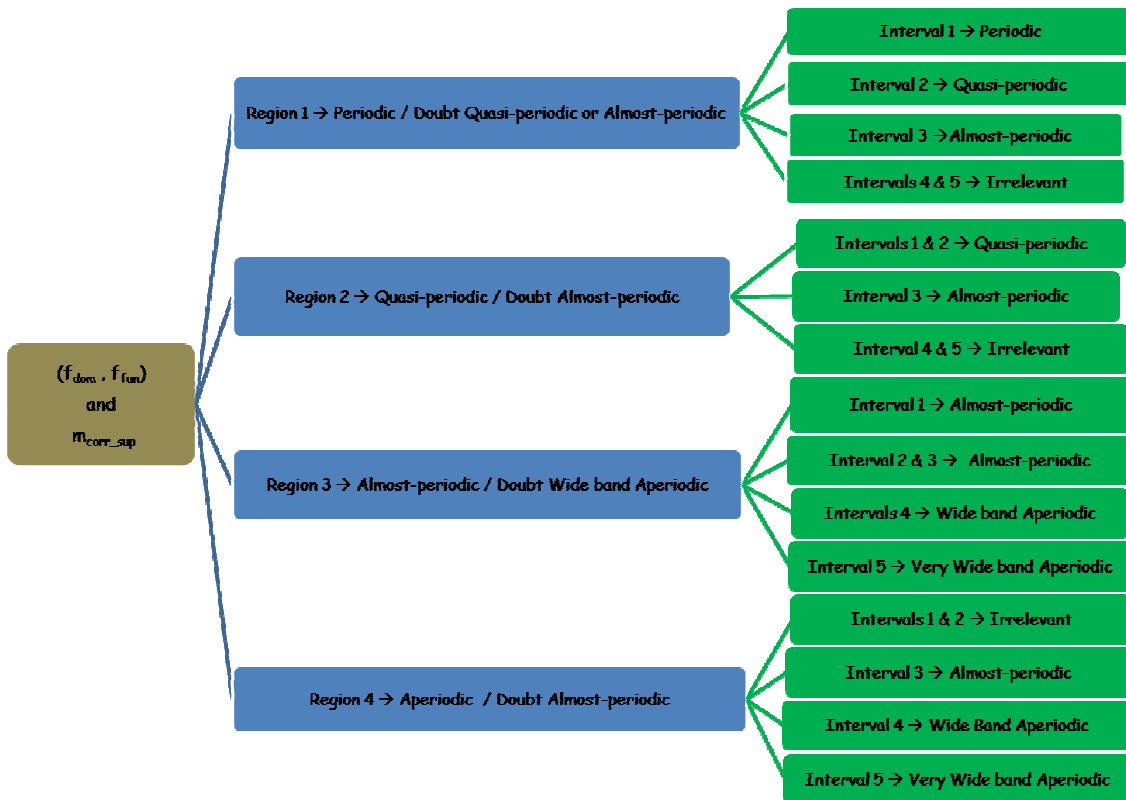


Figure 4. Classification of signals according to the periodicity test proposed. Region numbers come from the periodicity 2D-space (Figure 2) and interval numbers from the correlation support (Figure 3).

The addition of the correlation support information allows us to refine the estimation of the prevailing periodicity of the analysed signal. Figure 4 shows the final classification proposed in order to link the 4 regions defined in figure 2 and the 5 intervals defined in

figure 3. As previously stated, the validation has been investigated in a significant data bank of real-world and synthetic signals, as shown in the next section.

5. Results on simulations & real-world signals

The aim of this section is to apply the indicators proposed on simulated and real-world signals in order to estimate the signal periodicity in accordance with the definitions of section 2. All results are summed up in Table 1 and Table 2.

For all simulated signals, the classification is independent of the noise level, this result was expected according to the autocorrelation support of a white noise. It is less true for very low signal to noise ratio, around -10 dB, and for the damped sine. For the harmonic family, the fundamental frequency is estimated even if the fundamental is not present.

For real-world signals, the results obtained are complying with expert point of view. The fundamental frequency and power-dominant one can be similar or within the same range, which is the case for signals 6 and 7. In other cases the results are complementary.

Table 1. Results on simulations & real-world signals

	<i>Signal</i>	F_s	N	<i>Periodicity hypothesis</i>	n°
<i>Room Acoustics</i>	Impulse response	16 000 Hz	32 762	Wide band aperiodic and embedded in significant noise	1
<i>Lidar current</i>	Heterodyne current-coherent laser radar	40 500 Hz	81 200	Very wide band aperiodic and embedded in significant noise	2
<i>Magnetism</i>	Barkhausen noise	200 000 Hz	40 000	Wide band aperiodic and embedded in significant noise	3
<i>Bioacoustics</i>	Dolphin whistles	15 324 Hz	270 336	Very wide band aperiodic and embedded in significant noise	4
<i>Hydroacoustics</i>	Boat passing	5 000 Hz	100 000	Almost -periodic	5
<i>Vibratory 1</i>	Hydraulic noise in nuclear reactor pump		102 400	Periodic and embedded in significant noise	6
<i>Vibratory 2</i>	Hydraulic noise in oil station		50 000	Quasiperiodic and embedded in significant noise	7
<i>Vibratory3</i>	Vehicle cell	20 491 Hz	420 045	Quasiperiodic and embedded in significant noise	8
<i>Vibratory 4</i>	Rotating machine	6 365 Hz	65 536	Quasiperiodic and embedded in low noise.	9
<i>Simulated</i>	2_SW_periodic_80dB	1 000 Hz	100 000	Periodic and embedded in low noise	10
	2_SW_periodic_0dB	1 000 Hz	100 000	Periodic and embedded in significant noise	11
	2_SW_quasi_no_noise	1 000 Hz	100 000	Quasiperiodic and embedded in low noise	12
	2_SW_quasi_-10dB	1 000 Hz	100 000	Quasiperiodic and embedded in significant noise	13
	Damped_sine_80dB	1 Hz	10 000	Almost-periodic embedded in significant noise	14
	Damped_sine_0dB	1 Hz	10 000	Wide band periodic and embedded in significant noise	15
	Harmonics_no_noise_freq_fun_50Hz	1000 Hz	10 000	Periodic and embedded in low noise.	16
	Harmonics_10dB_freq_fun_50Hz	1000 Hz	10 000	Periodic and embedded in significant noise	17

6. Conclusions

A signal classification according to the signal periodicity is proposed. Autocorrelation domain is exploited in order to estimate characteristic frequencies with a confident level. Classes are currently chosen from the investigation of real-world and synthetic signals in a significant data bank. An improvement could be obtained by making profit of Monte Carlo simulations. Another lead to pursue would be the using of fuzzy logics in order to avoid nonlinear thresholds in the class definition.

Table 2. Items which lead to conclusions of Table 1.

n°	f_{fun}	C_{fun}	f_{dom}	C_{dom}	SNR	Region in 2D periodicity space	m_{cor_sup}	Interval along lag axis
1	492,31 Hz	6,9%	1656,96 Hz	31,3%	17,46 dB	3	1,6%	5
2	1760,87 Hz	8,1%	5389,93Hz	40,2%	-18 dB	3	0,0%	5
3	26 666,67 Hz	14,3%	13 123,36 Hz	40,0%	-12,33 dB	3	25,4%	4
4	901,41 Hz	5,4%	273,43 Hz	58,7%	-21,28 dB	3	3,3%	5
5	undefined	0,0%	472,08 Hz	48,4%	undefined	3	60,5%	2
6	2857,14 Hz	100,0%	2821,78 Hz	100,0%	-7,10 dB	1	100,0%	1
7	46,88 Hz	59,3%	46,77 Hz	100,0%	3,22 dB	2	100,0%	1
8	1,33 Hz	41,2%	351,76 Hz	91,6%	2,41 dB	2	100,0%	1
9	344,05 Hz	14,2%	46,97 Hz	100,0%	12,99 dB	2	100,0%	1
10	111,11 Hz	100,0%	109,88 Hz	100,0%	19,06 dB	1	100,0%	1
11	111,11 Hz	100,0%	109,99 Hz	100,0%	-0,09 dB	1	100,0%	1
12	22,22 Hz	69,0%	109,98 Hz	100,0%	16 dB	2	100,0%	1
13	38,46 Hz	21,6%	109,98 Hz	100,0%	-10,2 dB	2	100,0%	1
14	0,29 Hz	46,2%	0,30 Hz	69,0%	3,75 dB	3	36,0%	3
15	0,29 Hz	46,0%	0,30 Hz	69,0%	-2,86 dB	3	26,6%	4
16	50,00 Hz	100,0%	49,38 Hz	100,0%	73,35 dB	1	100,0%	1
17	50,00 Hz	100,0%	49,38 Hz	100,0%	9,99 dB	1	100,0%	1

References

1. Bendat J. S., Piersol A. G., Random data analysis and measurement procedures. Wiley, 1986.
2. Vandiver J.K., Dunwoody A.B., Campbell R.B., Cook M;F., A math. basis for the random decrement vibration signature analysis technique, J. of Mechanical, 1982.
3. Ibrahim S. R., Random decrement technique for modal identification of structures. J. Spacecraft Rockets 14, 696-700, 1977.
4. Shen F., Zheng M., Feng Shi D., Xu F., Using the cross-correlation technique to extract modal parameters on response-only data, Journal of Sound and Vibration, Volume 259, Issue 5, Pages 1163-1179, 30 January 2003.
5. Logan D., Mathew J., Using the correlation dimension for vibration fault diagnosis of rolling element bearings. MSSP, vol. 10, no3, pp. 241-250, 1996.
6. Benoist-Gueutal P., Mathématiques pour la physique, Ed. De Boeck &Larcier, 2005.
7. Corduneanu C., Gheorghiu N. and Barbu V., Almost periodic functions, Chelsea publishing Company, 1989.
8. Paley R. and Wiener N., Fourier transforms in the complex domain, New-York, American Mathematical Society, 1934.
9. Wong H., Sethares W. A., Estimation of pseudo-periodic signals, IEEE ICASSP, Montreal, Quebec, Canada, May 17-21, 2004.
10. Sethares W. A, Repetition and pseudo-periodicity, Tatra Mt. Math. Publ. 23, pp. 1-16, 2001).
11. Papoulis A., Probability, Random variables and stochastic processes, McGraw-Hill Book Company, 1965.
12. Oppenheim A. V., Schafer R. W., Digital Signal processing, prentice hall Editions, Englewood Cliffs, N.J., 1975.
13. Durnerin M., A strategy for interpretation in spectral analysis- Detection and characterization of spectrum components (in French), PhD, Grenoble INP, September 21, 1999.