CEAI, Vol. 9, No. 2, pp. 59-67, 2007 Printed in Romania

# SIMULATION MODELS OF DEFECT ENCODING VIBRATIONS

**Janetta Culita[1], Dan Stefanoiu[1], Florin Ionescu[2]**

[1] "Politehnica" University of Bucharest, Faculty of Automatic Control and Computer Science
Splaiul Independentei 313, Bucharest – 060042, ROMANIA
[2]University of Applied Sciences, Department of Mechatronics, Konstanz, GERMANY
E-mails: jculita@yahoo.com, danny@indinf.pub.ro, ionescu@htwg-konstanz.de

***Abstract:*** *Comparing to signals such as speech or seismic, the vibration generated by oscillations of a mechanical system in exploitation is quite regular. At a glance, good models of vibration can be designed, without special precautions. However, some basic characteristics of this signal should be considered when constructing the model. For example, vibration can serve to search for faults threatening a system. In this case, the main properties of mechanical vibration yielding identification and isolation of faults have to be a priori known. The paper goal is to overview these properties and to present the most used models that allow performing the automatic fault diagnosis.*

***Keywords:*** *mechanical vibrations, signal analysis, McFadden-Smith vibration models.*

## 1. INTRODUCTION

Vibration generated by mechanical systems is interesting, mainly for its capacity to encode information about the defects or faults. Several distinct efforts in detection of defects can be noted, but only in the last few decades the vibration has became crucial for automating this process. The earliest method, which dates back to the first days of machinery (and which is still in use today) is founded on a trained observer or listener referred to as *(expert) analyst*. A person with great deal of experience in working with a particular machine or engine can detect flaws in operating machinery, by simply "watching" or "listening" it. Very often, the resulting diagnosis, based on empirical observations and deductions, is amazingly accurate, but difficult to model. Other subsequent attempts became more systematic and used some parameters, such as: the lubricant temperature (which, unfortunately, gives a hint about defects after they are already severe), the oil cleanness (which requires an exhaustive and often inefficient analysis), the noise level of acoustic emission (which is in general enabled only by already fatigued elements), etc.

But the most efficient methods in early detection of defects are using Signal Processing (SP) techniques [13]. Differently from many typical SP applications, where the noise attenuation is a fundamental requirement, when using vibrations, exactly the noise is the most concerned part within the analysis. This is due to the fact that not the natural oscillations of machinery encode the defective behavior, but the noise corrupting them. Moreover, the applications revealed that the signal-to-noise ratio (SNR) is extremely small for vibrations encoding information about defects. Therefore, the models of vibration used in fault detection and diagnosis (fdd) are models of their noisy parts, encoding types and severity degrees.

The paper is organized as follows. The next section emphasizes the differences between *oscillation* and *vibration* concepts and reveals a general model of vibration birth. The most common models of vibrations and their sources are described in sections 3-5. Conclusions and a references list are completing the article.

## 2. GENERAL VIBRATION GENESIS MODEL

The Theory of oscillations refers a distinctive part of Mechanical Sciences and traditional Mathematics. This theory mostly relies on the study of differential equations describing systems forced to evolve at their stability limit (associated to resonance oscillation). The oldest oscillatory system model with one freedom degree is expressed as a two order differential equation, provided by applying the second principle of Dynamics [8], [5]. The solution of this equation is referred to as *oscillation*, that suggests the natural behavior of the mechanical oscillatory system.

In general, some of system parameters (mass, damping, elasticity) are unknown. Their estimation relies on measured data that usually, are affected by random noises. This is the case of non-deterministic model. The solution of the new equation is an oscillation that includes noises, referred to as *vibration*. Only the (measured) vibration data generates the model.

In Figure 1, one can clearly distinguish between a natural attenuated oscillation –left and a vibration data – right. In general, but depending on the SNR, the original oscillation is difficult to distinguish from the graphic of a vibration.

When measuring vibrations from a mechanical system, several signals are combined together, in order to generate the data. In Figure 2, a general model of vibration birth is depicted.

Thus, the running mechanical system under test can generate 3 signal classes: natural oscillations ($x$), interference signals ($u$) due to interactions between its different parts; defect encoding noise ($d$), indicating that something is wrong with one or more of its parts. Together with the environmental (background) noise ($e$), the signals above are mixed in a way that generates the *crude (mechanical) vibration* ($w$). This is converted into an *(electrical) vibration* ($v$) by means of a sensor connected to a transducer (which can also distort the crude vibration).

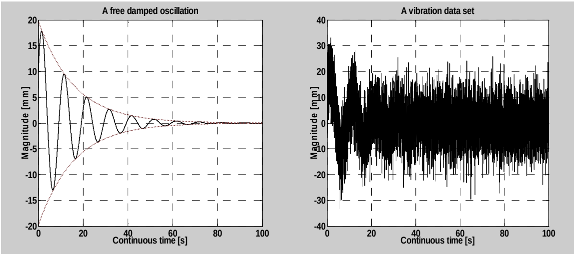
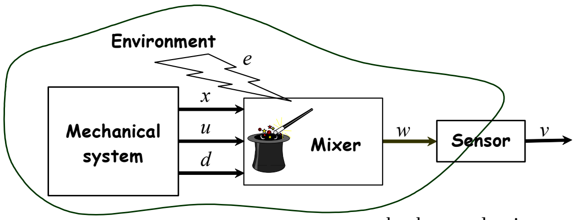



**Fig. 1.** A free damped oscillation and vibration data.




$x$ = natural oscillation
$w$ = crude vibration (mechanical)
$v$ = vibration data (electrical)
$e$ = background noise
$d$ = defect noise
$u$ = interferences

**Fig. 2.** General vibration genesis model.

## 3. ON MECHANICAL SYSTEM MODELS

The main problem in fdd using vibration data is to extract the defect encoding component (noise) and to remove or significantly attenuate all the other components (including the natural oscillations). If the machinery is defect free, then the resulting signal should be quasi white noise (as part of background noise). To solve this problem, several models of blocks in Figure 2 have to be accounted.

The model of mechanical system is often so complex that it cannot be really tackled. Fortunately, this model seems not to be so important in solving the problem above. Sometimes only simple equations or estimations of natural oscillation frequencies are sufficient. They can be used to identify the main harmonics of the oscillations to be removed. The other blocks are more important and some appropriate models are described within next two sections.

In order to illustrate how natural frequencies of a mechanical system could be estimated, take for example a radial-axial bearing with rolling balls. It consists of two races (inner and outer), several rolling balls and a cage aiming to preserve the balls clearance. The natural oscillation frequencies of bearing are intimately determined by its rotation speed and geometry.

In general, one of the two races is rotating and another one is fixed. The case when both races are rotating is unusual, but possible. Therefore, one can consider the general case when a relative rotation between races exists, without specifying for the races whether they are fixed or not. Denote by $\nu_r$ the frequency of relative rotation between races (in Hz). Then a set of 5 natural oscillation frequencies could be derived: the ball pass frequency on the outer race ($\nu_{out}$); the ball pass frequency on the inner race ($\nu_{in}$); the cage rotation frequency with respect to the outer race ($\nu_{cout}$); the cage rotation frequency with respect to the inner race ($\nu_{cin}$); the ball rotation frequency ($\nu_b$). Concerning the geometry, all manufacturers have to comply to the existing standards. Usually, manufacturers describe their products in catalogues [6]. The geometrical parameters necessary to derive bearings natural frequencies are the following (3 of them being illustrated in Figure 3): the *number of balls* $n_b$, the *pitch diameter* $\phi_p$, the *ball diameter* $\phi_b$ and the *contact angle* between ball and races when the bearing is under load, $\alpha$. This latter is also standardized, the values being listed in catalogues together with the other constructive characteristics.

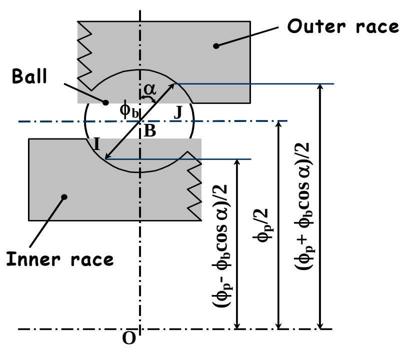



**Fig. 3.** Geometry of a rolling balls bearing.

The parameters above lead to the following natural frequencies for the bearing (by assuming the ball does not slide on the races):

$$\nu_{out} = \frac{n_b \nu_r}{2}\left[1 - \frac{\phi_b}{\phi_p}\cos\alpha\right], \nu_{in} = \frac{n_b \nu_r}{2}\left[1 + \frac{\phi_b}{\phi_p}\cos\alpha\right]$$

$$\nu_{cout} = \frac{\nu_{out}}{n_b}, \quad \nu_{cin} = \frac{\nu_{in}}{n_b}$$

$$\nu_b = \frac{\nu_r}{2}\frac{\phi_p}{\phi_b}\left[1 - \left(\frac{\phi_b}{\phi_p}\cos\alpha\right)^2\right] = 2\frac{\nu_{cin}\nu_{cout}}{\nu_r}\frac{\phi_p}{\phi_b} \quad (1)$$

Note that the highest frequency is $\nu_{in}$ (and thus the inner race is the most exposed to failures, together with the balls) and: $\nu_{cin} + \nu_{cout} = \nu_r$; $\nu_{in} + \nu_{out} = n_b \nu_r$. Beside the natural frequencies above, some multiples of their values are also present into harmonic contents of vibration, but, usually, their amplitudes decay quite fast. Practically, beyond 7-10 times the highest frequency ($\nu_{in}$), no harmonic due to natural oscillations is acting into vibration data. This means that the noisy part of vibration could be roughly isolated by high-pass pre-filtering of data, such that the natural frequencies and their significant multiples be removed.

For more complicated mechanical systems, the natural frequencies are often not derived from motion equations (unsuccessful attempt due to equations complexity), but simply estimated by measuring some parameters.

## 4. ON MIXER AND SENSOR MODELS

One might believe that the mixing mechanism is a kind of magic, since the algorithm is actually unknown. However, several plausible models have been proposed so far, especially in case of bearings and gears. For example, in [7], all the signals are added: $w \equiv x + u + d + e$, which constitute a coarse model, since the defect noise seems not only to add to, but also to modulate the natural oscillation. However, the authors have shown that, for their invention, the mixing model is not so important.

The concept of oscillation modulation, especially by defect noise, has been used in many papers such as [4], [1], [3]. Simple amplitude modulation has been considered in the beginning $w \equiv (d + e)(x + u)$ where the noises $e$ and $d$ play the role of *enveloping* signal. A more complicated model (claimed as "more realistic" by the authors) would be

$w \equiv (1+d+e)(x+u)$, where addition and modulation are both considered. This model have been considered when analyzing the vibration generated by sea waves, rolling element bearings, gears or working wheels of pumps (turbines).

Another more sophisticated model relies on frequency modulation instead of amplitude modulation as before:

$$w(t) = \sum_{k\in\mathbf{Z}} A\left(t + kT\lambda(t)\right), \quad \forall\, t \in \mathbf{R}, \qquad (2)$$

where $A \equiv x+u$ is the harmonic modulated signal with average period $T$ and $\lambda \equiv 1+d+e$ is the noisy frequency modulating part. This model has been used for analysis of vibrations produced by electrical machines with pulsating electromagnetic torque.

Note that in the first three models, the noise $e$ is just added to the noise of interest, $d$. Actually, the original expressions in the cited papers have not explicitly considered both noises. Only in case of defect free machinery, $e$ is present. There is no serious reason to consider that the noises are modulated each other, i.e. to replace $d+e$ by $d \cdot e$. Moreover, the additive noise model is suitable for SP analysis framework, since usually, the SNR ($d$ – signal and $e$ – noise) is now reasonable.

Other authors considered that the noise $d+e$ is input of an oscillatory system ([10], [11]). Moreover, the defect noise is considered as series of shock pulses with some period $T$ and constant amplitude $A$, produced by defects (especially on rolling parts in contact):

$$d(t) = A\sum_{k\in\mathbf{Z}} \delta_0\left(t + kT\right), \quad \forall\, t \in \mathbf{R}, \qquad (3)$$

where $\delta_0$ is the unit impulse. This model reveals a different approach: the noise is now filtered by (or convoluted with the impulse response of) an oscillatory system. But the periodicity of defect noise hypothesis is not actually viable. For example, a chop on a ball inside a bearing produces quasi-random non-periodical shocks. Also, this noise could be generated without shocks, especially by worn mechanical parts. However, the idea that the defect noise is basically generated by visible or microscopic quasi-random shocks has been largely accepted today. Actually, although the model introduced in [10] is regarding the bearings and the transmission path of vibration from defect to measured data, it has been adopted (and furthermore generalized) by many authors in the literature. More details on this model are described within the next section.

The key device in vibration data acquisition is the sensor. An exhaustive description of sensors (and their models) employed on this purpose is performed for example in [8]. Within applications, the seismic sensors with piezoelectric crystals (and especially the *accelerometers*) are the most preferred, for their good properties of sensitivity and linearity. Beside their small size and weight, there are at least 2 main reasons to take piezoelectric sensors into consideration. Firstly, they are able to introduce only an imperceptible distortion between the crude and measured vibrations in absence of defect noise. (Also, the very low frequency noises are attenuated.) Secondly, the defect noise, if present, is emphasized within measured vibration. This behavior could be easily seen by looking at the typical frequency response of such a sensor. The most accurate model is considered to be a rational transfer function with 3 poles and one zero (accounting both the mechanical and electrical parts):

$$H(s) = G\underbrace{\frac{s}{s-s_0}}_{electrical}\underbrace{\frac{1}{(s-s_1)(s-s_2)}}_{mechanical} \qquad (4)$$

Here, $s$ is the Laplace variable, whereas the amplification factor $G$ and the poles $s_0$, $s_1$, $s_2 = s_1^*$ ($\mathrm{Re}\, s_{1,2} < 0$) are derived from constructive specifications. The specific pole placement in model (4) leads to the frequency response drawn in Figure 4.

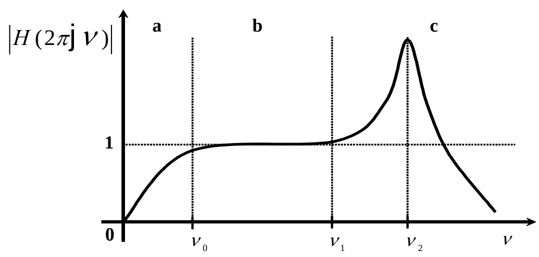



**Fig. 4.** Typical frequency characteristic of sensors.

There are 3 zones: (a) a quite small bandwidth attenuation zone for low frequency signals, up to $\nu_0 \in (0, 100]$ Hz (also removing the constant component, if present); (b) a large bandwidth quasi-constant pass band zone between $\nu_0$ and $\nu_1 \in [10, 60]$ kHz (that replicates the input signal with minimum distortions); (c) the "resonance" high frequency zone beyond $\nu_1$, with central natural/resonance pulsation $\nu_2 \in [60, 100]$ kHz and rapid attenuation of very high frequency signals. The electrical part of sensor (the transducer) controls the lowest frequency sub-band, but the resonance is due to the mechanical part. As mentioned before, the resonance zone is far away from vibrations localization band (usually, $[0, 20]$ kHz), with more than 90 % of energy concentrated in the first half ($[0,10]$ kHz). Obviously, if $h$ is the sensor impulse response then $v \equiv h * w$ (where '$*$' stands for *convolution*. ).

In fdd, the resonance zone (c) is the most concerned, since only the sensor resonance is able to extract the defect noise from crude vibration, independently of how the mixer works. This property is either not explained or the explanation is extremely unclear in the overwhelming majority of related publications, except for [7], where the authors have stressed that *the sensor resonance frequency is the only carrier of information about defects in mechanical systems*. Actually, only the defect noise is able to put the sensor in state of (attenuated) resonance. The other crude vibration components are replicated approximately as they are, except for the very low and very high frequency ones, which are attenuated (as proven by Figure 4). In general, very low and very high frequency noises are produced by environmental sources. So, the sensor performs moreover a preliminary useful denoising of crude vibration.

An intuitive representation of how the sensor works is illustrated Figure 5. Consider a crystal glass (or a musical tuning fork) that oscillates in front of a wall with high quality surface such that, when touching the wall, the glass speed is null (Figure 5(a)). Normally, the crystal sound could hardly be perceived. Eventually, if the oscillation frequency is sufficiently high, then the glass starts ringing with small intensity. Assume now that a (small) bump appeared from the wall in the contact zone with the glass (Figure 5(b)). Obviously, every time the glass touches now the wall, the crystal starts ringing with high intensity since the impact speed is non null. This resonance sound dominates the weak ringing before (that continue to exist however) and indicates the existence of wall irregularity. The defect severity is encoded by the acoustic intensity of dominating sound. Actually, Figure 5 also reveals another interesting phenomenon: the spectrum of vibration $v$ has been split into 2 sub-bands. The low frequency one groups all the components of crude vibration excepting for the defect noise, the spectral interference and the high frequency background noise, which are grouped inside high frequency (resonance) sub-band. Unfortunately, the separation is not sharp, the zones are strongly overlapped.

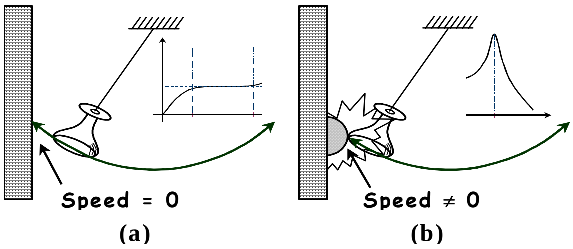



**Fig. 5.** Resonance effect of sensors – an intuitive image.

And yet, practically, two vibration components could be identified within $v$: a non-defect encoding one ($v_{nd}$) and a defect encoding one ($v_d$). How these components are mixed inside $v$, it is difficult to say, especially because their spectra are not disjoint. All one can say is that the energy of $v_{nd}$ is mainly localized at low frequency (since the spectrum of natural oscillations generated by the mechanical system decays rapidly with increasing frequency), whereas the spectrum of $v_d$ is mainly localized at middle or high frequencies (since the defect noise and, eventually, the interference forced the sensor to resonate).

A simple model of $v$ could be additive: $v \equiv v_{nd} + v_d$. If one accounts the different modulation models of $w$ (described before), another models could be: $v \equiv v_{nd} \cdot v_d$, $v \equiv v_{nd} * v_d$, etc. This model is less important in vibration analysis than the structure of defect encoding component, $v_d$. According to sensor behavior, shocks and interference are carried by the resonance signal. Hence, the signal $v_d$ could have the shape depicted in Figure 6.

Practically, $v_d$ is a modulated resonance signal:

$v_d \equiv v_{de} \cdot v_{dr}$, where $v_{de}$ is the *envelope signal* (low frequency) and $v_{dr}$ is the *carrier resonance signal* (high frequency). Depending on sensor mechanical characteristics, the natural resonance pulsation is $\omega_n = |s_2|$, the envelope decays as $\exp(\mathrm{Re}(s_2))$ and the carrier (damped) pulsation is $\omega_d = |\mathrm{Im}(s_2)|$.

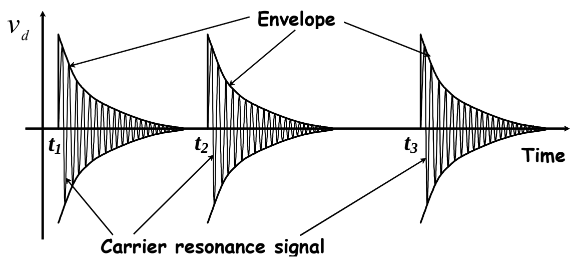



**Fig. 6.** Modulation of defect encoding vibration.

The shocks producing the resonance could not be necessarily equally spaced, as illustrates the variable distance between instants $t_1$, $t_2$ and $t_3$ in Figure 6. This corresponds well to the real shocks produced by defects.

Unfortunately, not only the shocks are stimulating the sensor to resonate, but the interference too. In this case, they are carried by the resonance signal as well, creating confusion between real and virtual defects. Interference signals could appear within mechanical systems with a certain degree of complexity, even they are defect free. But the interference effect could be neglected for simple machinery. A quite simple and efficient method to remove the interference components from vibration is described in [6]. This method mainly relies on *Fourier filtering,* a technique that yields canceling the energy inside some narrow bands of the spectrum. Moreover, the center frequencies of narrow bands (the interference frequencies, in fact) are automatically detected within this method by simply inspecting the real and imaginary parts of vibration *Fourier Transform* (FT) and selecting the "very extreme values" (according to an energy criterion) that have to be removed.

The concept of *envelope* used in this context is related to signal modulation techniques and not to original definition from SP, which relies on the *analytical signal* construction by using the *Hilbert filter* [13]. Constructing the envelope of a signal is easier than in SP, as described within [9] and [12]. It practically consists of 2 operations: *wave rectification* (*full*, by taking the absolute values or *half,* by taking the non negative values of data) and *smoothing* by (exponential) interpolation between maxima. A device performing this task is referred to as *peak follower* or *envelope detector* and is commercialized today as an integrated chip.

The envelope of a signal has the energy concentrated towards lower frequency sub-bands than the signal itself. This effect is mainly due to the rectification, but the interpolation also has some contribution. As mentioned above, the rectification could be full or half type, as suggested by Figure 7.

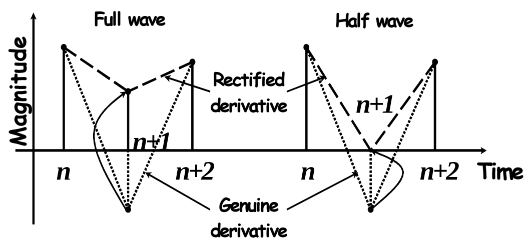



**Fig. 7.** Wave rectification principles.

The wave sample at current instant $n+1$ is either accounted without its sign (for full rectification) or totally ignored if negative (for half rectification), as illustrated by the bent arrows in the figure. Observe how the derivatives are changed from abrupt genuine values to smoother rectified values in Figure 7. It is well known that the signal derivative reflects the local (*instantaneous*) frequency contents: low instantaneous frequency signals have smooth derivatives, whereas within high instantaneous frequency signals the derivatives are abrupt. Hence, by rectification, the frequency content is shifted from high to low frequency sub-band. The interpolation is practically employed only for half wave rectification, because the full wave rectification provides in general not null values at the current instant (there are no gaps inside the resulted signal).

In case of high frequency signals (vibrations), full or half wave rectification is almost equivalent to taking the whole envelope (i.e., practically, the interpolation doesn't count any more). Actually, in many research reports, the authors just ignore the interpolation and deal only with full wave rectification. One can prove by simulation that, in case of these signals, the spectra of rectified signals are practically indistinguishable from the spectra of their envelopes.

## 5. MCFADDEN-SMITH DEFECT MODELS

One of the first sound models concerned with vibration generated by single point defects in bearings was introduced by McFadden and Smith in [10]. In this subsection, their model and its generalizations are succinctly presented, because all these models seem to describe how the defect is encoded by vibration in a more accurate and natural manner than the previous models.

The first model is based on the following assumptions (see Figure 8):

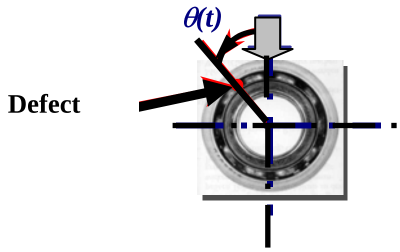



**Fig. 8.** Assumptions within McFadden-Smith model.

a) The outer race is frozen, whereas the inner race is rotating with constant frequency $\nu_r$ (the ball pass frequency on the inner race ($\nu_{in}$) is computed following (1)).
b) The defect is located only on the inner race (but extensions to another cases are possible).
c) The bearing has a radial load and the defect is located at the angle $\theta(t) = 2\pi\nu_r t$ (at instant $t \in \mathbf{R}$) by the vertical axis).

Denote when the defective inner race rotates, *in absence of load*. Then it is very likely that $p$ constitutes a series of equally spaced *shock pulses* of certain amplitude (in general determined by the *severity degree* of defect), produced when balls strike the defect:

$$p(t) = p_0 \sum_{k\in\mathbf{Z}} \delta_0(t - kT_{in}), \quad \forall t \in \mathbf{R}, \tag{5}$$

where: $p_0$ is the severity degree of defect; $\delta_0$ is the (continuous time) unit impulse ($\delta_0(t) = 1$ if $t = 0$ and null otherwise); $T_{in} = 1/\nu_{in}$ is the period between successive shock pulses. One can prove that the extension of FT to a series of periodic pulses like $p$ is also a series of periodic pulses, but in frequency. Thus:

$$P(\nu) = p_0\nu_{in} \sum_{k\in\mathbf{Z}} \delta_0\left(\nu - k\nu_{in}\right), \quad \forall \nu \in \mathbf{R}. \tag{6}$$

The load model is given by Stribeck equation:

$$q(\theta) = q_0 \left[1 - \frac{1}{2\varepsilon}(1-\cos\theta)\right]^n, \theta \in [-\theta_{\max}, \theta_{\max}] \tag{7}$$

where: $q_0$ is the maximum load intensity; $\varepsilon$ is the load distribution factor; $n = 3/2$ for ball bearings (or $n = 10/9$ for roller bearings); $\theta_{\max} > 0$ is the maximum angle for which the load is transmitted towards the inner race (see Figure 9).

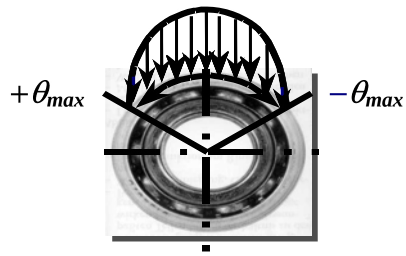



**Fig. 9.** Stribeck model of load distribution.

Since the inner race rotates (i.e. $\theta$ varies in time in (7)), the load on the defective zone is variable. By convention, the load is null when $\theta$ varies beyond the range specified in (7). Hence:

$$q(t) = q_0\left[1 - \frac{1}{2\varepsilon}(1-\cos(2\pi\nu_r t))\right]^n \tag{8}$$

for $(2\pi\nu_r t)\%2\pi \in \left[-\theta_{\max}, +-\theta_{\max}\right]$ and 0 otherwise. The notation $\alpha\%2\pi$ denotes the “reminder” of angle $\alpha$ in range $(-\pi, +\pi]$ (i.e. after all integer multiples of $2\pi$ have been removed). The FT of load, $Q$, could be derived in a complicated closed form that is not very interesting.

When the bearing is under load, the shock pulses are modulated by the load distribution. Thus, from (5) and (8), one could derive that the defect noise (denoted by $d$ in Figure 2) is given by $d \equiv p \cdot q$, which, by the Inverse Convolution Theorem [13], leads to $D \equiv P * Q$. This noise is combined with the main oscillation of the bearing, denoted by $x$ in Figure 2. As the inner race rotates, but the transducer is frozen on the outer race, the oscillation and the defect noise are received as signals varying in amplitude according the position of transducer. Moreover, the defect position directly affects the intensity of received signal. This is maximum when the defect is located at minimum distance to the transducer and minimum when the distance is maximum. Practically, the oscillation $x$ that modulates the defect noise looks like in Figure 10.

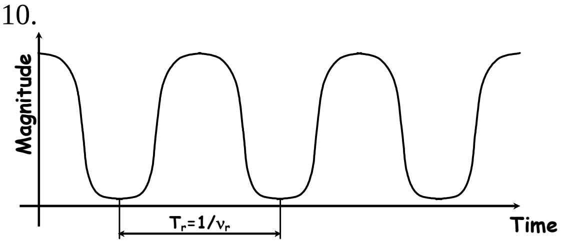



**Fig. 10**. Defect noise amplitude perceived by sensor.

The waveform is periodical with period $T_r = 1/\nu_r$ and thus, the crude vibration $w$ is expressed as the product between the oscillation and the defect noise (so, the defect noise is not additive) $w \equiv d \cdot x \equiv p \cdot q \cdot x$, or, equivalently, as a convolution between 3 FTs: $W \equiv P * Q * X$. Here, $X$ – the FT of oscillation $x$ – is simply a couple of *rays* (spectral lines) at $\pm\nu_r$.

The crude vibration is transformed into measured vibration data by means of the sensor. The sensor transfer function is mainly a first order linear one, due to electrical constructive part (usually, the mechanical part provides negligible time constants), as proven by equation (4). Thus, the sensor impulse response is very close to a decaying exponential with some (unknown) time constant $T_e > 0$:

$$h(t) \cong h_0 e^{-\frac{t}{T_e}}, \ \forall t \in \mathbf{R}_+ \tag{9}$$

This system is *causal* (null impulse response for negative instants). Its input is the crude vibration $w$, whereas, to the output, the vibration $v$ is generated:

$$\begin{cases} v \equiv h * w \equiv h * (p \cdot q \cdot x) & \text{(in time)} \\ V \equiv H \cdot W \equiv H(P * Q * X) & \text{(in frequency)} \end{cases} \tag{10}$$

The vibration spectrum reveals 2 types of rays:

a. located at multiples of $\nu_r$, such as: $\pm\nu_r$, $\pm 2\nu_r$, $\pm 3\nu_r$, ...;

b. located at frequencies such as: $\pm\nu_{in}$, $\pm\nu_{in} \pm \nu_r$, $\pm\nu_{in} \pm 2\nu_r$, $\pm\nu_{in} \pm 3\nu_r$, ..., $\pm 2\nu_{in}$, $\pm 2\nu_{in} \pm \nu_r$, $\pm 2\nu_{in} \pm 2\nu_r$ $\pm 2\nu_{in} \pm 3\nu_r$, etc.

In general, the defect is announced by rays at frequencies $\{m\nu_{in} + n\nu_r\}_{m,n\in\mathbf{Z}}$, i.e. the ball pass frequency on the inner race (where the fault appeared) is directly involved within the spectrum. In general, natural frequencies (1) could reveal defects located on the bearing constructive parts that generated them ($\nu_{in}$ for inner race, $\nu_{out}$ for outer race, $\nu_{cin}$ or $\nu_{cout}$ for cage, etc.). This idea seems to be very important for bearings fdd. It is thereby confirmed by one of the simpler and widely employed fdd methods in bearings industry: the Envelope Analysis [2]. The method also confirmed that the height of rays (the *spectral power*) or their decaying speed encodes the severity degree of defect.

The model (10) has been employed in many applications, in spite of its obvious limitations. It also was the source of some generalizations aiming to overcome its drawbacks. The same authors provided first generalizations in [11]. The main goal was to extend the model from single to multiple-point defects. In this aim, the following assumption has been added: *if the mechanical system is linear, the FT of crude vibration* ($w$) *is the superposition of FT produced by every defective point*.

Experimentally, this hypothesis has been verified in a satisfactory manner and only for 2 defective points located on the inner race. Authors noted however in [11] that the sum of FT could be distorted by another FT provided by interference signals ($u$ in Figure 3). One of such interference signals could be generated by cage rotation, which is distorting the FT phases. In this case, for a single point defect on inner race, the rays indicating the possible defect are located at frequencies $\{(mn_b + n)\nu_r - m\nu_{cout}\}_{m,n\in\mathbf{Z}}$ (since (1) involves: $\nu_{in} = n_b(\nu_r - \nu_{cout})$).

The idea introduced in [11] has been better exploited later by other authors (Su and Lin), in

[14]. Their model could be represented like in Figure 11. Although the mathematics are not very accurate in [14], the authors' point of view could be easily understood. The shock pulses $p$ (5) input the system in Figure 11. At the first tage, they are modulated by different loads applied on the mechanical system (also referred to as *contact energy* points): $\{q_i\}_{i\in\overline{1,m}}$ (for example, expressed by Stribeck equations (7) or (8)). The defect noise $p$ is the sum of all $m$ modulated shock pulses, as first manifestation of superposition hypothesis. This arrives to the sensor following different transmission paths generated by the relative location of defects: $\{x_i\}_{i\in\overline{1,n}}$. Each oscillation modulating now the defect noise is a finite sum of elementary harmonic signals with certain frequencies. The $n$ crude vibration components $\{w_i\}_{i\in\overline{1,n}}$ obtained so far are converted by the sensor into $n$ raw vibration signals.

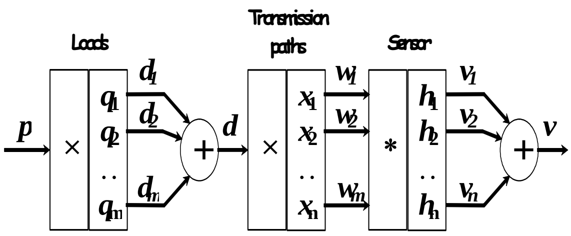



Fig. 11. Multiple-point defect model (Su and Lin).

The sensor impulse response function in (9) is replaced here by a set of $n$ weighting functions, harmonic, but exponentially decaying (looking like in Figure 1, left). The superposition assumption is invoked again to build the raw vibration $v$ from the resulting vibration signals.
This model is general enough, but quite complex. Many unknown parameters have to be estimated (or set), such as the frequencies of all involved harmonic signals, the sensor damping constants, etc.

## 6. CONCLUSION

Within this article, the idea that the spectrum of raw vibration is correlated to the natural frequencies produced by the defective parts in mechanical systems has been outlined. Unfortunately, there is no simple rule to put into correspondence the location of the defects and such frequencies. The problem of multiple defect diagnosis remains still open.

## REFERENCES

[1] Azovtsev, Y.A., Barkov, A.V., Carter, D.L. *Improving the Accuracy of Rolling Element Bearing Condition Assessment*, The 20-th An. Meet.of the Vibr. Instit, Saint Louis, Missouri, USA, June 1996, pp. 27–30.

[2] Barkov, A.V., Barkova, N.A. *Diagnostics of Gearing and Geared Couplings Using Envelope Spectrum Methods*, The 20-th Annual Meeting of the Vibration Institute, Saint Louis, Missouri, USA, June 1996, pp. 75–83.

[3] Barkov, A.V., Barkova, N.A. *Non-linear Signal Models in Vibroacoustic Machine Diagnosis*, The 21-st Annual Meeting of the Vibration Institute, Saint Louis, Missouri, USA, June 1997.

[4] Barkov, A.V., Barkova, N.A., Mitchell, J.S. *Condition Assessment and Life Prediction of Rolling Element Bearings. Part 2*, Journal of Sound and Vibration, Sept. 1995, pp. 27–31.

[5] Bedford, A., Drumheller, D.S. – *Introduction to Elastic Wave Propagation*, John Wiley & Sons Inc., Chichester, U.K., 1994.

[6] Carter, D.L. *Rolling Element Bearing Condition Testing Method and Apparatus*, United States Patent No. 5,477,730, December 26, 1996.

[7] Harting, D.R., Taylor, J.W. *Demodulated Resonance Analysis System*, United States Patent No. 3,862,663, October 22, 1974. http://www.uspto.gov/go/ptdl/

[8] McConnell K.G. *Vibration Testing. Theory and Practice.*, John Wiley & Sons Inc., NY, 1995.

[9] McFadden, P.D., Smith J.D. *Vibration Monitoring of Rolling Element Bearings by the High-Frequency Resonance Technique. A Review*, Tribology International, Vol. 17, No. 1, 1984, pp. 3–10.

[10] McFadden, P.D., Smith J.D. *Model for the Vibration Produced by a Single Point Defect in a Rolling Element Bearing*, Journal of Sound and Vibration, Vol. 96, No. 1, 1984, pp. 69–82.

[11] McFadden, P.D., Smith J.D. *The Vibration Produced by Multiple Point Defects in a Rolling Element Bearing*, Journal of Sound and Vibration, Vol. 98, No. 2, 1985, pp. 263–273.

[12] McFadden, P.D., Toozhy, M.M. *Application of Synchronous Averaging to Vibration Monitoring of Rolling Element Bearings*, Mechanical Systems and Signal Processing, Vol. 14, No. 6, 2000, pp. 891–906.

[13] Proakis, J.G., Manolakis, D.G. Digital Signal Processing. Principles, Algorithms and Applications., Prentice Hall, New Jersey, 1996.

[14] Su, Y.T., Lin, S.J. *On Initial Fault detection of a Tapered Roller Bearing – Frequency Domain Analysis*, Jour. Sound & Vibration, Vol. 155, No. 1, 1992, pp. 75–84.