THE ARCHIVES OF TRANSPORT

VOL. XXII NO 4 2010

10.2478/v10174-010-0027-9

Nonlinearity and Intermodulation Phenomena Tracking as a Method for Detecting Early Stages of Gear Failures

Stanisław Radkowski*
Marcin Jasiński**

Received November 2010

Abstract

An understanding of vibroacoustic signal is required for the robust and more effective detection of early stage gear failure. In the paper the possibility and method of time varying vibration decomposition are discussed. It is shown that analysing the coupling between the structure's components changes from linear to nonlinear or to other kind of nonlinearity together with intermodulation phenomena can be used as measure in structural health monitoring.
In addition on an analytical connection is investigated between the tracking method and the physics of the gear contact process based on the idea of higher-order spectra analysis, bispectral analysis specially.

Keywords: vibroacoustic diagnostics, modulation phenomena, bispectrum, gears

1. Introduction

The development of maintenance is determined on the one hand on relevant security of operated systems and the need for reducing the threat to the environment on the other.

In such circumstances it is natural to adopt Condition Based Maintenance (CBM), which means introduction of a system of evaluating the technical condition based on the collected data related to the parameters of a machine's operation and the

* Institute of Automotive Engineering, Warsaw University of Technology, ras@simr.pw.edu.pl
** Institute of Automotive Engineering, Warsaw University of Technology, jachuu@simr.pw.edu.pl

parameters of residual processes as well as performance of preventive maintenance based on the forecasted damage (failure) occurrence, which we could term as "just – in – time" maintenance.

From this point of view it is the implementation of proactive operational strategy that becomes particularly important.

Early detection of defects and determination of their causes occupies a special place among the diagnosis methods, specially the analysis of vibroacoustic. Let us note that the process of defect formation can lead both to intensification of non-linear phenomena as well as to occurrence of non-stationary effects even if in the early stages the intensity of defects is small while the growth of the level of vibration and noise is negligible in contrast with emergency situations. Here let us only note that the emergence of defects and the low-energy phases of their development are most often accompanied by local disturbance of the signal's run, which may result in tangible changes of the signal's frequency structure that are additionally variable in time. Such a situation inclines one to formulate the diagnosis of origin of defects while relying on the diagnostic information carried by non-stationary disturbance and non-linear effects.

Let us note that the analysis of low-energy pulse-type disturbance, causing broadband response with small amplitude, calls for accounting for not only the information on the changes of the signal's power but also on its phase, which points to the need for reaching beyond the information contained in the second-order process. This is so because even though the correlation function provides a sufficient description of the Gauss process with the mean value equal zero, still in the case of non-Gauss distributions of probability it is accordingly the correlation function of the power spectrum that provides partial information on the process.

For example the information on emergence of a defect can be contained in the low-energy components of the signal that are carried across the structure of a machine from the measurement source as a result of modulation of a relevant carrier function. Thus when examining signals attention is particularly devoted to the analysis of amplitude-and-phase modulation of a signal and occurrence of non-linear and non-stationary effects.

The actual signal can contain both, the component generated by the diagnosed kinematic pair as well as by the components transmitted over structure of the examined object which are generated by other kinematic nodes. This brings us to the necessity of solving the issue of relevant separation of diagnostically-useful information. It is connected with the issue of developing a relevant procedure which could be a part of the algorithm for diagnosing the low-energy phases of defect development.

As has been demonstrated by to-date research [1], significant diagnostic information is contained the higher order moments, which underscores the significance of non-linear phenomena in the detection of defect development. Such possibilities are not offered by the analysis of power spectrum which relies on the assumption of mutual independence of respective frequency components, which is a consequence of adopting the linearity and the superposition in the applied methods of power spectrum determination. In majority of cases the development of low-energy defects cannot be adequately presented by means of linear models. An example confirming this situation is the phenomenon of transformation of a harmonic function in which the additional frequency components are coupled. Similarly, the transition through a non-linear phenomenon with a square component of a signal consisting of two harmonics with various frequencies and various initial phases will lead to the emergence among the additional components, at the output points of system, also of components which maintain the same relationships between the original frequencies and phases as those found in the input signal. This phenomenon is called in publications [3] as quadratic-phase coupling.

The above presented introduction points to the need of broader reference to the methods of signal analysis that enable detection of phase relationships between the harmonics and the modulation, inter-modulation and mutual modulation effects which enable examination of resultant multi-dimensional signals.

2. Diagnostic Model in Detection of Low-energy Defects

While attempting to develop a model oriented on such defects one should on the one hand consider the issue of examining the signal's parameters from the point of view of their sensitivity of to low-energy changes of the signal and, on the other, the issue of quantification of energetic disturbances occurring in the case of defect initiation.

Let us assume that the degree of damage D is the dissipated variable that covers the changes of the structure's condition due wear and tear:

$$\mathrm{d}E_d(\Theta, D_0) = \frac{\partial E_d(\Theta, D_0)}{\partial D}\mathrm{d}D + \frac{\partial E_d(\Theta, D_0)}{\partial \Theta}\mathrm{d}\Theta \tag{1}$$

where:

$\mathrm{d}E_d = \dfrac{df(D, \Theta, \gamma(\Theta))}{d\Theta}$

$\gamma(\Theta)$ – the parameter describing how big a part of the dissipated energy $\mathrm{d}E_d$ is responsible for structural changes,

Θ – operating time.

Bearing in mind the possibility of diagnosis of the origin and the development of low-energy phases of defect formation, when the extent of the original defect can be different in each case, let us analyze this issue more precisely.

To examine this problem let us recall here the two-parameter isothermal energy dissipation model proposed by Najar [3] where:

$$\mathrm{d}E_{d_s} = \mathrm{d}E_d - \mathrm{d}E_{dq} = T\mathrm{d}s = \sigma_\Theta \mathrm{d}D \tag{2}$$

where:

dE_{d_q} – energy transformed into heat,
dE_{d_s} – energy responsible for internal structural changes,
T – temperature,
ds – growth of entropy.

The expression (2) shows that the growth of the dissipated variable D is attributable to the dE_{d_s} part of energy, which is the dissipated part of dE_d energy, that causes the growth of entropy ds.

The role of the multiplier determining the relation between the increments of dissipated variable and the entropy is played by the dissipation stress σ_Θ.

The assumption of T = constans results in independence of dissipation-related loss $dE_{d_s} = dE_{\Theta_s}$, thus following integration the expression (2) takes the following form:

$$E_{d_s} = T\Delta s \tag{3}$$

The derivative of defect development energy related to D, when $E_f(D_0) \leq \frac{1}{2}E\varepsilon^2$, means the boundary value of deformation energy and takes the following form:

$$\frac{dE_{d_s}}{dD} = \frac{E_f(D_0)(1-D_f)(1-k)D^{-k}}{D_f^{1-k} - D_0^{1-k}} \tag{4}$$

For a defined initial defect of D_0 and for a defect leading to damage D_f, relationship (4) will have the following form:

$$\frac{dE_{d_s}(D)}{dD} = (1-k)E_{D_0,f}(k)D^{-k} \tag{5}$$

Let us note that parameter $E_{D_0,f}$ is an exponential function of power k, similarly as the whole derivative. While referring to the second rule of thermodynamics for irreversible processes we will assume the following in the contemplated model:

$$\frac{dE_{d_s}(D)}{dD} \geq 0 \tag{6}$$

Thus for the assumed model to be able to fulfil condition (6), the exponent must meet the requirement of $k \leq 1$. In addition, while referring to the rule of minimization of dissipated energy, the conditions of permissible wear process [3] show that the change of exponent k is possible as the defect develops.

To examine this problem let us assume that the exponent shows a straight line dependence on the extent of damage:

$$k(D) = a + bD \tag{7}$$

For damage of small magnitude the linear approximation seems to be sufficient and enables description of defects whose emergence is characterized by small growth of defect energy (see Figure 1).

Thus while defining the set of diagnostic parameters we should pay attention to the need for selecting such a criterion so that it will be possible to identify defects whose emergence is characterized by small growth of defect-related energy.

While contemplating this issue let us assume that vibroacoustic signal is real and meets the cause-and-effect requirement, which means that it can be the base for creating an analytical signal.

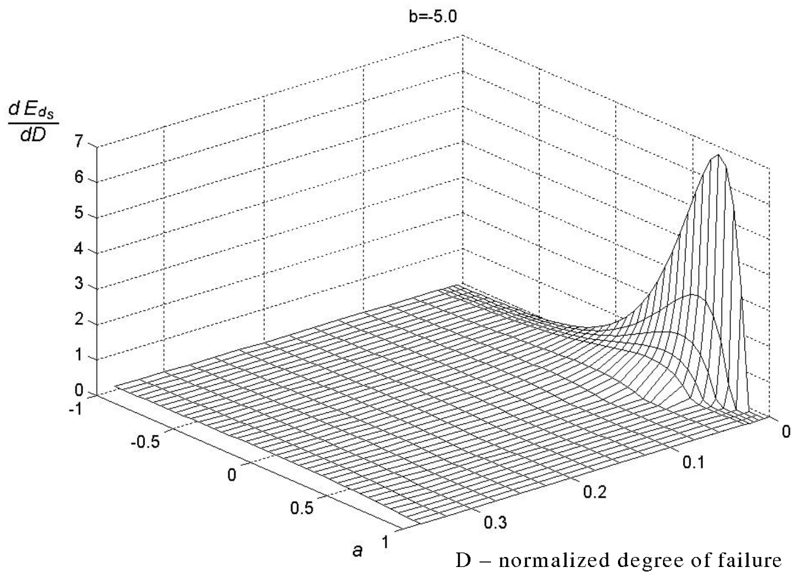

Fig. 1. Change of energy of defect development for small D

The central issue is how to extract the relevant diagnostic information and use it in the fore-casting process. Let us note that the measured vibroacoustic signal is a real signal which fulfils the requirement of causality. Thus, by using the measured signal $z(t)$ and a defined formalism, we are able, by means of addition of an imaginary part of $v(t)$, to form an analytical signal:

$$a(t) = z(t) + jv(t) \tag{8}$$

In accordance with the theory of analytical functions the real and the imaginary components are functions with two variables x and y.

Let us assume that the analysis of analytical signal is conducted on the basis of observation of the changes of the length of vector A and the phase angle of φ:

$$z(x, y) + jv(x, y) = A(\cos\varphi + j\sin\varphi) \tag{9}$$

Thus,

$$z = A\cos\varphi, \quad v = A\sin\varphi \tag{10}$$

which means that the measured signal is an orthogonal projection of the vector A on the real axis. Basing on Cauchy-Riemann condition, finally we get:

$$\frac{\mathrm{d}z}{\mathrm{d}\tau} = \frac{\mathrm{d}A}{\mathrm{d}\tau}\cos\varphi - A\sin\varphi\frac{\mathrm{d}\varphi}{\mathrm{d}\tau} \tag{11}$$

The obtained relationship, in accordance with our expectations, presents an equation which enables the analysis of the measured signal on the basis of observation of A and φ. What simultaneously captures our attention is the fact that for the low-energy processes, when we can disregard the changes of vector length and assume that $A \cong \text{const}$, the whole information about the changes in the measured signal is contained in the phase angle:

$$\frac{\mathrm{d}z}{\mathrm{d}\tau} = -A\sin\varphi\frac{\mathrm{d}\varphi}{\mathrm{d}\tau} \tag{12}$$

3. Intermodulation and Mutual Modulation Phenomena

Thus the occurrence of a defect and development of its low-energy phases are accompanied by a disturbance of the operation of kinematic node leading to change of power distribution between the spectrum components. The shares of respective components will be determined by multi-parameter modulation processes of four different modulated functions whose carrier frequencies correspond to the basic frequencies of harmonics appearing in the solution.

Let us note that the presented models showing the influence of defect origin and development are clearly associated with the development of the phenomenon of modulation of the signal's parameters. This is only partly confirmed by the results of analyses of the spectra generated by defective gears. Another effect that should be in addition taken into account is the occurrence of non-linear effects. For that reason the vibroacoustic signal generated by a defect should presented as a higher order components, which includes cases of non-linearity of the second, third or even fourth order:

$$y(t) = x(t) + \varepsilon(x(t))^2 + \sigma(x(t))^3 + \delta(x(t))^4 \tag{13}$$

It is sufficient in the case of systems or sets of machines with not so complex dynamic and kinematic structure. Let the results of simulation, as presented in [4], be an example of the fact that such a model of signal generation is unable to explain in a sufficient degree the spectrum's change in connection with a developing defect.

Figures 2-3 present the example of evolution of the spectrum simulated by the model of non-defective and defective two-stage toothed gear.

While explaining the problem let us note that by taking into account the second degree, we as result are able to follow the vibroacoustic signal generated by two pairs of toothed wheels, that is by two sources. This means that the measured

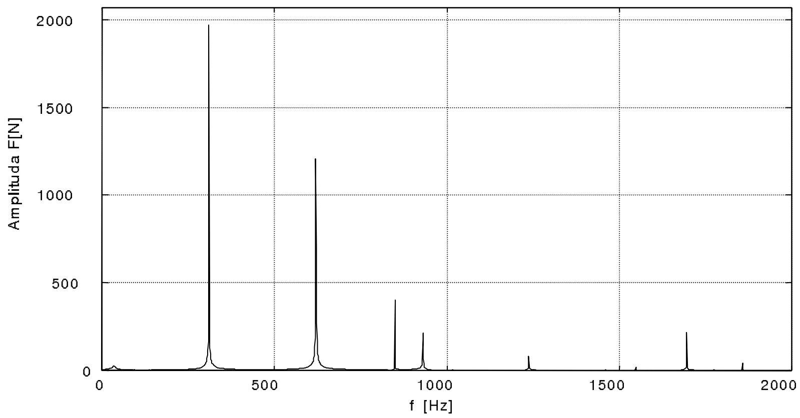

Fig. 2. Amplitude spectra of simulated response of bearings in an "ideal" toothed gear

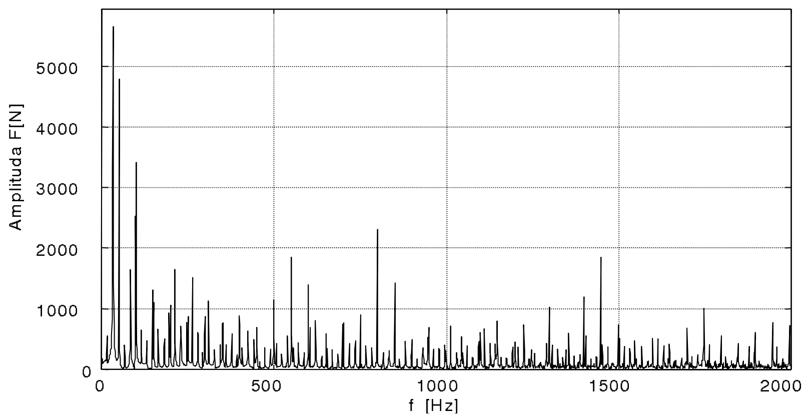

Fig. 3. Amplitude spectra of simulated response of bearings in a toothed gear with 2^{nd} degree pitch error the pinion

vibroacoustic signal is the sum of the minimum values from the two signals whose complexity, in this case the modulation of the parameters, depends on defect development. Bearing in mind the associated developing impact of non-linear effects one should expect coincidence of the influence of both phenomena. Thus one should include the phenomena of mutual modulation and inter-modulation in the description of changes of the frequency structure of the signal generated by the defective two-stage toothed gear.

The task associated with these phenomena of spectrum structure change demonstrates the operation of broadband amplifiers [5]. In accordance with the proposed terminology, mutual modulation denotes that the sum of minimum values of the two signals with equal amplitude, of which at least one is amplitude modulated, is subjected to influence of non-linear elements.

For a change, inter-modulation means the phenomenon of generation of additional frequency components as a result of non-linear influence on the sum of two non-modulated signals with the same amplitude and carrier frequencies of f_1 and f_2.

The additional components of $mf_1 + nf_2$ type which emerge as a result of interference, where m, n are integers, respectively cover the low and the high frequency bands and for the two signals are classified by a non-linearity function (of the second, third, fourth order, etc.).

In the case of non-linear interference of a bigger number of signals, the inter-modulation order is defined by the sum of harmonics that make up a defined component:

$$I.M. = n_1 f_1 \pm n_2 f_2 \pm n_3 f_3 \pm \ldots \tag{14}$$

where:

n_1, n_2, n_3, ... – integers defining the harmonic of the next harmonic.

f_1,f_2, f_3, ... – basic frequencies of interfering signals.

The value of the amplitudes is the factor having significant influence on the components generated by higher order inter-modulation. This is so since along with the growth of the order of inter-modulation, the amplitude of the generated component decreases. Let us consider the influence of these phenomena and the non-linearity of various types on the modulated and non-modulated signals:

$$y(t) = A_1 \left(1 + M_1 \cos 2\pi f_{r1} t\right) \cdot \cos 2\pi f_1 t + A_2 \left(1 + M_2 \cos 2\pi f_{r2} t\right) \cdot \cos 2\pi f_2 t \tag{15}$$

$$\begin{gathered} y(t) = A_1 \left(1 + M_1 \cos 2\pi f_{r1} t\right) \cdot \left(\cos 2\pi f_1 t + m_1 \sin 2\pi f_{r1} t\right) + \\ + A_2 \left(1 + M_2 \cos 2\pi f_{r2} t\right) \cdot \left(\cos 2\pi f_2 t + m_1 \sin 2\pi f_{r2} t\right) \end{gathered} \tag{16}$$

$$x(t) = y(t) + \varepsilon \cdot y^2(t) + \sigma \cdot y^3(t) + \delta \cdot y^4(t) \tag{17}$$

Then, while illustrating the influence of the discussed phenomena let us analyse the changes of the spectrum's frequency structure for both types of modulation based on relationships (15)-(16) and the non-linearity model having the form of relationship (17). The obtained results are presented in Fig. 4-5.

It is common knows that the power spectrum based methods cannot detect the phase relationship between different frequency components and additionally suppresses the phase information. It is therefore necessary to explore spectral measures of higher order, like the bispectral measures, to detect various forms of phase coupling between frequency components. Investigating this possibility we try to write, the bispectrum in form [4, 6]:

$$B\left(f_x, f_y\right) = E\left[S\left(f_x\right) S\left(f_y\right) S^*\left(f_x + f_y\right)\right]. \tag{18}$$

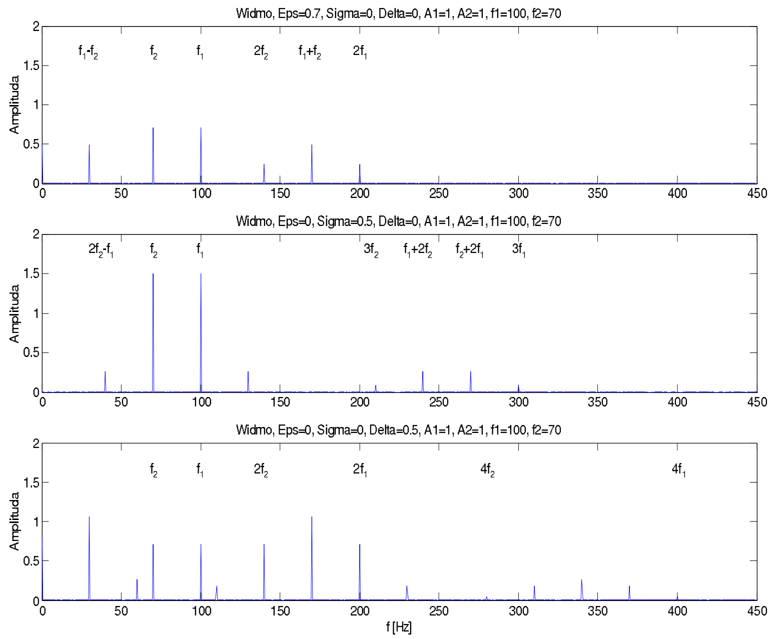

Fig. 4. Inter-modulation of various orders

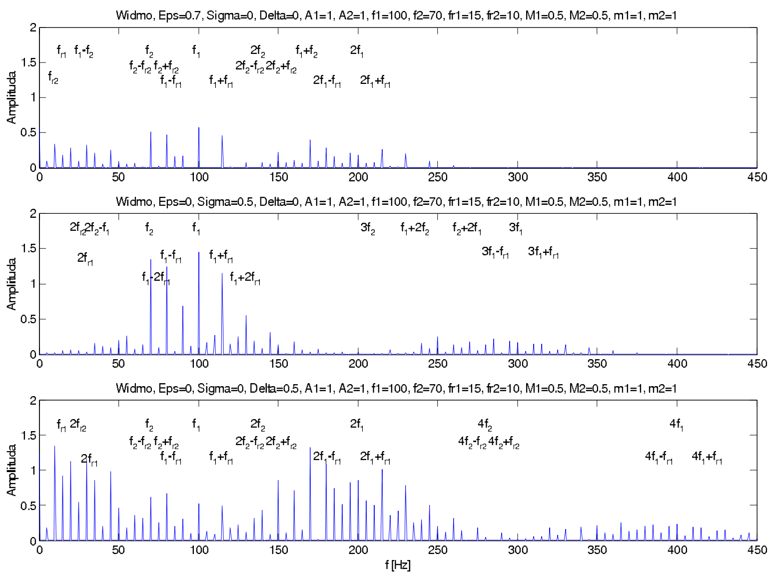

Fig. 5. Mutual modulation of a signal for various orders of non-linearity

It is easy o see the bispectrum is complex and that the bispectral values depend on two frequencies f_x and f_y. Writing the Eq. (6) in terms of amplitude and phase quantities one becomes:

$$B\left(f_x, f_y\right) = |S(f_x)|\left|S\left(f_y\right)\right|\left|S\left(f_x + f_y\right)\right| e^{j\Theta_\beta\left(f_x, f_y\right)} \tag{19}$$

where $\Theta_\beta\left(f_x, f_y\right) = \Theta(f_x) + \Theta\left(f_y\right) - \Theta\left(f_x + f_y\right)$ and is called the biphase.

Using the fast Fourier transform (FFT) algorithm it is possible to calculate the raw bispectrum:

$$B_i\left(f_x, f_y\right) = S_i(f_x) S_i\left(f_y\right) S_i^*\left(f_x + f_y\right) \tag{20}$$

The raw bispectrum can be estimate over the inner triangular region $0 \leq f_y \leq f_x$, $f_x + f_y = f_\mu/2$. This is sufficient for a complete description of the bispectrum, since, due to symmetry in the $f_x - f_y$ plane of the bispectrum, all of the significant information is contained in the principal domain that consists of the inner and outer triangles [6].

In addition to the basic bispectrum, the bispectrum diagonal slice is defined as:

$$B(f, f) = E\left[S(f) S(f) S^*(2f)\right] \tag{21}$$

with $f_x = f_y = f$.

The bispectrum diagonal slice is specially useful in detection of nonlinear effect.

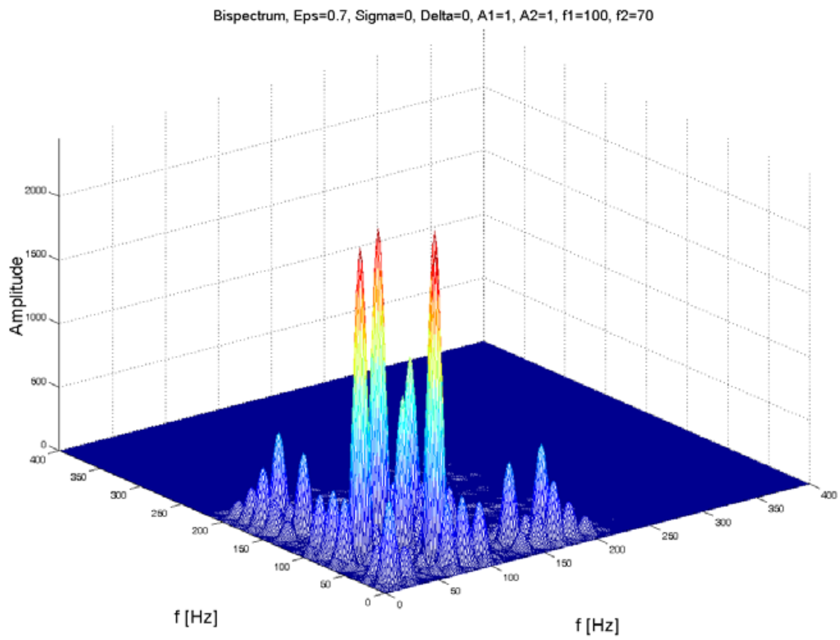

Fig. 6. Bispectrum of a signal with mutual modulation and squared non-linearity

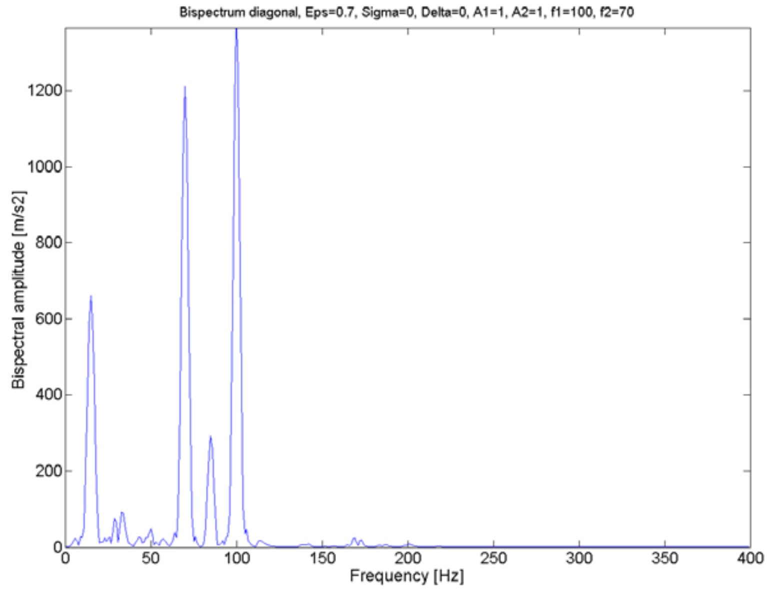

Fig. 7. Bispectrum diagonal of a signal with mutual modulation and squared non-linearity

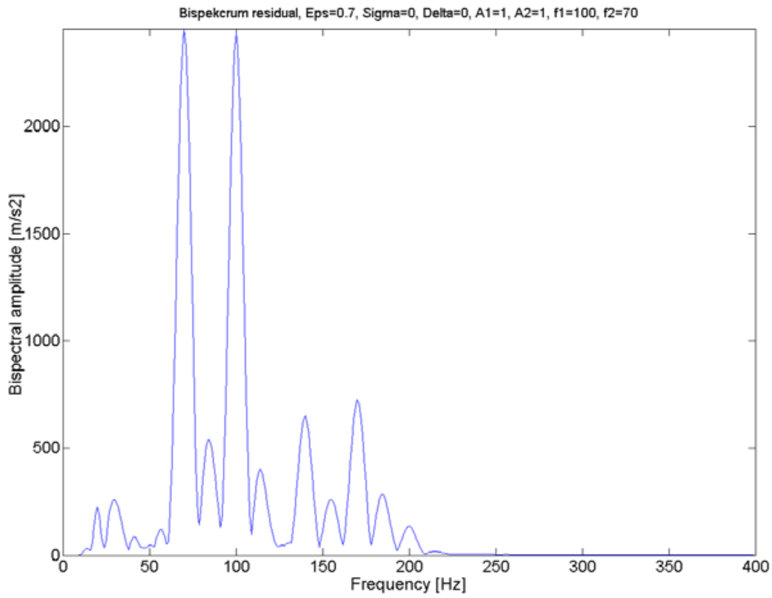

Fig. 8. Bispectral residual of a signal with mutual modulation and squared non-linearity

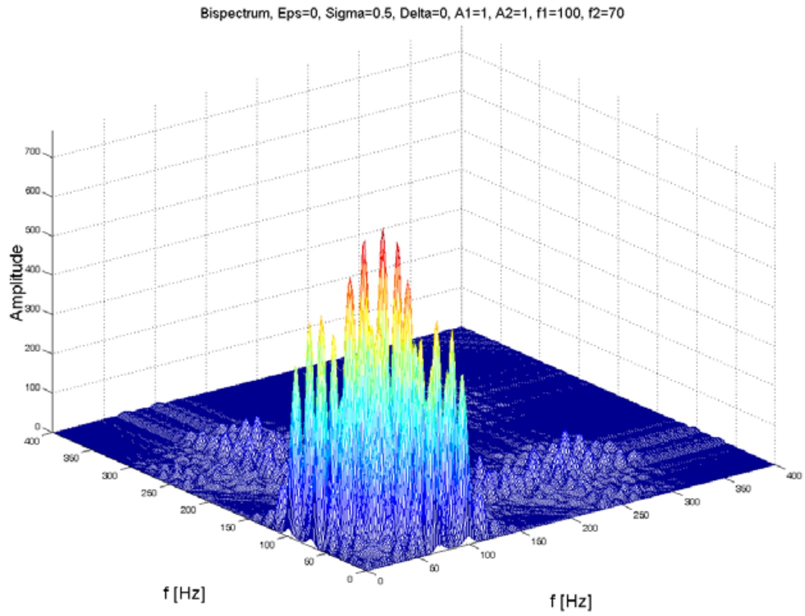

Fig. 9. Bispectrum of a signal with mutual modulation and non-linearity of the third order

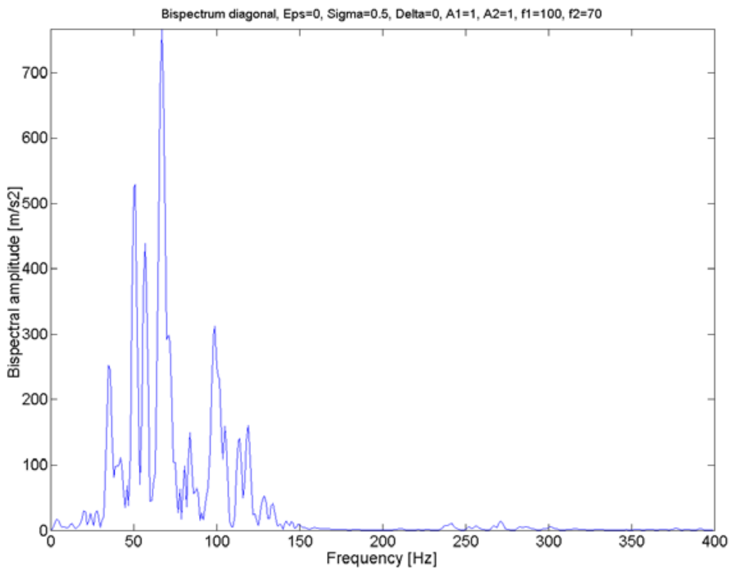

Fig. 10. Bispectrum diagonal of a signal with mutual modulation and non-linearity of the third order

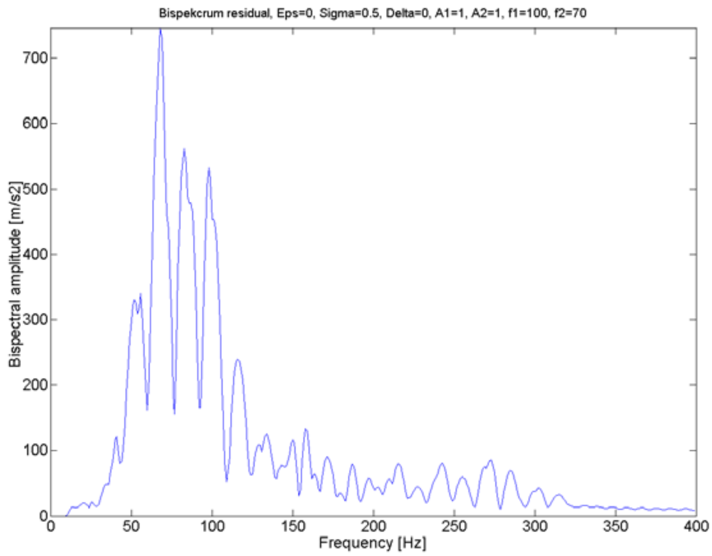

Fig. 11. Bispectral residual of a signal with mutual modulation and non-linearity of the third order

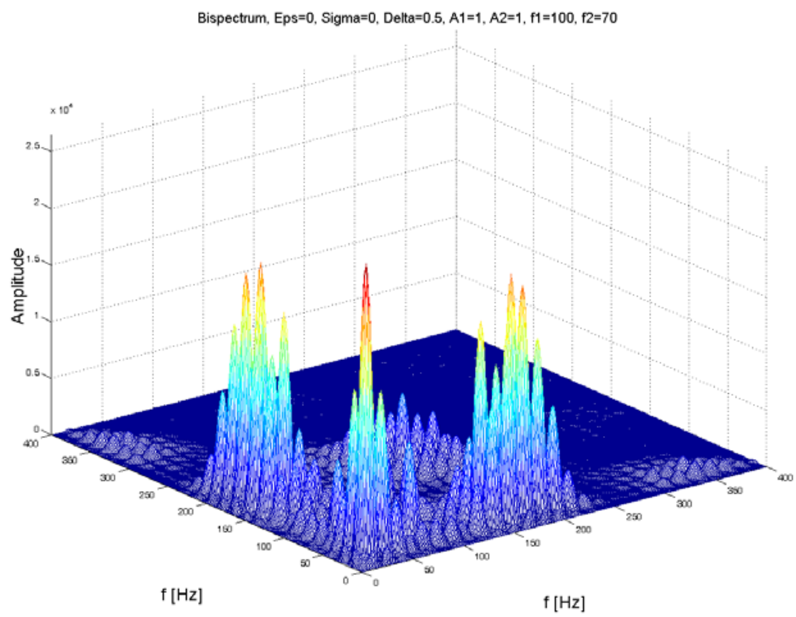

Fig. 12. Bispectrum of a signal with mutual modulation and non-linearity of the fourth order

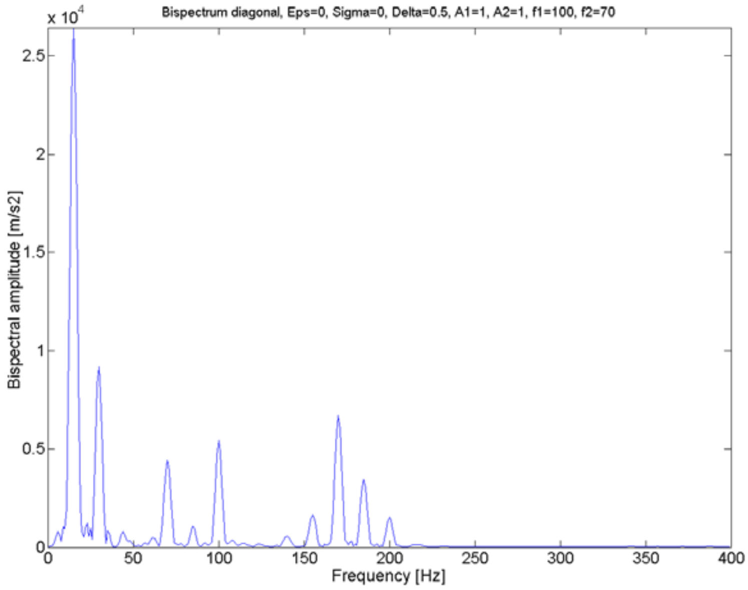

Fig. 13. Bispectrum diagonal of a signal with mutual modulation and non-linearity of the fourth order

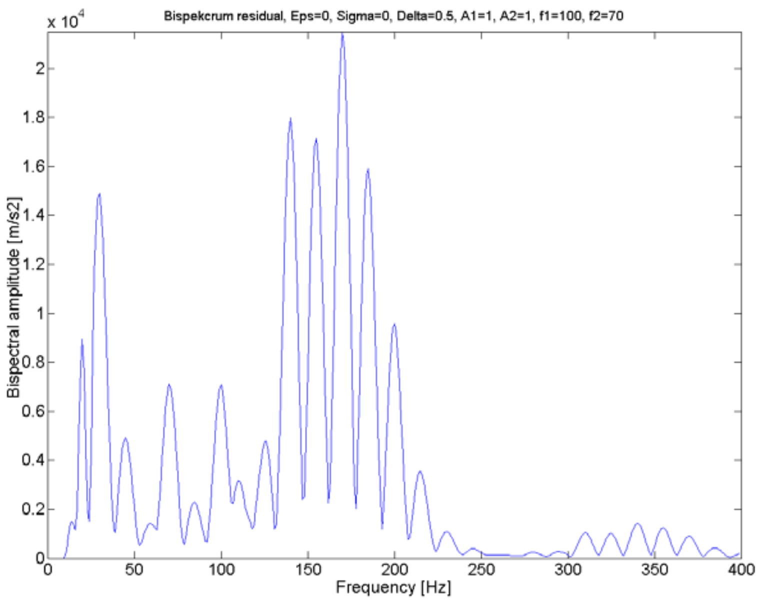

Fig. 14. Bispectral residual of a signal with mutual modulation and non-linearity of the fourth order

Let us note that the phenomena of inter-modulation and mutual modulation enable, on the one hand, the explanation of the mechanism of emergence of additional components in the spectrum, and on the other they constitute interesting basis for diagnostic inference. Above all thanks to the analysis of the relations between the emerging defect and the developing modulation effect we can observe the growing role of mutual modulation in shaping the frequency structure of a vibroacoustic signal.

For comparison, Figures 4 and 5 present the changes of the frequency structure of a single signal defined by formulas (15) and (16), that has been subjected to non-linear influence in accordance with the relationship (17). What draws one's attention here is the definitely simpler frequency structure of thus created signal and hence the significance of the mutual modulation phenomenon in shaping the structure of the disturbed signal.

The paper presents the issue of relation between the development of defects and the modulation phenomena as well as the occurrence of non-linear effects. We have proven that the complex structure of the diagnosed object can cause the effect of inter-modulation, and in the case of occurrence of defects it can also cause the effect of mutual modulation, thus leading to substantial complication of the signal's structure if it is accompanied by the phenomenon of non-linear effects. Due to this, the diagnostic inference based on the vibroacoustic signal's spectrum is becoming much more difficult. Assuming that numerous components are related to each other due to phase coupling, one should expect that significantly better results of diagnostic inference will be brought by applying multi-dimensional spectra, especially the bi-spectrum.

The relevant results of analyses are presented in Figures 6÷14. What is worth noting is the extension of the frequency band as the degree of non-linearity increases. Also the structure of bands which are characteristic for a square phase coupling can be a distinctive feature for a given type of non-linearity.

Particularly interesting results have been obtained for a diagonal bispectral measure (Fig. 7, 10 and 13) and for measure created from vector of maximum values of triangular matrix separated from bispectrum matrix by removing main diagonal of this matrix – residua bispectrum [6] (Fig. 8, 11 and 14). Fig. 6, 9 and 12 presents the results of bispectral analysis until occurring the intermodulation phenomena.

4. Conclusions

Construction and assembly imprecision as well as material-related faults influence the overall technical condition of an object and can manifest themselves in a certain specific way in the signals associated with the functioning of the machine. Most often this is associated with the need for examination of the changes occurring in the multi-dimensional description of the signal. Thus, while bearing in mind the possibility of detecting structural, assembly and material-related imperfection in the signal generated by a multi-stage toothed gear, it is necessary to develop a relevant simulation model enabling detection of diagnostic information.

One of the methods is to analyze physical phenomena on the example of diagnostic dynamic models. For example, to learn the influence of respective stages of a toothed gear, to identify potential coupling, we need to extend the models of a single-stage toothed gear to multiple-stage toothed gears. Accordingly a two-stage toothed gear is a dynamic object that is much more complex than the presented model of a single-stage toothed gear. This results not only from the bigger number of components (two pairs of toothed wheels instead of one) but also from the dynamic coupling between the elements that do not exist in a single-stage toothed gear.

Scentific research project financed from the scientific research for years 2009-2011.

References

1. Radkowski S.: Vibroacoustic Diagnostics of Low-energy Failures, ITE Radkom, 2002 (in Polish).
2. Mendel J.M.: Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications. Proc. of the IEEE, Vol.79 No 3, 1991, pp. 278-305.
3. Najar J.: Continuous damage of in: Continuum damage mechanics theory and applications. Editors: Krajcinovic D., Lemaitre J. International Centre for Mechanical Sciences Courses and Lectures, 295, 1991, pp. 234-293.
4. Chudzikiewicz A., Radkowski S.: Failure – oriented procedure of vibroacoustic signal analysis. Proceedings of the Third European Workshop Structural Health Monitoring 2006. Juli 5-7 2006, pp. 1300-1307.
5. Meyer R., Shensa M., Eschenbach R.: Cross Modulation and Intermodulation in Amplifiers at High Frequences. IEEE Journal of Solid-State, 1972.
6. Jasiński M., Radkowski S.: Use of bispectral-based fault detection method in the vibroacoustic diagnosis of the gearbox. Proceedings of the 4rd World Congress on Engineering Asset Management (WCEAM 2009), 27-30 September, 2009, Athens, Greece, pp. 32 (CD-ROM).