INVESTMENTS IN EDUCATION DEVELOPMENT

VIBRATION DIAGNOSTICS

Autorské dílo v rámci projektu Rozvoj jazykových kompetencí pracovníků VŠB-TUO: InterDV, registrační číslo projektu: CZ.1.07/2.2.00/15.0132

OSTRAVA 2012

ALENA BILOŠOVÁ
JAN BILOŠ

CONTENTS

LIST OF SYMBOLS

a	acceleration [$m \cdot s^{-2}$]
a_n, b_n	Fourier coefficients
c_n	Fourier coefficient (amplitude)
δ	decay constant [s^{-1}]
Δf	frequency resolution [Hz]
f	frequency [Hz]
f_s	sampling frequency [Hz]
f_{max}	Nyquist frequency [Hz]
$f(t)$	excitation force as a function of time [N]
F	excitation force amplitude [N]
g	acceleration of gravity ($\cong 10\ m \cdot s^{-2}$)
φ	phase shift
φ_F	initial excitation force phase shift
ϕ_n	Fourier coefficient (phase)
k	stiffness [$kg \cdot s^{-2}$]
m	mass [kg]
N	number of time samples
t	time [s]
T	period, length of the time record [s]
v	velocity [m/s], [mm/s]
$x(t)$	displacement as a function of time [m], [mm], [µm]
X	displacement amplitude [m], [mm], [µm]
x_a	amplitude (of an arbitrary quantity)
x_{RMS}	root mean square value (of an arbitrary quantity)
x_{mean}	mean value (of an arbitrary quantity)
x_{p-p}	peak-to-peak value (of an arbitrary quantity)
ω	angular excitation frequency [s^{-1}] (= [rad/s])
Ω	natural angular frequency of vibration [s^{-1}]
ζ	damping ratio [-]

LIST OF ABBREVIATIONS

A/D	analog/digital
CW	clockwise
CCW	counter clockwise
CPM	cycles per minute
ČSN	Czech standard
FFT	Fast Fourier Transform
MIMOSA	Machinery Information Management Open Systems Aliance
ISO	International Standard Organization
RMS	Root Mean Square

ACKNOWLEDGEMENT

This text was created with financial support of the European Social Fund in the scope of the project No CZ.1.07/2.2.00/15.0132. My preparations for this text were also supported by funding the official foreign business travel to the International Conference on Noise and Vibration Engineering that was held from 17th to 19th September 2012 in Leuven, Belgium. I had the opportunity to discuss technical terms used in this text with the leading specialists in vibration diagnostics from all over the world and I believe that this contributed significantly to the quality of the text.

FOREWORD

Dear students, the purpose of this textbook is to give you an insight into the area of measuring vibrations and the use of measuring vibrations in vibration diagnostics. Vibration diagnostics is one of the non-destructive methods used for condition monitoring of machines in operation. All the machines while operating vibrate more or less, and with most of them the vibrations are unwanted and the effort is to minimize them. Only with some types of machines, vibrations are directly a working principle of the machine and are caused deliberately (e.g. vibrating screeners). Though, this group of machines is not of interest to vibration diagnostics.

Diagnostic work can be thought of by analogy with activities of a practising physician who during preventive inspection detects and evaluates one's medical condition. Basically, three situations can occur: You will learn that 1) you are healthy and you can live as before, 2) you have high blood pressure and you should start taking the medication for its reduction and/or change your lifestyle, or 3) your condition requires hospitalization and a more detailed examination and/or a surgery. Machines are at exactly the same situation. Based on a diagnostician’s assessment they can either continue in operation, or a tiny intervention is necessary, or they need to be shut down and repaired thoroughly. Purpose of all this is, in case of both humans and machines, to save the cost of repair or to prevent a disaster and its associated costs.

As the name *vibration diagnostics* suggests, machine condition is diagnosed on the base of an analysis of vibration. Successful application of vibration diagnosis requires in practice staff with considerable degree of knowledge and experience. Routine work in data collection may be carried out by trained personnel without academic qualifications, but data processing and assessment of the state of a machine is a task for an engineer who has knowledge in various areas (design of machines, dynamics, mathematics, signal processing, etc.) and who is able to use this knowledge in context. A graduate in Applied Mechanics specialization is an ideal candidate for becoming a skilled vibration diagnostician after several years of practice.

This text is almost your first encounter with the experimental mechanics. We believe that we will convince you that it is a beautiful and promising area which should become an integral part of your engineering practice and mastering of which will contribute to your becoming a full member of the team of experts addressing complex technical problems.

1 VIBRATION DIAGNOSTICS – PRELIMINARY CONSIDERATIONS

Each machine, if it has to work reliably throughout its planned life, must be maintained. For all large and expensive equipment, to which the vibration diagnostics mainly applies, operational life is an essential and often neglected part of the life of the machine. The machine life can be divided into the following stages. Duration time of individual stages is given here for huge machinery such as turbo-generators:

- Period of creation
  - design: duration depends on the designed part; usually 1 to 3 years
  - production: usually half a year to 1 year
  - assembling: several months
  - setting in operation: 1 to 2 months
- operation: 25 years or even more (100 000 to 200 000 operating hours)

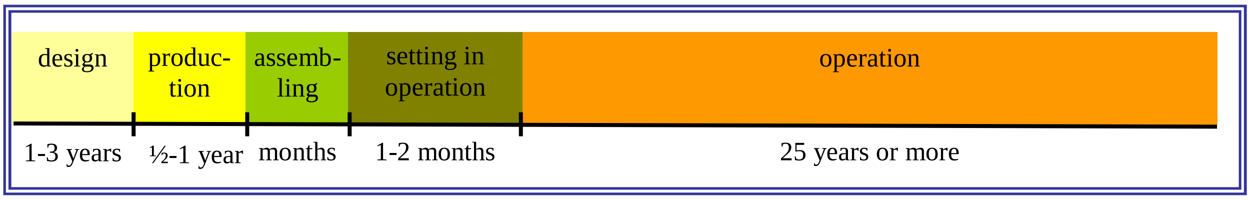


Fig. 1.1 – Scheme of Machine Life

Simplified graphical representation of the total machine life in Fig. 1.1 aims to highlight the large discrepancy between the duration of a machine’s creation, when the development, design, manufacturing and assembling involved a large group of specialists from various disciplines (computational, engineers, technologists, assemblers, test technicians) and much longer operating time during which the machine works flawlessly, if possible, without faults and with permanently great efficiency. Appropriate maintenance during itsoperation is just as important for a reliable machine’s operation as proper design, manufacturing and assembling.

In Fig. 1.2 there is a view on the assembly of a complex machinery unit (turbo-generator) for a power station. All essential parts are manufactured with certain tolerances or even with allowances, so preliminary assembly is done in the manufacturer’s plant in order to ensure that the entire device can be mechanically assembled. Whenever possible, the device that is factory-assembled is not disassembled any more. For larger systems, it sometimes applies only to some parts, in this case to a high pressure turbine part (Fig. 1.2 left) which is transported assembled to the power station. To decide whether this is possible, it is necessary to consider the possibility of transport (dimensions) and the way of transport. At a construction site, crane capacity and dimensions of access openings to the building are determinative (sometimes they must be, at least temporarily, increased).

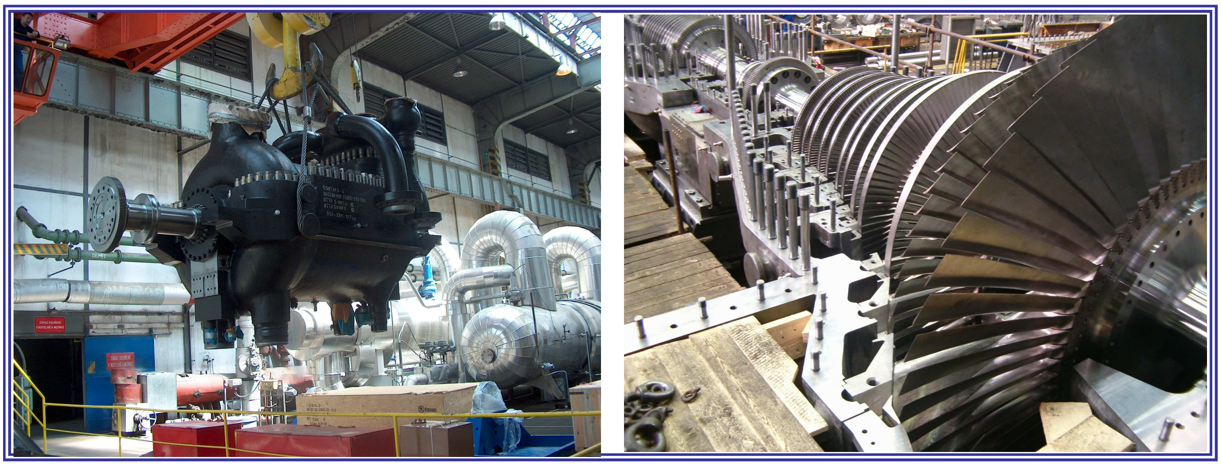


Fig. 1.2 – Preliminary Assembly of a Complex Machinery

1.1 Role of Maintenance

The role of maintenance is not to repair damaged equipment, but to prevent its damage. Moreover, we want the machines to work efficiently, reliably and safely. Goal of the maintenance can be expressed through three interrelated requirements:

1. Achieve maximum productivity:
   - Ensure continuous and satisfactory operation of the machine throughout its proposed lifetime – or even longer.
   - Achieve higher machine utilization with minimal downtimes for maintenance and repairs.
   - Continually improve the production process.
2. Optimize machine performance – Smooth and efficiently running machines cost less and produce higher quality products.
3. Ensure operation safety.

Each of you may imagine a car example – when you neglect maintenance, your car will not only be unreliable, but can also be dangerous.

1.2 Maintenance Types

Equipment maintenance is essential for long-term trouble-free operation. In the course of technological development, several types of maintenance have been established, the application of which depends on a number of circumstances that must be considered. Basic maintenance tasks are listed in the preceding paragraph. In considering them, however, costs should always be considered together with safety. Therefore, small and backed up equipment is still used in *maintenance-free way* – i.e. *operation to failure*. Examples of this type of maintenance are household appliances (we do not perform regular inspection of a vacuum cleaner or microwave); in industry these may be small (and backed up) pumps, etc. This type of maintenance is called **reactive maintenance**.

For more expensive equipment with more costly operation, the method with *periodic* maintenance inspections or repairs has been established, which is called **preventive maintenance**. Examples of this type of maintenance may be cars that have a service book and certain service tasks are distance-based or are performed in certain periods of time. Number of large industrial facilities is treated in this way as well. The aim is to prevent failure of the machine. Time to repair is determined by the rate of failures of similar equipment - failure rate reflects the so called *mean time to failure*. This constitutes a certain weakness of this method because it is difficult to estimate time between repairs - some devices have a failure before the planned repair, some are revised or repaired “uselessly” (they were all right). For example, routine repair of the 70 MW turbo-generator, costing about CZK 500 million, is carried out every 4 years, the overhaul every 8 years. Costs of the routine repair are around CZK 4 million, including spare parts, overhaul costs are approximately CZK 15-17 million + costs of spares range from CZK 4 to 8 million.

Largely because of an effort to prevent unexpected failures, but also because of an effort to optimize maintenance costs, which in medium-sized enterprises account for about 1/3 of all costs (hundreds of millions of CZK per year), two other ways of maintenance are developed:

Predictive maintenance – a machine is repaired when its condition requires a repair rather than at predetermined intervals. Of course, it is necessary to know this condition, and therefore watch the machine in operation, i.e. to perform monitoring and diagnostics. This approach helps us to avoid unplanned shutdowns and failures. The key idea is *the right information at the right time*. If we know which part of the equipment requires replacement or repair, we can order spare parts, arrange it for staff, etc., and perform shutdown at the appropriate time. Such a planned shutdown is shorter and less costly than a shutdown forced by a failure of equipment or even an accident. An increase of equipment lifetime, increased safety, fewer accidents with negative consequences for the environment, optimized management of spare parts, etc. are other advantages of predictive maintenance.

Proactive maintenance - In addition to the previous type of maintenance this one also includes addressing the root causes of the aggravated condition. Corrective actions do not focus on current symptoms of the fault (e.g. damaged bearing), but the key idea is to identify and address the root cause of the fault (e.g. damage of the bearing has arisen due to a bad alignment of the machine).

There are also other modern maintenance methods such as RCM - Reliability Centred Maintenance, which is used in aviation, and others.

When predictive and proactive maintenance is applied, it is necessary to determine the current condition of the machine. Maintenance process can be divided into five stages:

1. **Determining the initial condition -** A thorough measurement of the machine is performed at the time when it is in good state, which provides the basic reference values for subsequent comparison.
2. **Monitoring -** On the machine, points are defined at which vibrations are measured at regular time intervals. Usually the overall value of vibration is measured. This activity can be carried out by a trained worker without diagnostic knowledge.

3. **Detection -** Data obtained through monitoring are simply quantitatively evaluated. For each measured quantity, alarm limits are set. Exceeding the programmed alarm limit means warning about a problem.

4. **Analysis (diagnostics itself) -** After detecting the problem, detailed measurements and analyses (evaluation of the trend, FFT analysis, phase analysis, etc.) are carried out allowing a clearer view of the problem and its underlying cause.

5. **Recommendation -** Once the basic cause of the problem has been detected, economically acceptable corrective actions can be recommended and implemented.

1.3 Diagnostics

The term *diagnostics* is usually used for monitoring and evaluation condition of a machine during operation (i.e., points 2, 3 and 4 stated above). This paper deals with vibration diagnostics, where detection of the machine condition is based on its vibration. In practice, it is advisable or even necessary to use other parameters for monitoring as well. Most of the procedures are described by international standards ISO (see list of references). Types of diagnostics according to the type of parameters analyzed are:

- **Operation diagnostics -** All available measured operating parameters, which allow assessment of the machine condition in operation, are used. For very expensive equipment, on-line systems are used with large databases and possibly with analysis software. For less important machinery, parameters are periodically recorded or some tests during operation are performed to verify proper operation. Standard ISO 17359 - Condition monitoring and diagnostics of machines - General guidelines deals with this issue. Table 1 is taken from this standard and it shows how various operational parameters are associated with various machine faults. In the standard, there are several such tables for different types of machines. For rotating machines, the majority of defects manifest themselves in change of amount and spectral content of vibrations.

- **Tribo-diagnostics** (analysis of lubricants) - It fulfils two main tasks:

  - <u>Monitoring the condition of the lubricant</u> - A lubricant degradation can occur for various reasons (oxidation, penetrating of water or other substances, etc.).

  - <u>Analysis of impurities and wear particles</u> (ferrography) – On the base of the material and shape of particles present in the lubricant, an assessment about the place where the machine is damaged is carried out.

- **Thermo-diagnostics** (measurements of temperature, thermal imaging) - Using local or surface temperature measurements, sites with different temperature can be determined and the cause of the elevated temperature can be deduced (excessive friction, high electrical resistance, etc.). Thermo-diagnostics is widespread in inspections of electrical switch-gears, high voltage lines, hot water pipes, in the steel industry (brick lining of furnaces and chimneys), etc. Figure 1.3 is an example of applying thermo-diagnostics to detect misalignment in coupling (when the coupling is misaligned, greater loss of transmitted power occurs that is converted to heat, warming the coupling and the adjacent bearings). Picture of thermal imaging camera is usually supplemented with common photo to be clear what kind of equipment is observed.

- **Ultrasonic diagnostics -** Based on the physical fact that the dry friction generates ultrasound. It is also produced when the flow occurs - the leakages due to leaks and friction in seals, etc. In addition, electrical discharges produce ultrasound as well and therefore this method and instruments based on it are also used by specialists in the field of electrical equipment.
- **Electro-diagnostics -** Based on the analysis of electrical quantities (e.g. power supply) to detect faults of electrical machines (e.g. broken rotor bars).
- **Vibration diagnostics -** Vibration signal involves information about the cause of vibration and through its analysis using different methods, an emerging or developing fault can be detected. For rotating machines, this is usually the method that covers most possible faults (see the example of the standard in Table 1.1). Vibration diagnostics is described in more detail in ISO 13373-1: *Condition monitoring and diagnostics of machines - Vibration condition monitoring - Part 1: General procedures* and ISO 13373-2: *Condition monitoring and diagnostics of machines - Vibration condition monitoring - Part 2: processing, presentation and analysis of vibration data*.

Table 1.1 - Example of Operational Parameters Monitoring (ISO 17359)

Machine types: **pumps**	**Symptom or parameter change**									
Examples of faults	Fluid leakage	Length measurement	Power	Pressure or vacuum	Speed	Vibration	Temperature	Coast down time	Oil debris	Oil leakage
Damaged impeller		•	•	•	•	•	•	•	•	
Damaged seals	•	•		•	•	•				
Eccentric impeller			•	•	•	•	•	•		
Bearing damage		•	•		•	•	•	•	•	•
Bearing wear		•				•	•	•	•	
Mounting fault						•				
Unbalance						•				
Misalignment		•				•				
• Indicates symptom may occur or parameter may change if fault occurs.										

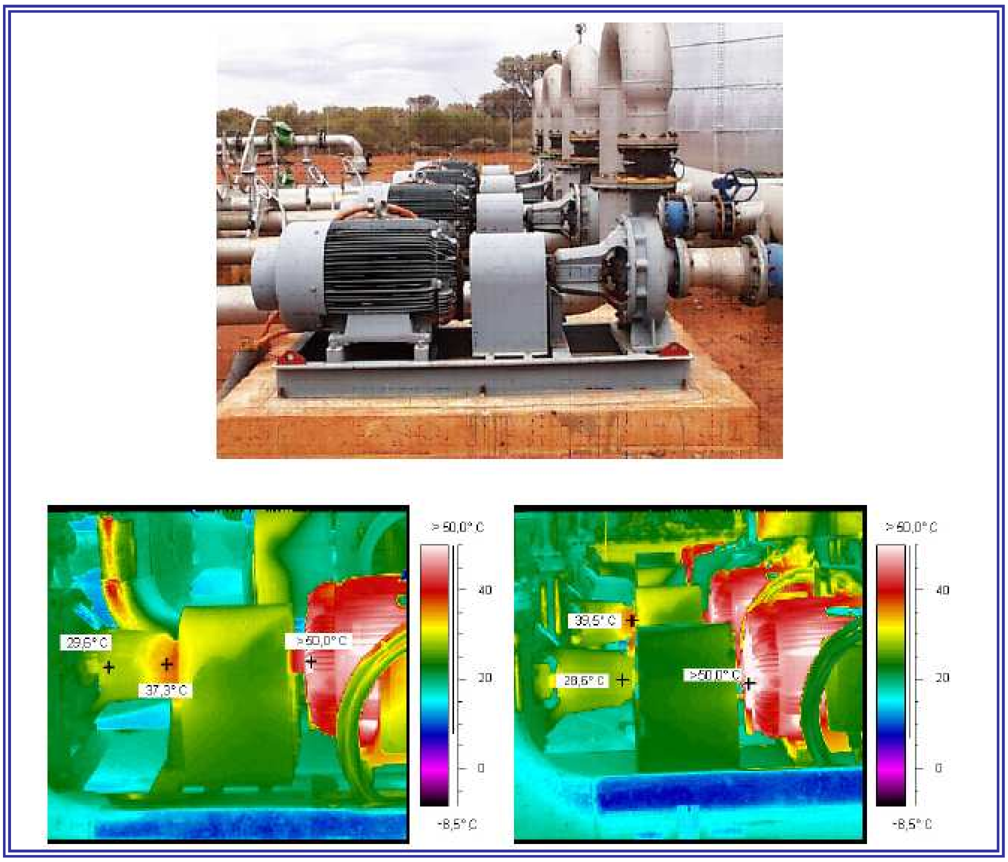


Fig. 1.3 - Example of Thermal Imaging Measurement

In the Czech Republic, there is a Technical Diagnostician Association which brings together diagnostic specialists in their respective sections by the type of diagnostics. It also performs certification of specialists (see more at www.atdcr.cz).

1.4 Excitation Force and Vibration Response

Basic problem of the application of each type of diagnosis is the fact that we analyze only the response to the acting causes that are essential to establish the method of repair. In the case of vibration diagnostics, this response is represented by vibrations, the character of which depends on the applied force. Common types of excitation force are:

- periodical
- impulse
- random

Periodical excitation force

The simplest case of periodic force is a harmonic force. In engineering practice, harmonic force is very rare, but most of the real forces occurring in rotating machinery can be expressed as a sum of harmonic forces (see Chap. 1.5 and Fig. 1.10). Therefore, it is possible to describe the properties of periodic force and its influence on the vibration response using harmonic force and response. If a harmonic force

$$f(t) = F \cdot \sin(\omega t + \varphi_F) \qquad (1.1)$$

acts on a flexibly supported body,

where F ... excitation force amplitude [N]
ω ... excitation force angular frequency [rad/s]
t ... time [s]
φ_F ... excitation force initial phase shift,

the steady movement of the body is also harmonic with the same angular frequency $\boldsymbol{\omega}$, but generally with different amplitude (see Fig. 1.4). This vibration is called *forced vibration*. Displacement of such a vibration can be expressed as:

$$x(t) = X \cdot \sin(\omega t + \varphi_F - \varphi) \tag{1.2}$$

where X ... amplitude of forced vibration
φ ... phase shift - lag between the displacement and the acting force

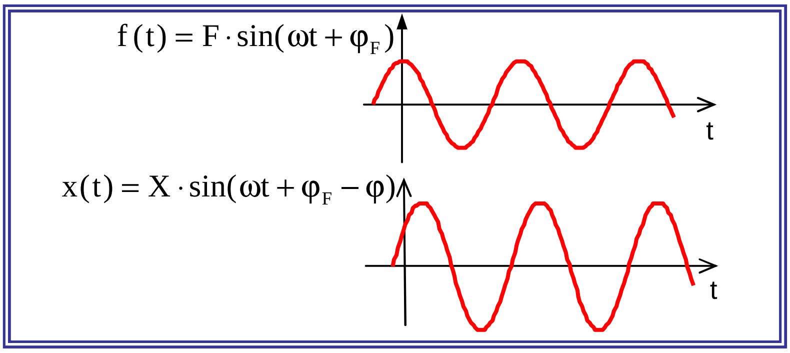


Fig. 1.4 - Forced Vibration Caused by Harmonic Excitation Force

The type of vibration when the vibration excitation force and response are periodic occurs for instance when there is rotor unbalance or coupling misalignment.

Impulse Excitation Force

When an impulse force is acting on the body, it diverts the body from the equilibrium position which causes subsequent free vibration on one or more of its natural frequencies (see Fig. 1.5). A common example may be hitting a glass (solid glass sounds different than a cracked one), ringing a bell, etc. In technical practice, we use intentional impulse excitation performing "bump test" or a modal test. The unintentional impact excitation is associated with defects in rolling bearings (see Chap. 4.5.3).

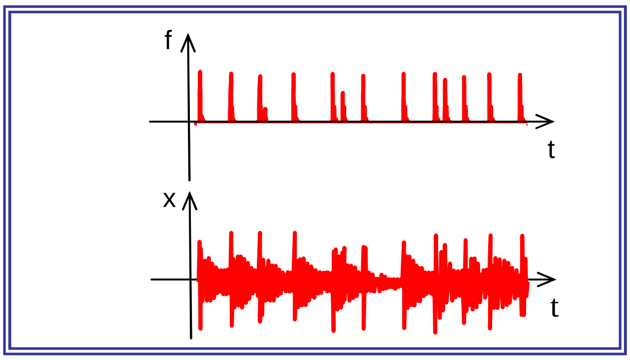


Fig. 1.5 - Free Vibration Caused by Impulse Excitation Force

Excitation force of random waveform

When random force acts on a body, the response is also random (see Fig. 1.6). Moreover, similarly to impulse excitation, natural frequencies can be excited (any abrupt change in force would excite the free vibration on natural frequencies). It should be realized that the random excitation is always present, mostly as noise only, but occasionally it should be considered even in the standard vibration diagnostics, e.g. when unwanted turbulence flow occurs.

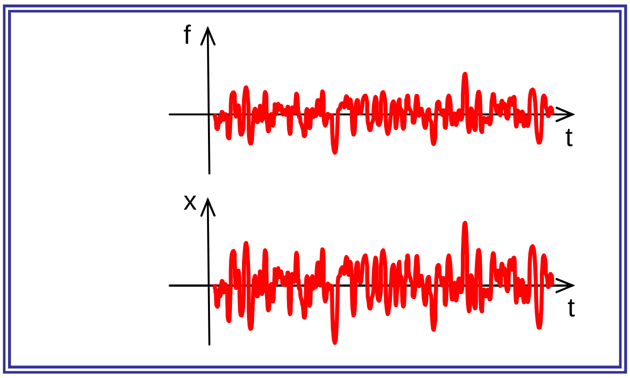


Fig. 1.6 - Vibration Excited by Force of Random Waveform

Self-excited Vibration

Self-excited vibration is potentially a very destructive phenomenon when aerodynamic or hydrodynamic forces acting on the object excite vibration of the object on one of its natural frequencies. This phenomenon is called flutter and, e.g. in aviation, it should be unconditionally avoided. In a non-destructive form, this phenomenon is quite common - it causes for instance vibrations of blinds caused by a draft, laundry flapping on a clothesline, etc.

Flutter can occur in any object within a strong fluid flow, under the conditions that a positive feedback occurs between the structure's natural vibration and the aerodynamic forces. That is, the vibrational movement of the object increases an aerodynamic load, which in turn drives the object to move further. If the energy input by the aerodynamic excitation in a cycle is larger than that dissipated by the damping in the system, the amplitude of vibration will increase, resulting in self-exciting oscillation. The amplitude can thus build up and is only limited when the energy dissipated by aerodynamic and mechanical damping matches the energy input, which can result in large amplitude vibration and potentially lead to rapid failure. Mathematically it can be described similarly as a free damped oscillation, where the damping is negative, see Fig. 1.7). Because of this, structures exposed to aerodynamic forces - aircraft wings, turbine and compressor blades, but also chimneys and bridges - are designed carefully within known parameters to avoid the flutter. In complex structures where both the aerodynamics and mechanical properties of the structure are not fully understood, flutter can only be assessed through detailed testing (e.g. a new aircraft without ground vibration test is not allowed for operation). An example of a structure that was destroyed by self-excited vibration is Tacoma Narrows Bridge at Fig. 1.8.

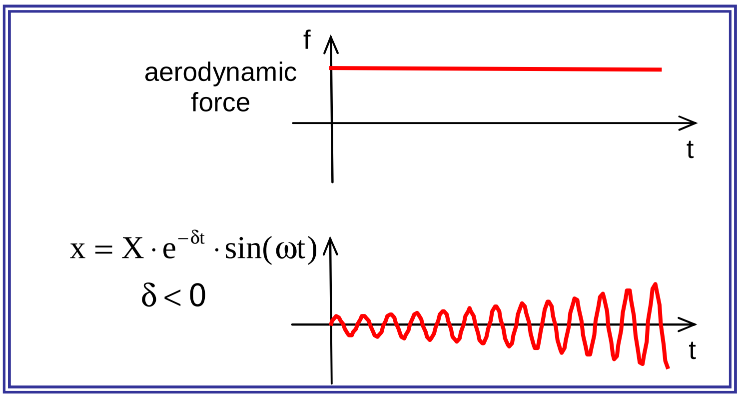


Fig. 1.8 - Self-excited Vibration

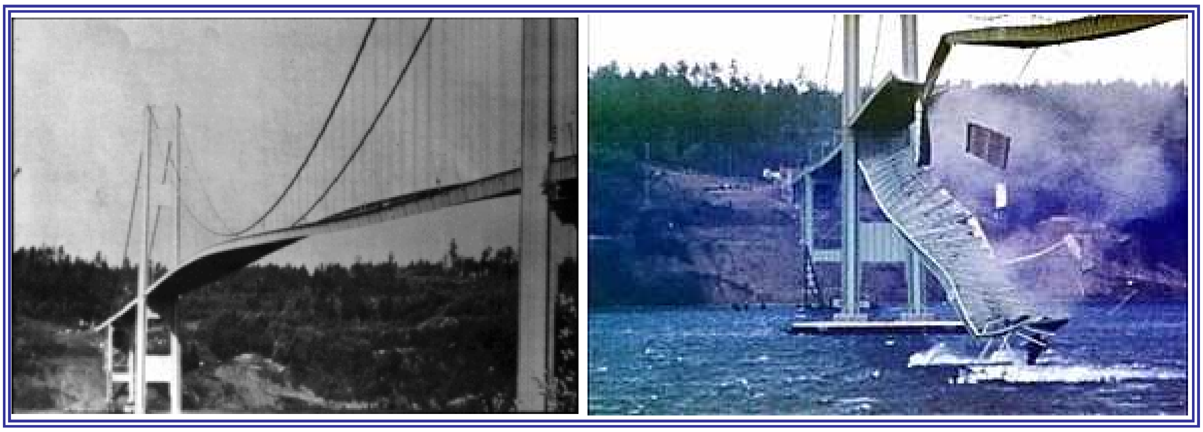


Fig. 1.8 - Destruction of Tacoma Narrows Bridge

1.5 Basic Quantities Describing the Oscillatory Movement

Mass of **m** supported on a spring of stiffness **k**, after being displaced from its equilibrium position, performs harmonic oscillating motion. If damping is neglected, the mass oscillates with natural frequency $\Omega = \sqrt{k/m}$ and the course of displacement is a sine wave with amplitude of $\mathbf{x_a}$ (see Fig. 1.9), thus:

$$x(t) = x_a \cdot \sin(\Omega t - \varphi) \qquad (1.3)$$

where x_a ... amplitude of harmonic oscillation [m]
Ω ... angular natural frequency [rad/s]
φ ... initial phase shift (is determined by the initial displacement)

In technical practice, frequency **f** expressed in hertz (i.e. in number of complete cycles per second) is used more often than angular frequency $\mathbf{\Omega}$ (or $\boldsymbol{\omega}$) expressed in radians per second:

$$f = \frac{\omega}{2\pi} \quad [\text{Hz}] \qquad (1.4)$$

Reciprocal value of a frequency **f** is a period **T**:

$$T = \frac{1}{f} = \frac{2\pi}{\omega} \quad [\text{s}] \qquad (1.5)$$

Other characteristics, rather than amplitude, are often used to describe the harmonic signal (see Fig. 1.9), namely (Note: **x** here can mean any quantity, not just the displacement):

peak value (= amplitude for harmonic signal) x_a

rms (Root Mean Square) *value* = 0,707 × amplitude $x_{RMS} = 0{,}707 \cdot x_a$ *(1.6)*

mean value = 0,637 × amplitude $x_{mean} = 0{,}637 \cdot x_a$ *(1.7)*

peak-to-peak value = 2 × amplitude $x_{pk-pk} = 2 \cdot x_a$ *(1.8)*

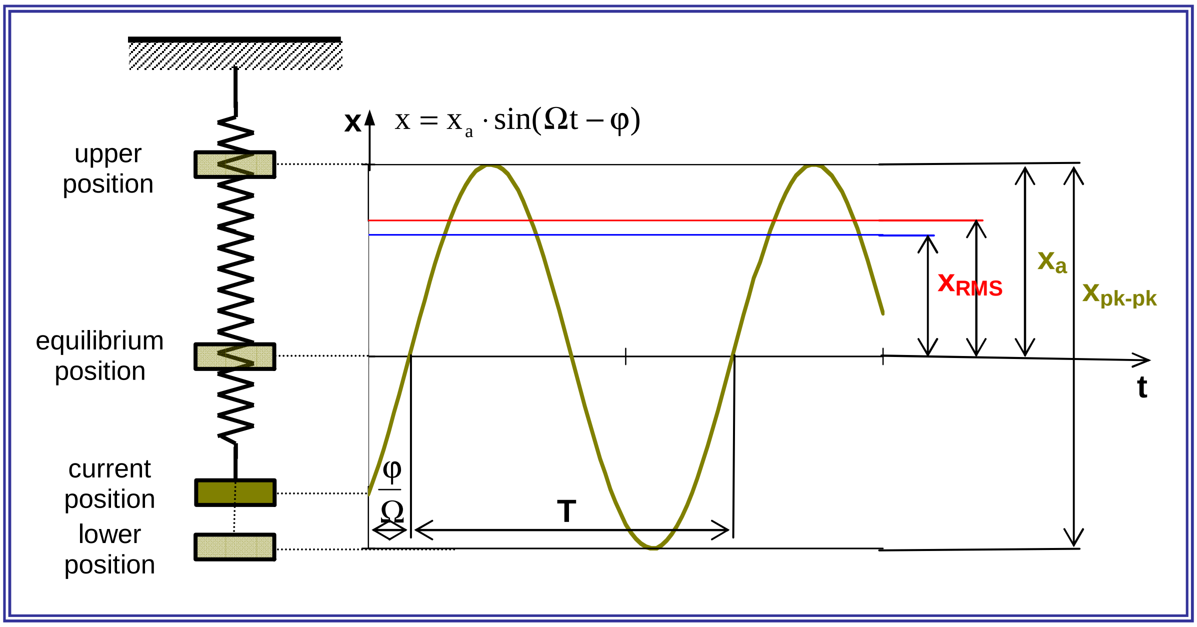


Fig. 1.9 - Quantities for Harmonic Motion Description

Similar characteristics are also used for signals that are not harmonic. For them, the concept of amplitude loses its sense, but expressions for the rms value and mean value are still valid. RMS value is often used to describe the vibration signal. It represents the average power of the measured quantity. Procedure for obtaining rms value is the following:

- a signal for a certain measurement time T is recorded (generally, it needs not to be a period)
- the signal is rectified (it may have both positive and negative values which would cancel out when simply summarized) - numerically it is done by raising to the power of two
- the values are summarized
- the sum is divided by measurement time T giving the average value
- the result is root extracted

The above described procedure can be expressed as:

$$x_{RMS} = \sqrt{\frac{1}{T} \cdot \int_0^T x^2 dt} \qquad (1.9)$$

and the name *root mean square* is obvious from this. It should be remarked that if the measurement time is not equal to the period (what usually is not the case in practical applications), repeated measurements would not get exactly the same value, even while keeping all the rules for correct measurements. It is a consequence of the fact that the rms value will be calculated of a random waveform at each measurement.

Mean value, which is much less important in technical practice, can be determined by:

$$x_{mean} = \frac{1}{T} \cdot \int_0^T |x| dt \tag{1.10}$$

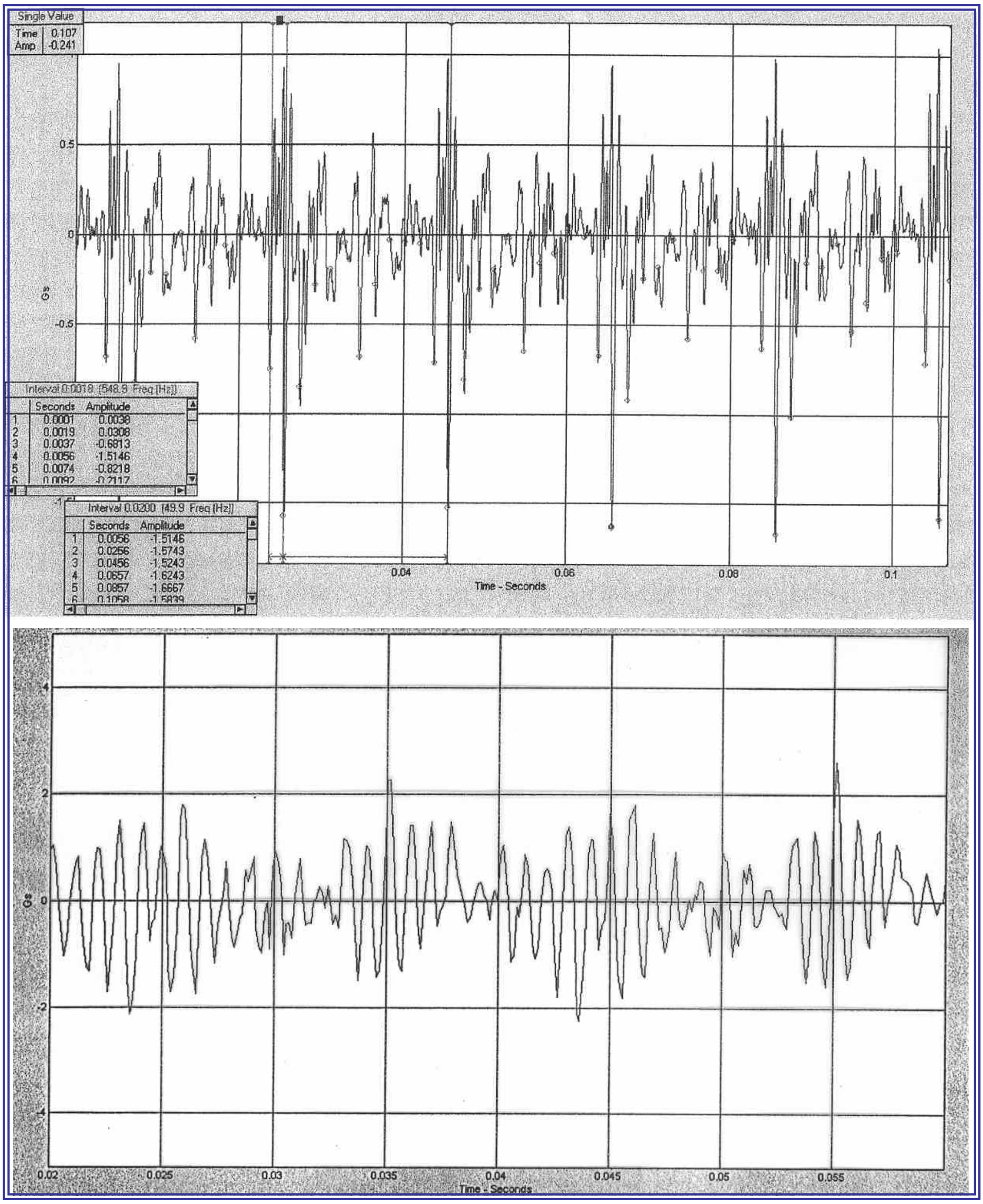


Fig. 1.10 - Practical Examples of Vibration Waveform
(Above: Acceleration Vibration Waveform of a Turbine Pedestal,
Below: Vibration Waveform Caused by Misaligned Gears)

Examples of waveform for practical applications are presented in Fig. 1.10. Above is a waveform of vibration at a turbine pedestal between turbine cases. At rotational speed 3000 rpm, the period is 0.02 sec and on the picture, marked peaks with this period can be observed. However, in between these peaks, vibrations are low. Below is a waveform of vibration at a front turbine pedestal, where a simple gear drive of the oil pump is placed. From the pictures it is obvious that an experienced diagnostician has to decide which quantity he would use for vibration assessment - whether the peak value, rms value or some other criteria. The right choice depends on the particular application and the measured quantity.

Using the ratio of peak to rms value, the prevailing shape of the signal waveform can be presumed. This ratio is called *crest factor*.

$$CF = \frac{x_a}{x_{RMS}} \tag{1.11}$$

When the CF is small (approximately up to 3.0), the prevailing character is sinusoidal; when the value is higher, the impulse character is prevailing and this is one of the methods for assessing the condition of rolling bearings.

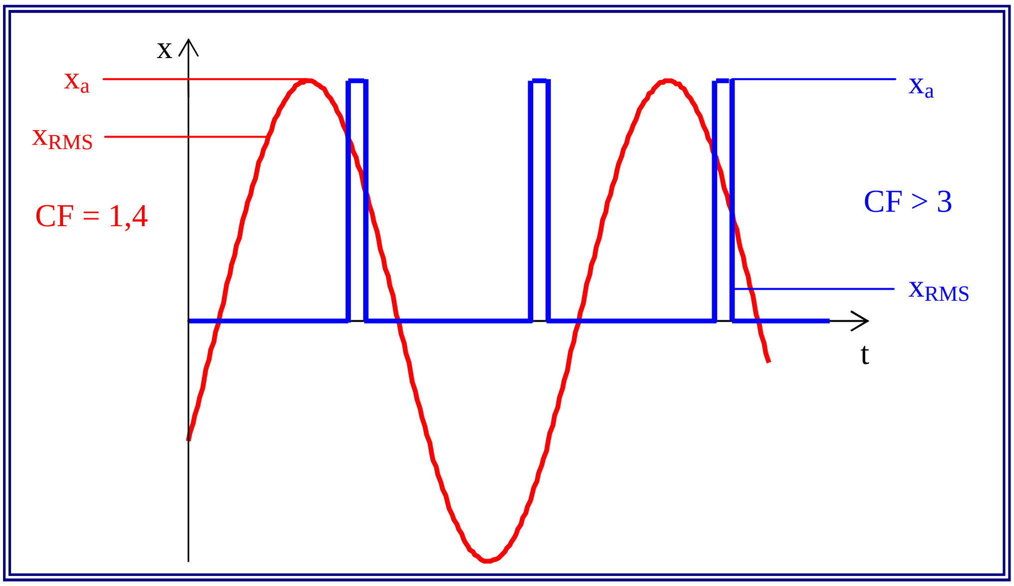


Fig. 1.11 - Peak Value, RMS Value and Crest Factor

1.6 Measured Quantities

In mechanics, movement can be described by displacement, velocity or acceleration, and these variables are linked by mathematical relationships. From this perspective, it does not matter which variable is chosen to describe the vibrational behaviour, it is just a matter of scale and time shift (phase).

displacement of vibration is usually stated in micrometers [μm]

velocity is the first derivative of displacement with respect to time
(velocity of displacement change); it is usually stated in *mm/s*

acceleration is the second derivative of displacement with respect to time
(velocity of velocity change); it is usually stated in m/s^2 or in *g*

In the example at Fig. 1.12 (where X = 1 mm, ω = 2 rad/s), waveforms of these quantities during one period are shown:

$$x(t) = X \cdot \sin(\omega t)$$

$$v(t) = \frac{dx}{dt} = X \cdot \omega \cdot \cos(\omega t)$$

$$a(t) = \frac{dv}{dt} = -X \cdot \omega^2 \cdot \sin(\omega t)$$

It can be seen that theoretically it is enough to know one of the variables, and the remaining two can be easily computed. The velocity always lags for 90º behind the deflection and the acceleration lags for further 90º behind the velocity.

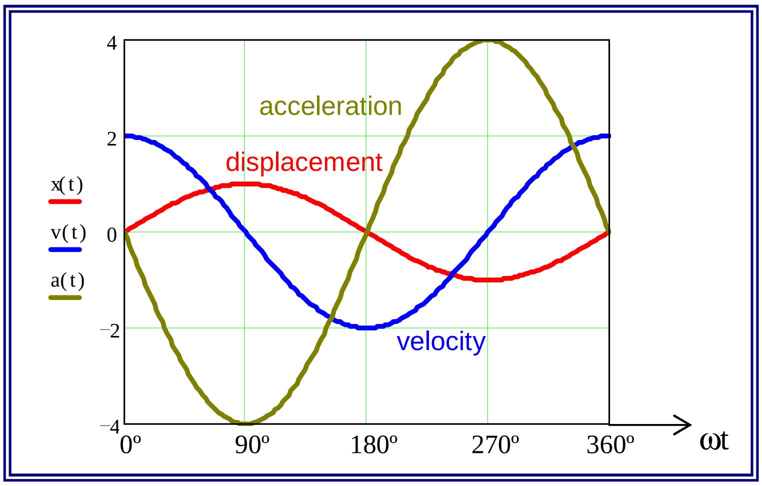


Fig. 1.12 - Relations between Displacement, Velocity and Acceleration

In contrast to the calculations, however, measurements should also take into account adverse factors that influence the measurement accuracy and therefore it is advisable to choose the measured value to give sufficient signal to noise ratio. Noise is always present in the measured data and for weak signals it means more inaccuracies (measurement errors).

Fig. 1.13 indicates why velocity is used for common measurements in the frequency range 10 Hz to 1000 Hz, acceleration is preferred for higher frequencies and displacement is preferred for lower frequencies. If a constant amount of vibration at all frequencies is considered (e.g. 7.6 mm/s, which is the common value for measurements of rotating machines), the deflection decreases with increasing vibration frequency and acceleration increases. Frequency range of interest is one of the factors that determine the type of measured value. If the measured frequency range includes high frequencies (such as gear mesh frequencies), the best choice would be to measure acceleration. Conversely, if the measurement frequency is limited to the running speed, the best choice would be measuring displacement or velocity (depending on application). When measuring the velocity of vibration, there is no need to care about the frequency (speed) at which the value was measured; when measuring the other two variables, it is necessary to indicate at what rotational speed (frequency) the value was measured. Otherwise, it is not possible to assess the condition of the machine.

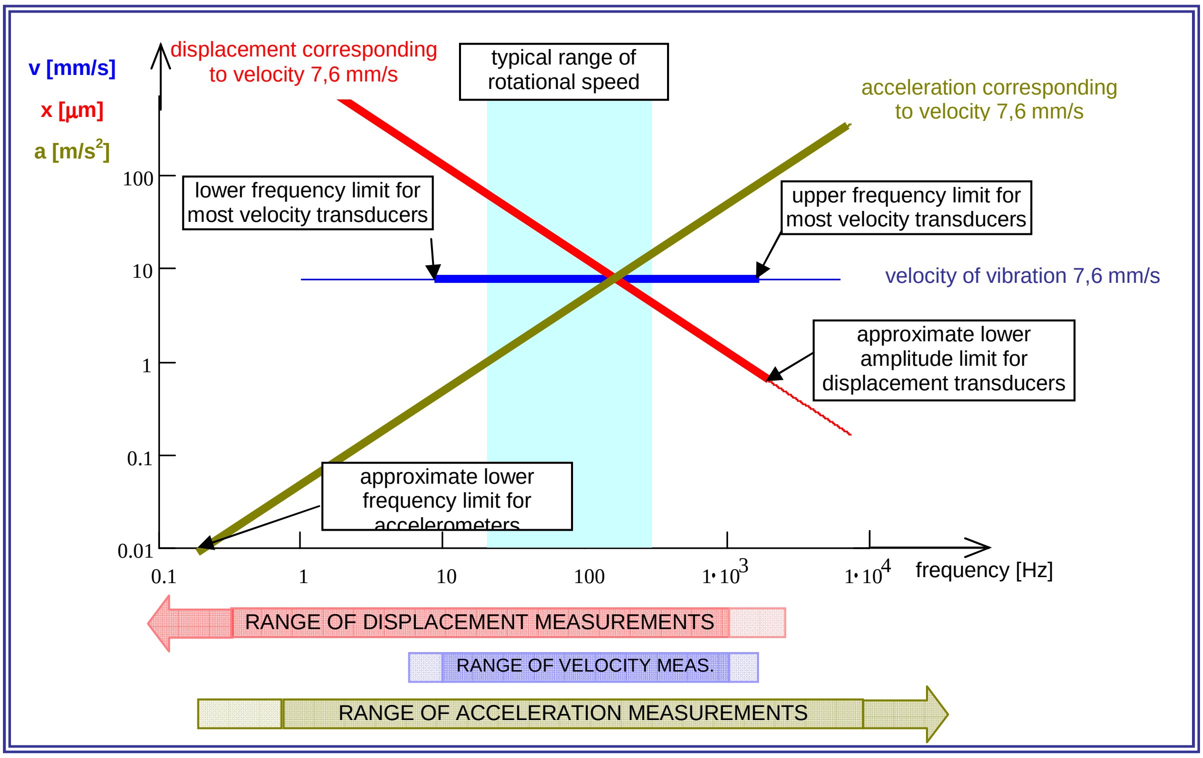


Fig. 1.13 - Measurement Limitations

2 VIBRATION MEASUREMENTS

To be able to measure vibration of a machine, some technical equipment is necessary. In practice, various tools are used, from simple instruments measuring overall vibration to multi-channel analyzers equipped with numerous features that facilitate not only the measurement itself but also the analysis of the measured data.

In this chapter, the typical scheme of vibration analyzer and various types of sensors that are used for vibration measurements will be introduced. Furthermore, analysis and evaluation of the vibration measurements will be described.

2.1 Analyzer

Basic scheme of the analyzer used for vibration measurements is in Fig. 2.1. The analogue signal from the vibration sensor passes through the input amplifier, anti-aliasing filter and A/D converter, where it is digitized and enters the data buffer. From the buffer it can be displayed either as a time waveform or can be further processed by the Fourier transform to obtain frequency spectrum. Individual functional units of the analyzer will be discussed in detail in the following chapters.

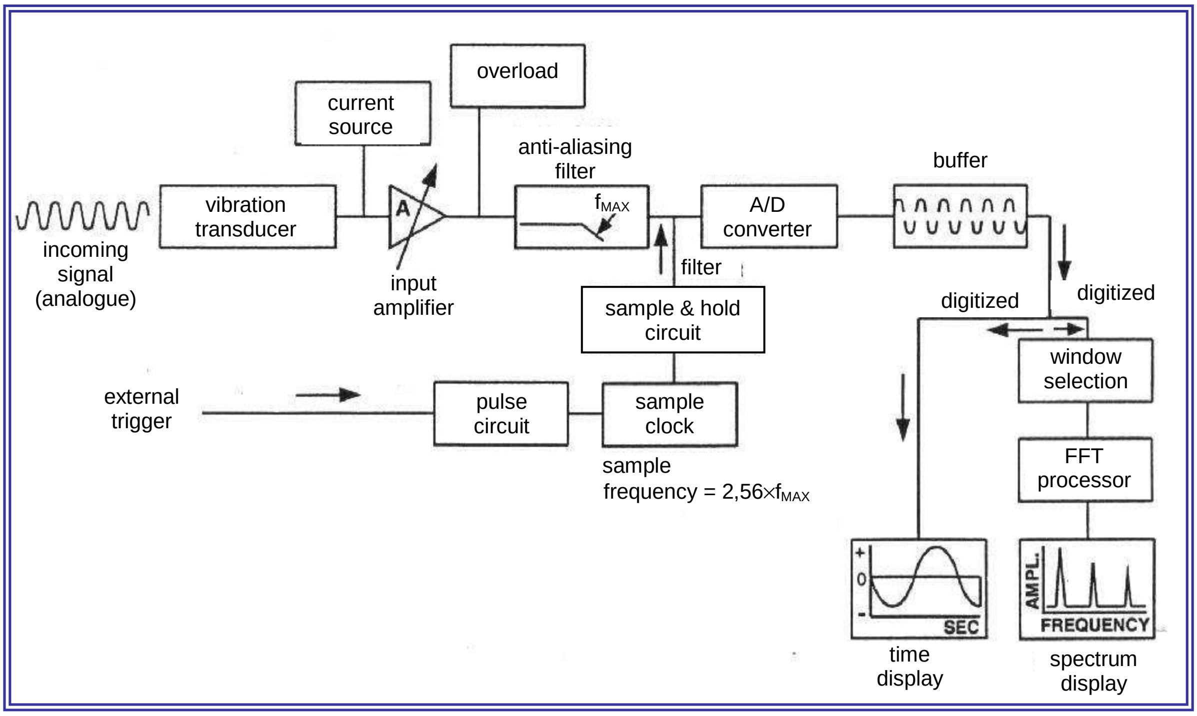


Fig. 2.1 - Vibration Analyzer Scheme

2.2 Vibration Transducers

In chapter 1.6 it was stated that any of the three quantities describing the vibratory motion can be measured. Depending on the measured quantity, sensors are divided into:

- displacement transducers (proximity probes)
- velocity probes (velometers)
- accelerometers

Usable frequency response and dynamic range differ for various types of sensors. The *dynamic range* of a sensor is the range of amplitudes of the measured quantity that can be measured by the sensor. Choosing the right type of sensor depends both on the application (for example, whether shaft vibration or vibration of machine case are measured) and on the frequency range of interest. As shown in Figure 1.13, the non-contact displacement sensors have the upper frequency limit at approximately 2000 Hz. But already in the range from 1000 to 2000 Hz, measurements performed by non-contact proximity probes are very suspicious because it is not possible to adequately eliminate the influence of unevenness of the shaft surface, which is comparable to the measured displacements.

Velocity transducers are limited because of their design to frequencies of approximately 10-1500 Hz. Accelerometers that can measure frequencies lower than 1 Hz to about 30 kHz have the widest frequency range.

Further on, the individual types of sensors, their typical applications and mode of operation will be described.

2.2.1 Displacement Transducers

There are several types of sensors to measure displacement, distance or position. The oldest type is probably a contact mechanical slider; nowadays, the often used type is a non-contact sensor based on eddy currents - proximity probe which operates on the base of change of Foucault currents - the resistance of the material changes due to the change in distance. Other types, such as laser, ultrasonic, capacitive or inductive sensors, also exist. Displacement sensors are quite complex systems; so, they are only used for shaft vibration measurement - they measure vibrations of shaft relative to a part of stator, usually relative to the bearing housing.

The proximity probe based on eddy currents measures the distance between the sensor tip and a conductive surface. The measuring system comprises the sensor and the proximitor (see Fig. 2.2). The oscillator in proximitor generates high-frequency alternating current that passes through a coil embedded in the sensor tip and creates high-frequency electromagnetic field around the tip of the sensor. Bias voltage used to be -10 Vdc, but may be up to -24 Vdc (depending on the manufacturer); an alternating component has a frequency of about 1.5 MHz (depending on the manufacturer). The electromagnetic field in the coil induces *eddy (Foucault) currents* in the conductive material. These eddy currents absorb energy from the system, resulting in the change in impedance of the coil. Instant distance to the target surface would modulate itself onto this wave and then is demodulated. With respect to the high frequency of the electromagnetic field, the entire measurement is strongly dependent on the total resistance (all of ohmic, inductive and capacitive resistance). Cables leading the high-frequency signal are produced in strict tolerances of electrical values and their length cannot be modified. Any damage to the cable or the shield threatens the quality of measurements. After affecting the carrier wave and eddy current by the variable distance of the target surface during the vibration, the signal is led back to the demodulator. Then - now already low-frequency - the signal is led to the evaluation unit.

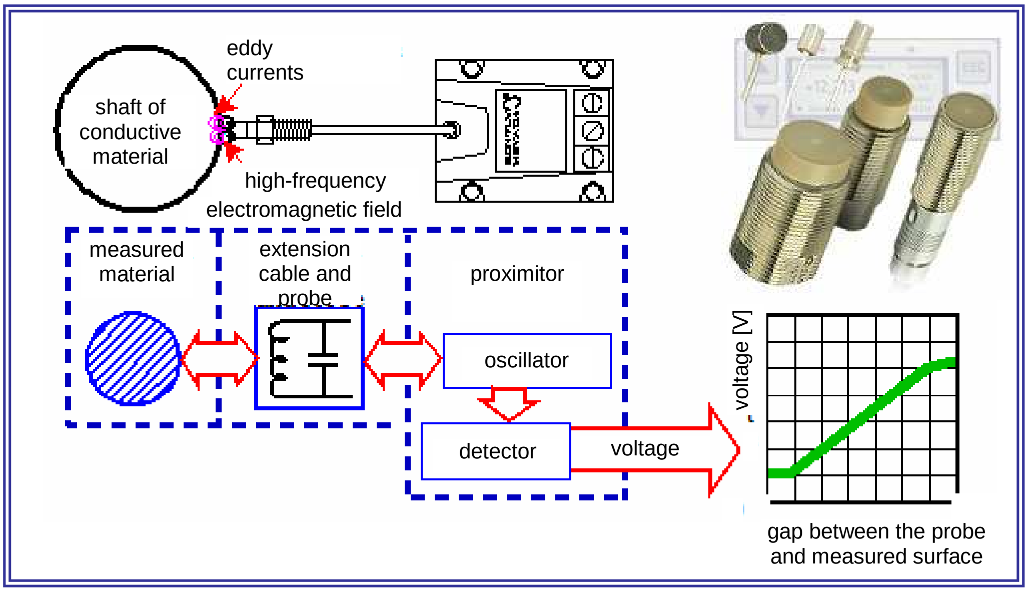


Fig. 2.2 - Scheme of the Proximity Probe System Based on Eddy Currents

If the distance between the tip of the sensor and the conductive surface is constant, output voltage depends on the frequency of the electromagnetic field, the conductivity of the measured material and its magnetic permeability. It is obvious that sensors of this type are supplied according to a particular shaft material and may not be used for the shaft made of different material. In Fig. 2.3 an example of the sensitivity characteristics of the same probe to different target materials is shown. A common value of sensitivity is 8 mV/µm.

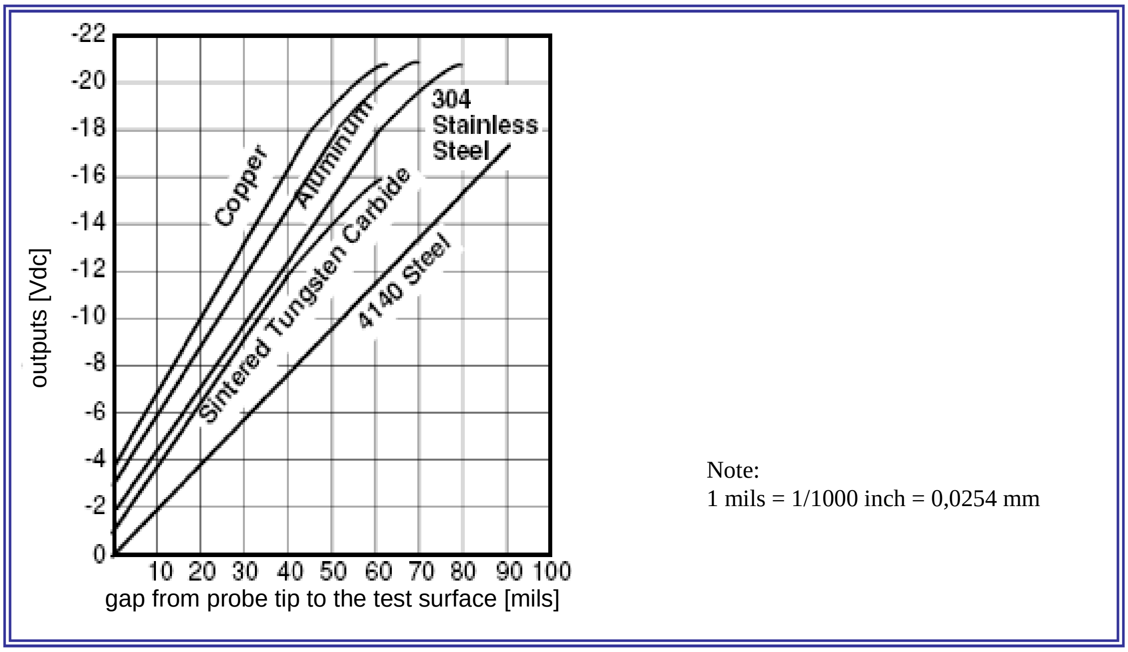


Fig. 2.3 - Example of the Probe Sensitivity to Different Materials

It is therefore appropriate to have a sizing agent, which is simple - the micrometer screw and a bracket for the target material. In addition, it is very important to specify the type of probe with regard to the target material while purchasing the probe.

Note: These probes have a relatively small range of distances in which the output signal is linear (typically in the range 0.25 to 2.3 mm). From a user perspective, this means that the probe should be set in the middle of the zone with linear response at rest. The manufacturer gives recommendations - both in mm and in bias voltage on the probe at rest. Graph of voltage as a function of the gap (calibration curve) is supplied together with the probe.

Proximity probes are compact devices without any moving parts, so they provide the same output regardless of the position in which they are mounted. Usually, two proximity probes rotated by 90° are used (see Fig. 2.10). The use of proximity probes will be described in more detail in the section 2.4 dealing with the shaft vibration.

2.2.2 Velocity Transducers

Velocity transducers have been used as vibration sensors for rotating machinery for a long time and they are still in use in many applications. They operate on the principle of electromagnetic induction: when the coil moves in a magnetic field, voltage is generated on the coil outlets. This induced voltage is caused by energy transfer from the magnetic field to the coil. The amount of the induced voltage is directly proportional to the relative speed between the coil and the magnetic field. Velocity transducer is designed so that this relative speed reflects velocity of vibration of the measured machine.

Velocity transducer (see Fig. 2.4) needs no external device for its functioning. It has two main components:

- Permanent magnet that is fastened to the sensor's housing and therefore with the case of the measured machine (when properly mounted).
- Coil (hollow bobbin wrapped with wire) which is hanged on very soft springs and which due to its inertia remains at rest, while the permanent magnet vibrates with the case. Sometimes the wire is wound on the bobbin two-way to distract external electric fields.

Usable frequency range of the velocity transducer is determined by mechanical parameters of its components. Spring stiffness and material damping and the weight of the coil determine the sensor response at low frequencies. Usually, the resonant frequency of the sensor is below 10 Hz and the usable frequency range is 10 to 1000 Hz. To increase the resonant frequency damping, the sensor housing can be filled with oil.

During installation, the sensitivity axis of the transducer should be considered - due to the effects of gravitational forces, sensors specified for measuring in vertical direction should have different design from those specified for horizontal direction. Velocity transducers are very sensitive to lateral vibration (other than in the sensitive axis direction) - the coil could get in touch with the sensor housing and cause damage of the sensor. The place for sensor mounting should be adjusted so that the surface was straight and slightly larger than the base of the sensor.

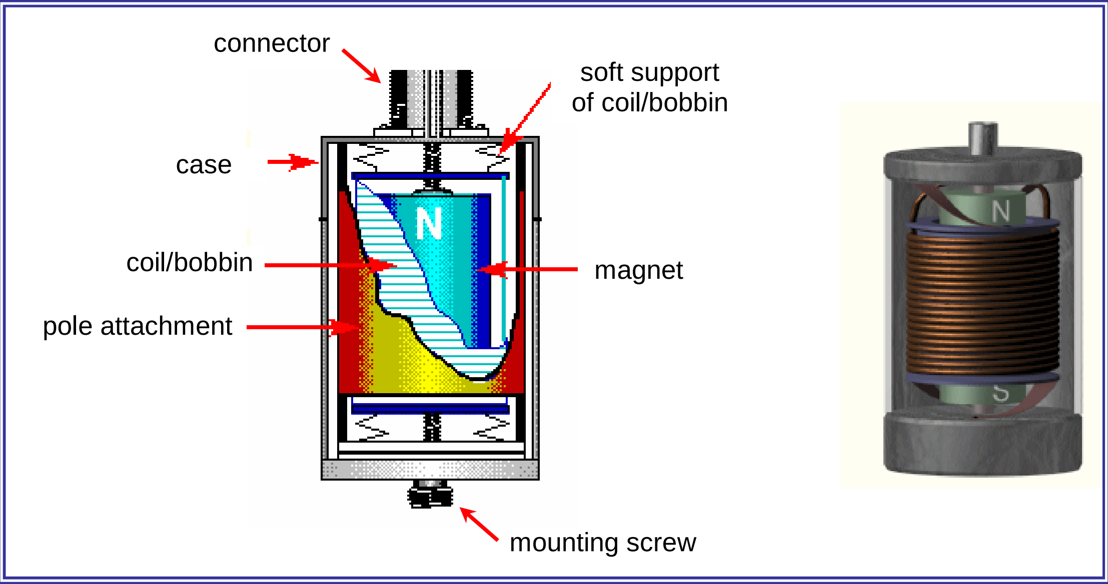


Fig. 2.4 - Velocity Transducer

Velocity transducers have the advantage that they are low-cost and quite sensitive; the disadvantage is that because of their fragile design they are susceptible to shocks and are not suitable for "manual" measurements, they are used only permanently mounted on the machine case using screw.

Velocity transducers of modern type are laser sensors that operate on the base of the Doppler effect. They are very expensive and not commonly used in diagnostics.

2.2.3 Accelerometers

Nowadays, an accelerometer is used as a basic vibration sensor, particularly for measuring on stationary parts of the equipment (rotating machinery), as it has minimum disadvantages in comparison with previous types of sensors (such as the one that it cannot measure at 0 Hz). If velocity or displacement is needed, this information can be obtained by integrating the signal from the accelerometer.

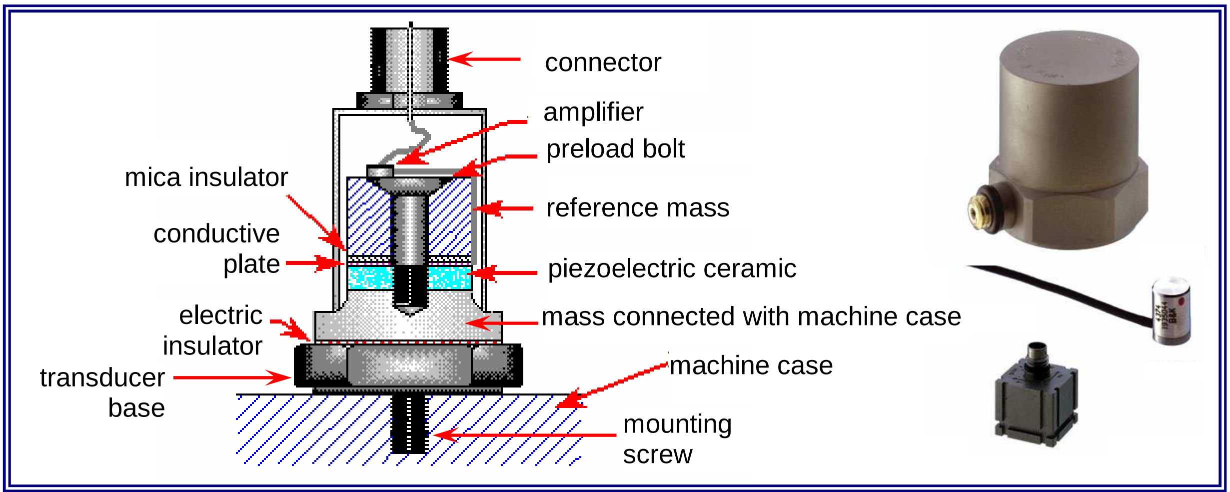


Fig. 2.5 - Accelerometer (Pressure Type)

The principle of accelerometer operation is that deformation of a plate of piezoelectric ceramic material results in an electrical charge, magnitude of which is directly proportional to the deformation. The ceramic plate is placed between two masses, one of which is inertia (seismic) reference mass and the other is firmly attached to the sensor case and therefore to the case of the measured machine (see Fig. 2.5). The seismic mass is flexibly supported; pre-stressed bolt with piezoceramic plate is here considered as a spring, so that the stiffness to the mass ratio of this system is large and the resonant frequency of the sensor itself is very high.

When the measured machine vibrates, the mass firmly attached to the transducer case also vibrates, while the reference inertial mass remains at rest. This creates a force applied to the piezoelectric element that deforms and creates an electrical charge that is proportional to the acceleration of the machine case. Thus, the accelerometer measures absolute vibrations (stationary inertial mass is the reference) unlike proximity probes that measure relative vibrations (vibration of the rotor relative to the stator, which itself may vibrate).

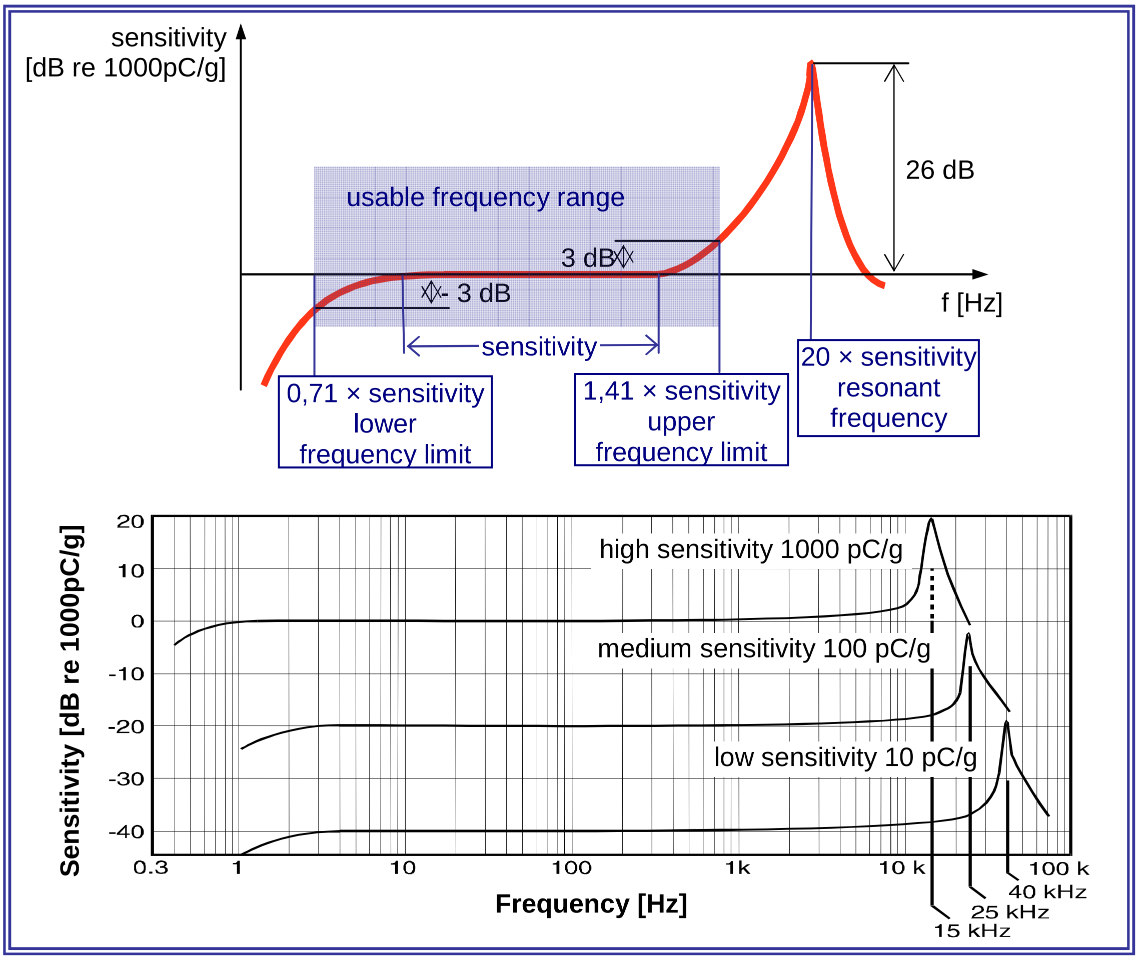
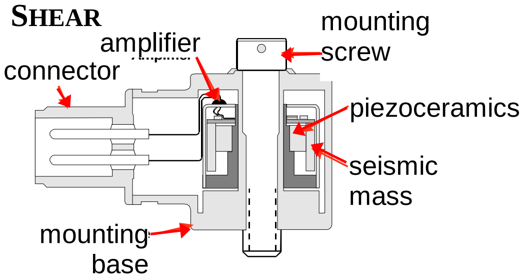
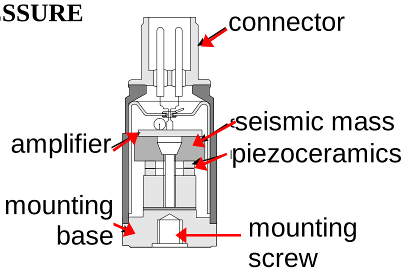
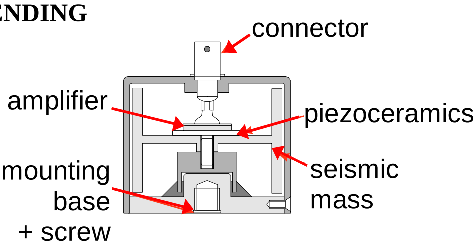


Fig. 2.6 - Calibration Curve of an Accelerometer

An integral part of delivery of each accelerometer is the *calibration curve*, from which the sensor's sensitivity and frequency range of its applicability can be read (see Fig. 2.6 above). The sensitivity of the accelerometer is given in pC/ms^{-2} or in pC/g. If the measured structure vibrates on the frequency below the lower frequency limit of the used sensor, the inertial mass moves together with the casing and does not generate an electric charge. This lower frequency limit of the sensor is given by the stiffness of elastic support, by material damping and weight of the reference mass. On the calibration curve, it is indicated by the decrease in nominal sensitivity by 3 dB. Generally, the larger the sensor is, the better its sensitivity is and the lower frequencies can be measured. The upper frequency limit of the sensor is in about one third of its resonant frequency. On the calibration curve, it is indicated by the increase in nominal sensitivity by 3 dB. The actual calibration curves for three different sensors are shown in Fig. 2.6 below. The greater the weight of the sensor (or sensor's reference mass) is, the higher the sensitivity and the lower the resonant frequency are.

Electric charge is a quantity that cannot be transmitted over long distances. That is why older accelerometer types were supplemented with an additional external device - a charge pre-amplifier, which transformed the electric charge to the voltage. Modern accelerometers already have such a preamplifier integrated in its case and then their sensitivity is expressed in mV/g.

Table 2.1 - Accelerometer Types

<table>
<tr><th>ACCELEROMETER TYPE</th><th>ADVANTAGES</th><th>DISADVANTAGES</th></tr>
<tr><td>SHEAR

</td><td>wide frequency range
quite durable
low temperature influence</td><td>lower sensitivity</td></tr>
<tr><td>PRESSURE

</td><td>wide frequency range
durable to shocks</td><td>sensitive for temperature effects
sensitive for base deformation</td></tr>
<tr><td>BENDING

</td><td>very low frequency measurements
very high sensitivity</td><td>fragile, sensitive for shocks</td></tr>
</table>

The design allows accelerometer to be attached to the machine case in any orientation - the mass/spring system is stiff enough so that the orientation does not matter (unlike the velocity transducer). However, it is important that the piezoelectric element in the sensor is not exposed in any other way than from vibrations. Therefore, the place where the transducer is to be mounted should be smooth and flat to prevent the deformation of the transducer mounting base. Also, temperature changes and excessive torque on mounting screw could cause deformation of the base and therefore a false signal. Pressure type of accelerometer, which is shown in Fig. 2.5, is the most prone to these parasite loads. Other two versions of accelerometers - bending and shear types - are much more resistive to them. Advantages and disadvantages of each type are shown in Table 2.1. Shear type is nowadays produced e.g. as *delta-shear*, in which three piezoceramic elements are arranged in a triangle and are oriented so that they are minimally affected by deformation of the base. This type of accelerometer is sensitive enough and durable and does not have disadvantages of the pressure type. It is the most common type of transducer used for absolute vibration measurements.

Due to the fact that accelerometers do not contain any moving parts, they are durable and reliable and do not require frequent calibration check as velometers. Their installation on the machine case is easy, they can be used in a wide range of frequencies (from 0.1 Hz to 30 kHz) and they have large dynamic range. Designs for high operating temperatures are also available.

2.2.3.1 Mounting of Accelerometers

An important aspect to be considered when using any transducer is the attachment to the machine case. It should be realized that each sensor measures only what is happening to itself. Attachment must ensure that this would be the same to what is happening to the machine case the vibration of which is of interest. Any attachment adds to the sensor (which is a system with one degree of freedom - mass and spring) additional mass and spring, consisting of flexibility of the joint. Inappropriate attachment can completely degrade the measured data, or at least significantly reduce the usable frequency range of the sensor. The way of attachment is chosen so that the sensor worked safely in the frequency range of interest. There are several ways to mount the accelerometers, which are in detail described in ISO 5348, e.g.

- *screw* - This is the most reliable attachment which only marginally reduces the usable frequency range of the sensor. The surface under the mounting base should be clean and flat, the whole base should be in contact with the machine surface, the screw hole should be perpendicular to the surface and deep enough and the correct thread should be used (transducers often have a different type of thread than the metric one). Requirements for quality of machining and perpendicularity of the threaded hole are quite strict. It is difficult to comply with these requirements on already manufactured machines. Therefore, adhesive pads of stainless steel are often used (rollers with carved thread for mounting screw, but magnetic - for indicative measurements with the accelerometer attached by a magnet). They have the advantage that they can be easily replaced in case of a damage (e.g. during repairs).

- *glue (adhesive)* - When suitable adhesive is used (e.g. Loctite Depend, Loctite Liquid Metal) and the surface for sensor attachment is properly adjusted (clean, roughened), the usable frequency range of the sensor is also preserved. Note: Some adhesives have intentionally an additional soft component (to compensate for inequalities, etc.), which significantly reduces the frequency range, so it is necessary to use only the recommended adhesives. In practice, the two-component adhesive HBM X60 for strain gauges is the most used one. Typically, pads with carved thread for mounting screw are glued to the machine surface rather than the sensor itself, because its removal when glued directly could result in its damage. Attention must be paid at the maximum operating temperature at which the glue softens - it is about 80-100°C.

- *double-sided adhesive tape* - quick method, less reliable, not too common.
- *magnet* - Frequent and quick way of mounting a sensor, but suitable only for common operational measurements in the frequency range up to about 2 kHz. Again, there must be good contact surface (flat, without paint, scales, etc.). The transducers are supplied with special rare-earth magnets, which are stronger than conventional magnets.
- *beeswax* - A quick way of mounting a sensor, which is mainly used for laboratory measurements and for smaller sensors. It is not applicable in diagnostics. Usable frequency range of the sensor is slightly reduced by this attachment. It is applicable only to the temperature up to 40 ºC.
- *touch needle* – Applicable only for indicative measurements up to 1 kHz.

2.2.3.2 Cabling

Cabling is an integral part of the measurement chain. Through cabling, signal from the accelerometer is transmitted into the analyzer and effects that devalue the measured data can occur as well. Even after the charge is converted to voltage, the signal is very weak. In an industrial environment, relatively strong electromagnetic fields come into account, and, moreover, static electricity generated by dry friction appears. Signals from the accelerometers are transmitted by special shielded cables, but it must be ensured that the shield was grounded at only one end of the cable; otherwise, this may lead to closure of a ground loop. As a protection against ground loops, a sensor with insulated base can be used (such as the one in Fig. 2.5). If not available, insulated screws or mica pads can be used with accelerometers. If measurements are performed with attaching accelerometers using a magnet, it is often enough to insert a sheet of paper between the magnet and the sensor and observe how the signal changes (improvement of the signal means that measurement was really influenced by the static electricity). Ground loops are very uncomfortable especially for multi-channel measurements when transducers for simultaneous measurements are placed on different parts of machines with different static potential.

2.2.3.3 Dynamic Range

After ensuring that the analyzer was fed with a high-quality signal from the sensor by following all the above principles, this signal must be adjusted to suit the internal electronic circuits. Therefore, the input to the analyzer includes an input amplifier/attenuator of the signal that modifies the signal so that it is strong enough, but overload is avoided. This makes optimum use of the dynamic range of the measurement chain. The input amplifier can be set:

- AUTORANGE - It is most common and is recommended for most steady mode measurements. During the first measurement, the instrument evaluates the signal strength and sets the input amplifier according to this measurement (with some safety margin). For purposes of measurement itself, this first measurement is lost; so, autorange cannot be used with transient events. Also, when the signal strength changes significantly during the measurement, the instrument needs to re-adjust the input amplifier and thus some data may be lost.

- MANUAL - It is used for transient (impact or shock) measurements, measurements of run-up and coast-down and other unrepeatable events. The user manually adjusts the estimated maximum value that the measured quantity could reach. When the signal exceeds this value, the analyzer would report OVERLOAD, but it does not measure and the data are not distorted.

2.3 Vibration Measured on Non-Rotating Parts of a Machine (Absolute Vibration)

Basic task of vibration diagnostics is not to measure, but to assess the machine condition. To assist in this task, standards - currently created for the evaluation of the overall rms value of vibration velocity or of the peak amplitude - are used. These standards represent guidelines that are generally not obligatory. However, when they become a part of trade agreements between the supplier and the buyer, they become obligatory.

Basic principles recommended for measuring and assessing the absolute vibration are described in ISO 10816-1 *Mechanical vibration - Evaluation of machine vibration by measurements on non-rotating parts - Part 1: General Guidelines* and its amendment 10816-1/Amd.1. In other parts of the standard - ISO 10816-2 to 7 - assessment of machines according to their type is elaborated (pumps, reciprocating machinery, etc.).

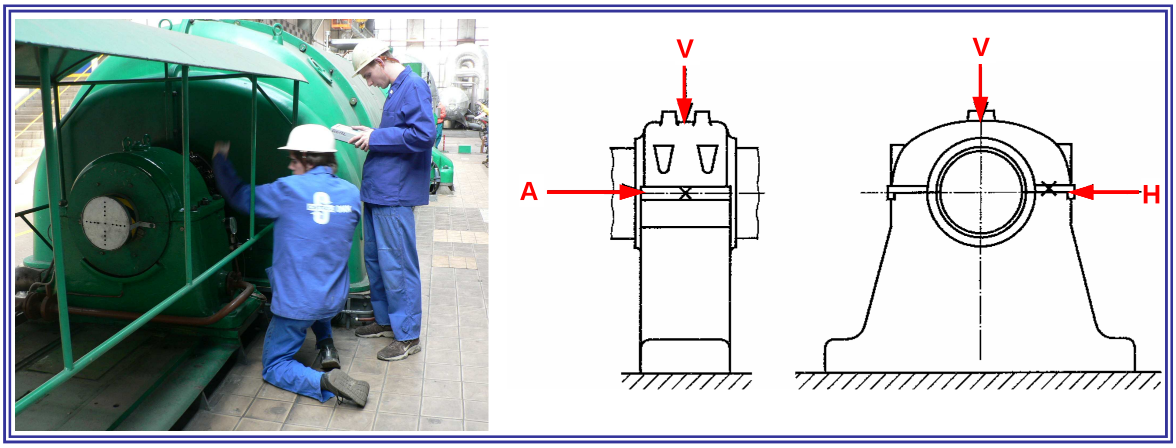


Fig. 2.7 - Selecting Measurement Points and Directions on the Machine

2.3.1 Selecting Measurement Points

The first step that should be done in machine monitoring is selecting the appropriate measuring points, i.e. selecting places which periodic vibration measurements will be performed in. These points are selected principally at locations, where the forces are transmitted from the rotor to the stator parts, usually on the bearings (on the bearing housings). However, safety of personnel carrying out the measurements should be taken into account as well. Measurement points are usually prepared so as to ensure that the authorized employee will always carry out measurements in the same place. This means that either pads with thread for mounting screw are glued on the machine case, or at least flat and clean spots on the surface are created (for attaching the transducer with the help of a magnet). In all the defined measurement points, one to three directions are measured by the needs and the availability: A - axial direction, H - horizontal direction, V - vertical direction (see Fig. 2.7). For numbering and labelling of the measured points, to use the MIMOSA convention is recommended. MIMOSA = Machinery Information Management Open Systems Alliance, which is stated in appendix D of the standard ISO 13373-1 (*Condition monitoring and diagnostics of machines - Vibration condition monitoring, Part 1: General procedures*). According to this convention, bearings are numbered successively, starting from the free end of the drive shaft (see the example in Fig. 2.8).

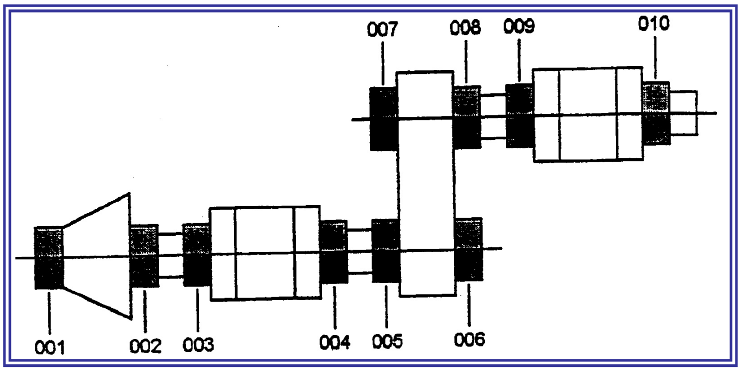


Fig. 2.8 - Labelling of Bearings According to MIMOSA

2.3.2 Criteria of Assessment according to ISO 10816

Standard ISO 10816 provides guidance for the assessment of machine condition for different types of machines based on two criteria:

I. vibration magnitude
II. change in vibration magnitude

2.3.2.1 Criterion I: Vibration Magnitude

Standard ISO 10816 is based on measuring the ***total rms value of vibration velocity in the frequency range from 10 to 1000 Hz***. The highest value of measurements at different locations is called the *vibration severity*. The standard defines evaluation zone limits of the vibration severity. Based of these limits, a machine can be classified according to its state into one of 4 zones:

zone A - Vibration of newly commissioned machines would normally fall within this zone.

zone B - Machines with vibration within this zone are normally considered acceptable for unrestricted long-time operation.

zone C - Machines with vibration within this zone are normally considered unsatisfactory for long-term continuous operation. Generally, the machine may be operated for a limited period in this condition until a suitable opportunity arises for remedial action.

zone D - Vibration values within this zone are normally considered to be of sufficient severity to cause damage to the machine.

Classification of the machine into the zone would help to decide about the future operation of the machine and to propose necessary actions (e.g. immediate shutdown, operation to the next scheduled repair, etc.). Zone limits are indicative values rather than strict ones and may be adjusted on the base of experience of the manufacturer or operator.

Parts 2 to 7 of the ISO 10816 standard define the limits of vibration severity for various types of machines. Machines that are not listed in these sections are assessed according to Amendment to Part 1: ISO 10816-1/Amd.1:

- zone A/B boundaries 0.71 - 4.5 mm/s
- zone B/C boundaries 1.8 - 9.3 mm/s
- zone C/D boundaries 4.5 - 14.7 mm/s

This amendment also applies to models of machines that will be measured in the laboratory of Department of Mechanics. It can be seen that the zone boundaries are set quite widely. The standard also states that small machines (such as electric motors with output up to 15 kW) tend to be at the lower end of the scale and larger machines (such as drives on supports which are flexible in the direction of measurement) tend to be at the upper end of the range.

Standard ISO 10816 assesses machines on the base of measurements in the frequency range 10 Hz to 1000 Hz, where, according to the experience, most of the information about the machines occurs. However, there are applications where the use of this frequency band would be incorrect. Frequency of 10 Hz means 600 rpm, but lots of machines operate at a lower speed. In such cases, lower frequency limit of the measurements needs to be changed. It is recommended that the lower frequency limit was under 1/3 of the running speed. The upper frequency limit of 1000 Hz is also in some cases inadequate, especially if gear transmissions are part of the machine. It should be considered that the gear mesh frequency can be above this limit. If, for example, the rotational frequency is 50 Hz (i.e. 3000 rpm) and a pinion has 23 teeth, then the gear mesh frequency would be 50 x 23 = 1150 Hz and is thus above the measured frequency range. In both these cases, it is necessary to use measurements in frequency range that differ from those defined in the standard for assessment of machines. Therefore, the measurement frequency range should be explicitly stated in the measurement report.

2.3.2.2 Criterion II: Change in Vibration Magnitude

This criterion provides an assessment of a change in vibration magnitude from a previously established reference value measured under steady-state operating conditions. A significant increase or decrease in broad band vibration magnitude may occur, which requires some action even though the zone C of Criterion I has not been reached. Such changes can be instantaneous or progressive with time and may indicate that damage has occurred or be a warning of an impending failure or some other irregularity.

To assess a machine according to the criteria II requires long-term monitoring. As stated in section 1.2, the diagnostic activity has two significantly different phases:

1. Continuous *monitoring* of the machine. The purpose is to determine the deviation from the normal state - to detect the failure.

2. The actual *analysis* of the problem which the term *diagnostics* in the narrower sense is used for. The purpose is to determine the cause of the deviation from the normal state and to determine the root cause of the failure development.

One of the most important tools used in monitoring is the analysis of *trends* of the values of the monitored quantities, namely the development of the monitored quantity over time. Fig. 2.9 shows four possible trends:

A - consistently good condition, no failures

B - a sudden change (breaking off a part of equipment, failure, etc.). It should be noted that both sudden increase in vibration and its sudden decrease is considered to be a serious change. For example, if the constant vibration level was caused by unbalances, then a sudden drop in vibration may indicate that the piece of equipment has broken off, which caused a spontaneous balancing - but still, the broken part may cause a serious subsequent failure.

This vibration trend may sometimes be caused by improper maintenance action during a repair of the machine. Therefore, it is necessary to keep the principle that the vibrations are measured both before the repair and shortly after the repair.

C - a typical trend for progressive damage, e.g. due to wear

D - absurd "trend" - for example, it may be caused by improper selection of measurement points, by poor transducer mounting, etc.

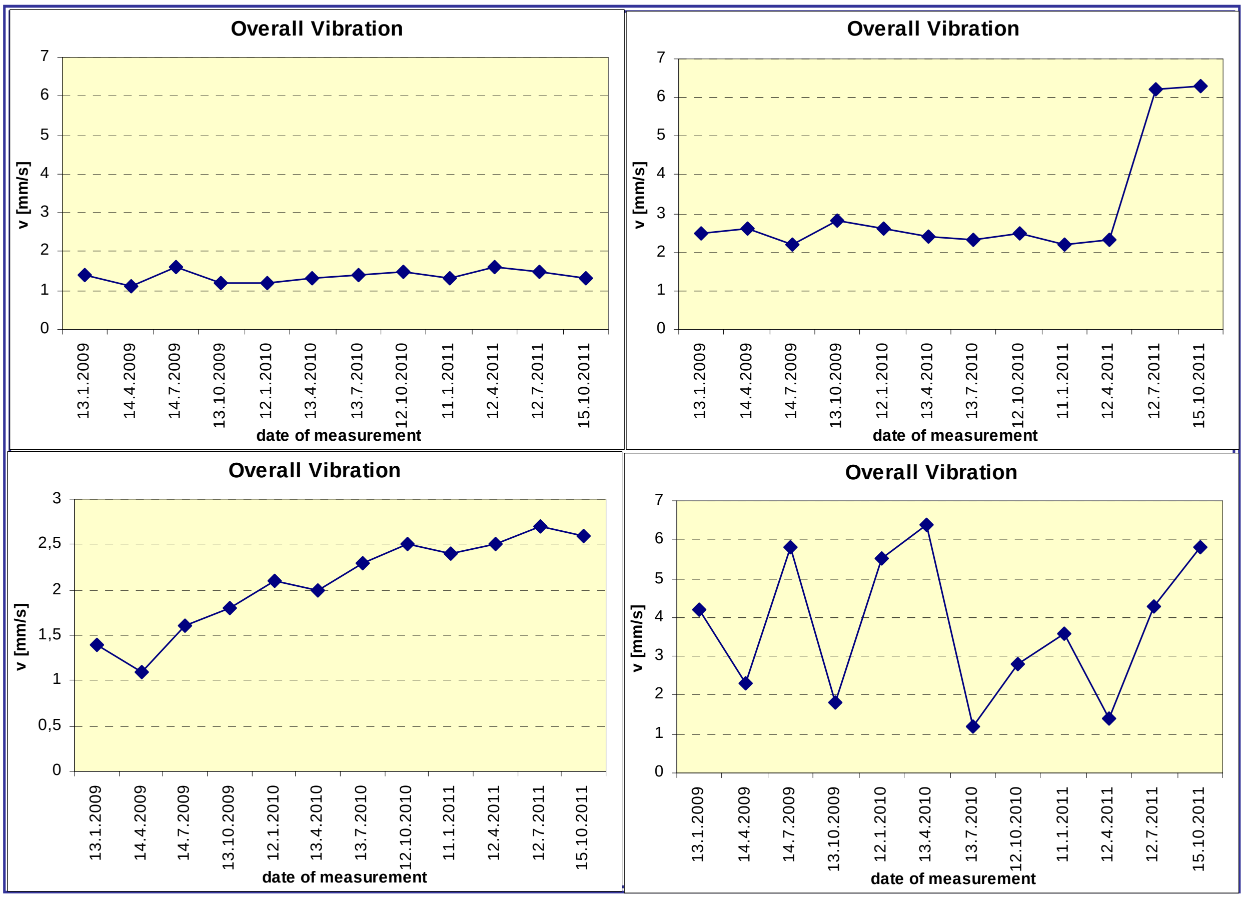


Fig. 2.9 - Trends of Overall Vibration Values

2.3.2.3 Operational Vibration Limits

For long-term operation, it is a common practice for some machine types to establish operational vibration limits. These limits take two forms:

ALARM - To provide a warning that a defined value of vibration has been reached or a significant change has occurred at which remedial action may be necessary. In general, if an ALARM situation occurs, operation can continue for a period whilst investigations are carried out to identify the reason for the change in vibration and define any remedial action.

TRIP - To specify the magnitude of vibration beyond which further operation of the machine may cause damage. If the TRIP value is exceeded, immediate action should be taken to reduce the vibration or the machine should be shut down.

Different operational limits, reflecting differences in dynamic loading and support stiffness may be specified for different measurement positions and directions. In the ISO 10816 standards, there are guidelines for setting criteria for both operational limits for certain types of machines. One of the recommendations for the ALARM limit says: On the base of experience, the reference value is set and 0.25 of the zone C lower limit is added to it.

2.4 Shaft Vibration

Basic forces that excite the vibration are applied to the machine rotor. However, when the vibration response is measured on bearing housings or on a machine case, it is in fact a response to those forces that is transmitted from the rotor to the stationary parts of the machine. If the mass of the machine casing is much greater than the mass of the rotor, even relatively large forces acting on the rotor and causing its adequate response are significantly damped after transmission to the case. This case occurs in high-pressure machines (turbines, compressors, etc.) that have much more massive casing than the rotor. Therefore, in such cases measurements of rotor (shaft) vibration are performed.

The almost exclusive measured quantity is the **vibration displacement** in micrometers (microns) [**µm**]. Usually, the maximum displacement or peak-to peak value is measured (see Fig. 1.9). There are two basic ways of measuring the shaft vibration (according to ISO 13373-1):

- <u>Relative shaft vibration</u> is the shaft vibration towards the stator part (mostly towards the bearing housing) which also vibrates itself.

Such measurement is important particularly for assessing the relation between shaft vibration and clearances at different parts of the rotor system against the stator - to avoid contact between rotating and stationary parts. The scheme of measuring relative shaft vibration is in Fig. 2.10 on the left. The typical location of two displacement transducers (45° from the vertical) and their alternate location (horizontal and vertical) can be seen there.

- <u>Absolute shaft vibration</u> is the shaft vibration with respect to the non-moving ground.

They are used for evaluating the vibration from the point of view of stress load to the shaft/rotor components by dynamic forces. Their measurement is much more difficult than the relative vibration measurement. In case that this type of monitoring is necessary (for large machines - over 1000 MW), this measurement is performed as a combination of measurements of relative rotor vibration using proximity probes and measurements of absolute case vibration using seismic transducers (accelerometers or velometers), see Fig. 2.10 on the right. This combination is exacting as to the accuracy of both systems and particularly to necessity to remove those phase differences of the two signals that are due to the electrical effects of the measuring system and are not related to the process mechanics. Such a system has not been installed in the Czech Republic yet.

An alternative way of measuring absolute shaft vibration is to use a slider - shaft vibration is taken with a slider (usually in the shape of rod which touches the shaft) which transmits vibrations to the seismic sensor, placed on this slider. The disadvantage of this method is relatively rapid wear of the contact place which is formed by carbon (the same as is used e.g. for brushes of electric machines). At present, it is rarely used in diagnostics, just in case of serious problems. It is not used for monitoring.

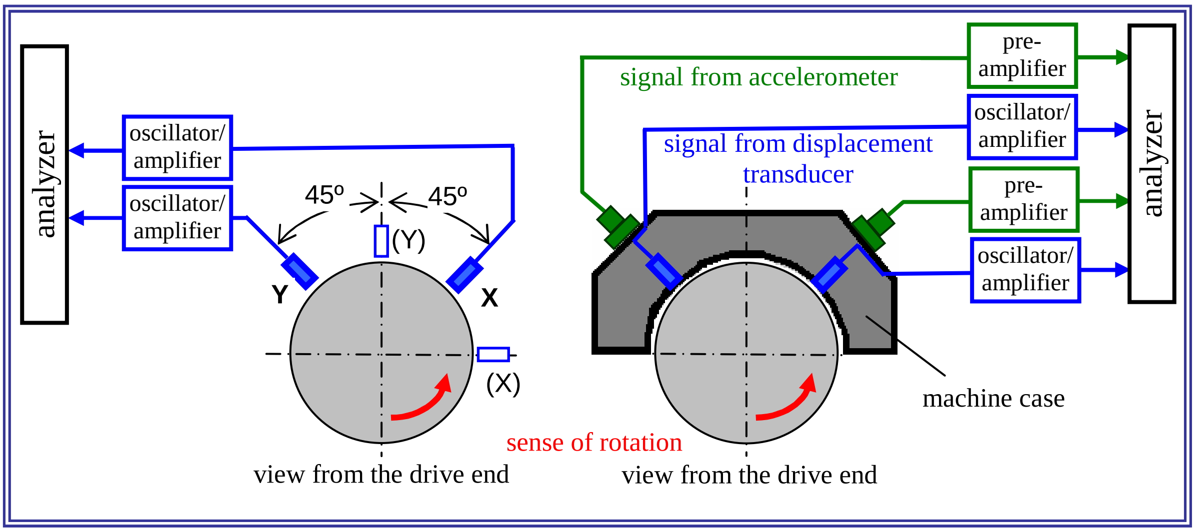


Fig. 2.10 - Scheme of Shaft Vibration Measurement (left - relative, right - absolute)

2.4.1 Proximity Probes Installation

Installing the displacement sensors (proximity probes) requires compliance with certain rules. Sensitivity of the sensors to the material of the measured shaft was already discussed in section 2.2.1. This section focuses on aspects of the sensor mounting itself. In terms of installation, two rules are important:

- No metal should be in the vicinity of the tip where a coil creating eddy currents is. Fig. 2.11 on the left shows a typical arrangement - drilling around the probe must be at least twice the diameter of the probe tip.
- Created eddy current fields should not influence each other (see Fig. 2.11 on the right). Such a problem could arise namely when shafts of small diameters are measured (e.g. on Bently Nevada model in the laboratory). In this case, the adjacent probes should be placed to different radial planes so that the distance between them is increased.

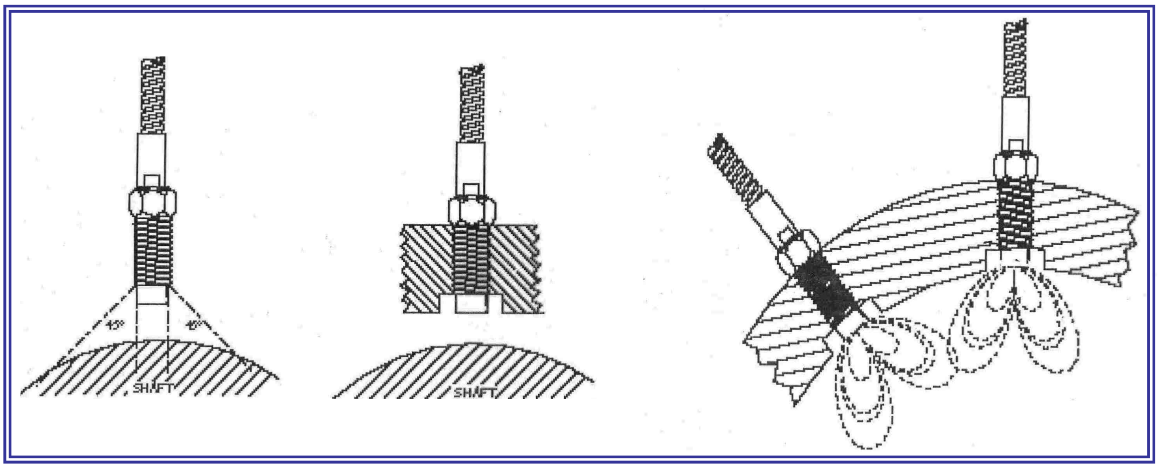


Fig. 2.11 - Installation of Proximity Probes with Respect to Eddy Currents Influencing

Accuracy of measurements using proximity probes also significantly depends both on the mechanical machining of the target surface and, with regard to high frequency currents, on its electrical properties. The key influences are:

- **mechanical runout**
  - off-roundness
  - bent shaft
  - irregular surface (scratches, etc.)
- **electrical runout**
  - residual magnetism
  - metallurgical segregations (distinct grain boundaries, etc.)
  - local stress concentrations

These issues are discussed in detail in application note of Bently Nevada Comp. [21] and a number of recommendations from manufacturers of measuring systems.

With regard to the great importance of this influence, standard ISO 7919-1 sets a limit value of the runout. Runout should be lower than the greater of the following values:

- 6 µm
- 25% of the allowed vibration level for a new machine (boundary A/B)

The total runout is measured at low rotor speed (200 to 600 rpm) when the influence of centrifugal force is not large. To achieve a small value of the total runout, adjustments of the target areas are made on the manufactured rotor and it is recommended to protect these areas during installation or repair of the machine.

Example from practice: Just placing a small magnetic stand (e.g. during the preliminary alignment of rotors) may result in reporting such high vibration during operation that the machine must be shut down, dismounted (this is time consuming and costly) and the area must be repaired (degaussed). Scratches have similar effects (causing sharp peaks in the signal), etc.

Some monitoring systems perform compensation of signal from the vibration sensor to the total runout. The principle is that the signal from the runout in one turn is subtracted from the signal measured during rotation. Although it seems to be correct, many critics highlight the influence of inaccuracies in measurements of small voltage signals and thus do not recommend this method.

2.4.2 Evaluation of Relative Vibration

For assessing the total value of the relative vibration displacement, series of standards ISO 7919-1 to 5 are applied; they are developed by the same principles as the series of standards for assessing the overall vibration velocity measured on the stationary parts of the machine.

- Peak-to-peak displacement value is evaluated (or S_{max} value - will be explained later)
- The same principal of evaluation as for ISO 10816:
  - vibration magnitude (zones A to D)
  - change in vibration magnitude (of 25% of B/C zones boundary value), change in 1X vector position may be evaluated as well

Zone A/B boundary limit:

$$S_{(p-p)} = \frac{4800}{\sqrt{n}} \ [\mu m]$$

Zone B/C boundary limit:

$$S_{(p-p)} = \frac{9000}{\sqrt{n}} \ [\mu m]$$

Zone C/D boundary limit:

$$S_{(p-p)} = \frac{13200}{\sqrt{n}} \ [\mu m]$$

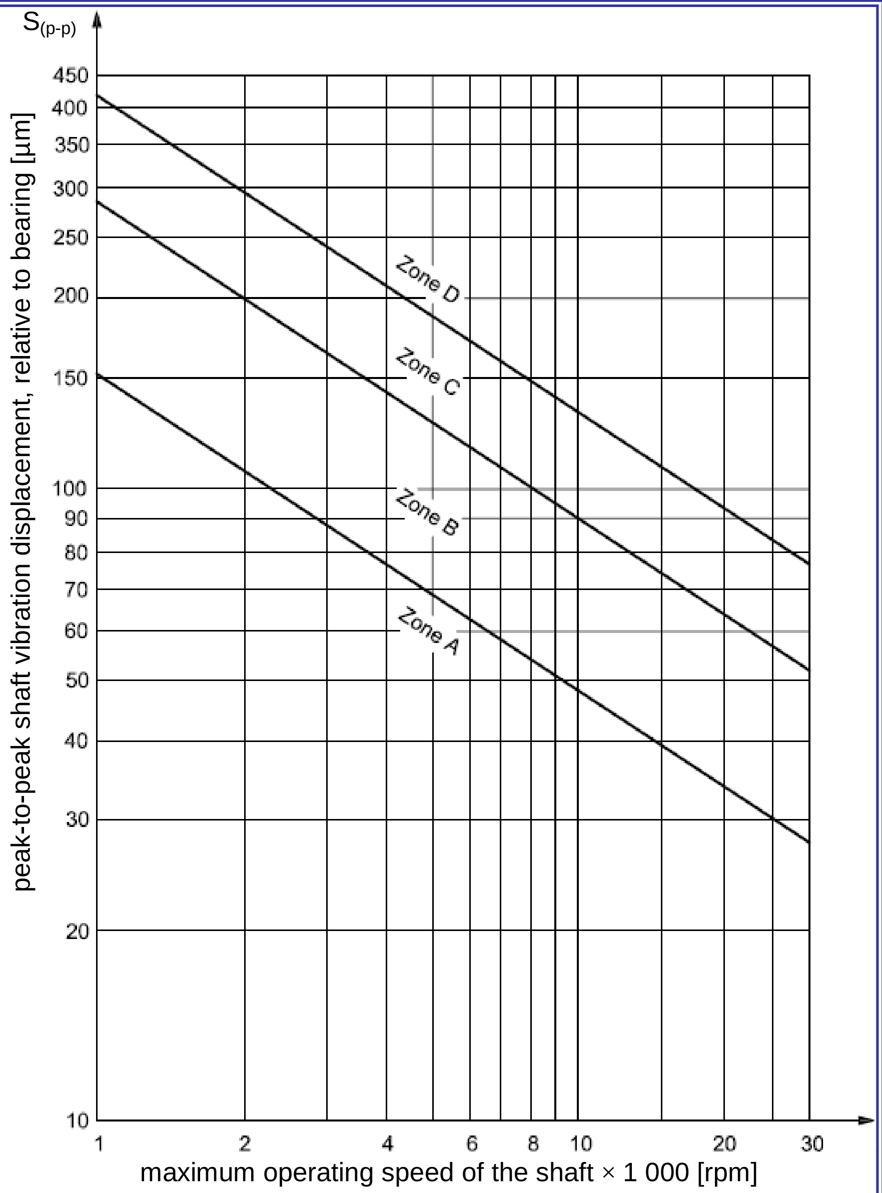


Fig. 2.12 - Zone Boundaries (example from ISO 7919-3 for industrial machines up to 50 MW)

A significant difference between the two sets of standards lies in the fact that the acceptable relative (and absolute) shaft vibration displacement depends on machine running speed. This dependence is expressed both by formulas and graphically (see example in Fig. 2.12). In this example, the boundary values for peak-to-peak displacement for n=3000 rpm are:

Zones A/B boundary 88 μm, B/C 164 μm, C/D 241 μm

As the centre of the shaft is not in one place during rotation - the shaft exhibits a precession movement (the trajectory of the shaft centre is commonly called *orbit*) - there is a problem to place the proximity probes so that the peak displacement value was measured. The solution is usually to use two sensors which are perpendicular to each other. A simpler method of evaluation is based on an evaluation of the larger of the two measured values.

Some companies recommend evaluating a specific value, which is called S_{max}. It is the largest distance from the central position of the shaft to the orbit. Its determination requires the use of special software, which is delivered by these companies only. ISO standards prefer a method based on an evaluation of the larger of the two measured peak-to-peak values. For more information about the value S_{max} see ISO 7919-1.

2.4.3 Interpretation of the Proximity Probe Signal

As the proximity probe has a bias voltage and it measures the instant distance between the probe tip and shaft surface, it gives actually two values - DC component corresponds to the static distance between the probe tip and the shaft, AC component corresponds to the vibration in the direction of the probe axis. Both these parts of information have their diagnostic significance and are evaluated during measurements.

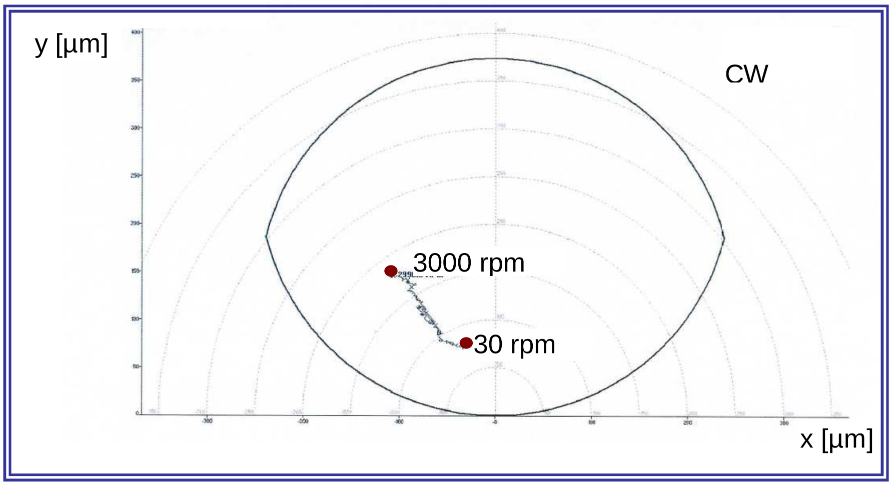


Fig. 2.13 - Display of the Centre of the Shaft Position

2.4.3.1 DC component

As stated above, two proximity probes spaced on the shaft by 90° are commonly used. DC components from both these probes are added and the result corresponds to the position of the centre of the shaft. Well-established way of displaying this position is shown in Fig. 2.13. Contour curve for cylindrical or segment bearings is a circle. This curve indicates the clearance, not the entire hole in the bearing. For bearings with non-uniform clearances in horizontal and vertical directions (e.g. elliptical bearing), the program draws a "shell", which depends on the clearances ratio. At zero speed, the position of the centre of the shaft is shown at the lowest point of this curve.

Figure 2.13 shows the change in position of the centre of the shaft in an elliptical bearing while the rotational speed is increasing (the direction of rotation is clockwise - CW; CCW means counter clockwise). The displayed curve corresponds to the correctly functioning bearing with a well formed oil wedge.

2.4.3.2 AC Component

After removing the DC component, the AC component of signal is obtained. In Fig. 2.14 on the left, unfiltered AC component of a signal, which also contains other frequency components (and sometimes also effects of scratches on the target surface), is shown. On the right, the signal is filtered so that only the AC component is visible. Typically, the filter gives only the signal with rotational frequency. In basic assessment of shaft vibration, this component is usually taken into account. It is also possible to filter for harmonics or around the sub-harmonic component (if any).

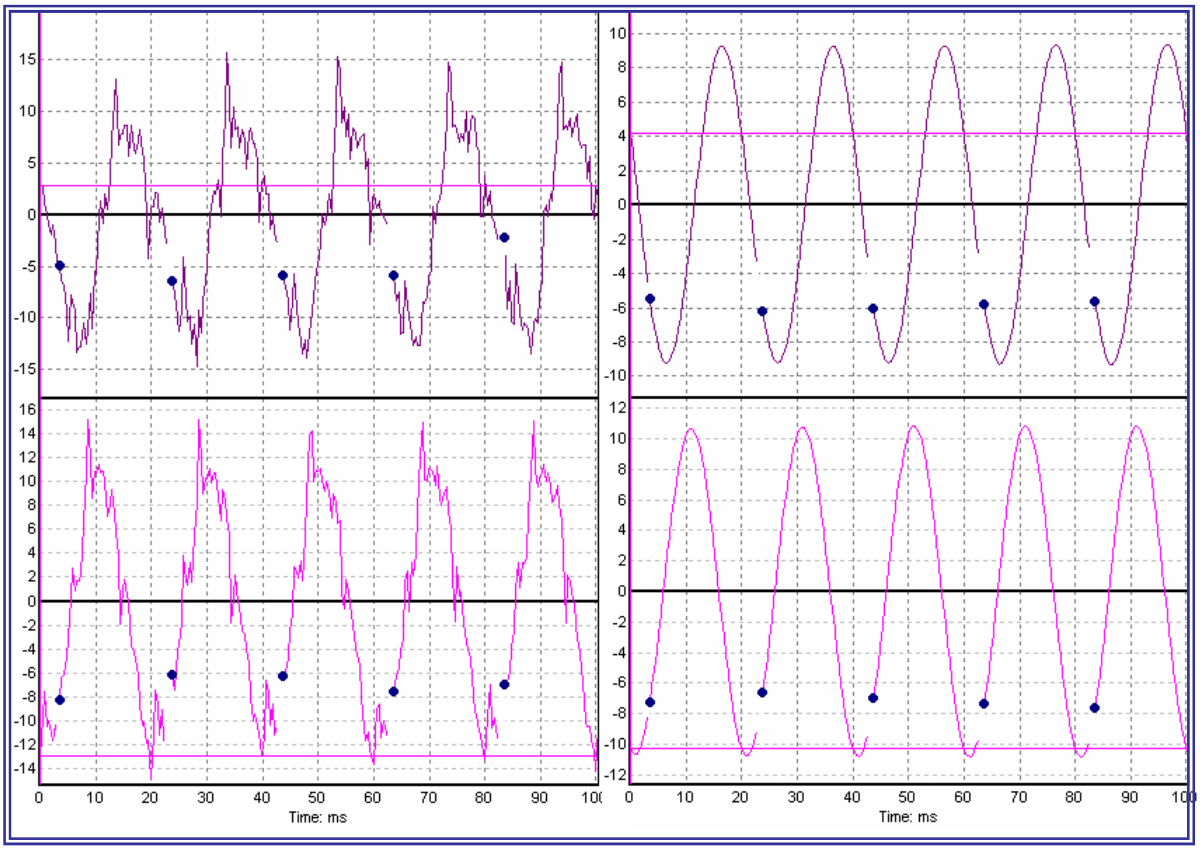


Fig. 2.14 - Non-filtered (left) and Filtered (right) Time Signal

For the shafts in rotating machinery with journal bearings it is typical that the centre of the shaft is deflected from the position at rest, as already explained, but this centre is not "still", but it exhibits the precession motion around its equilibrium position. Representation of this motion is obtained by adding these two signals. In Fig. 2.15, creation of the pattern by adding the two signals from the two perpendicular sensors is shown.

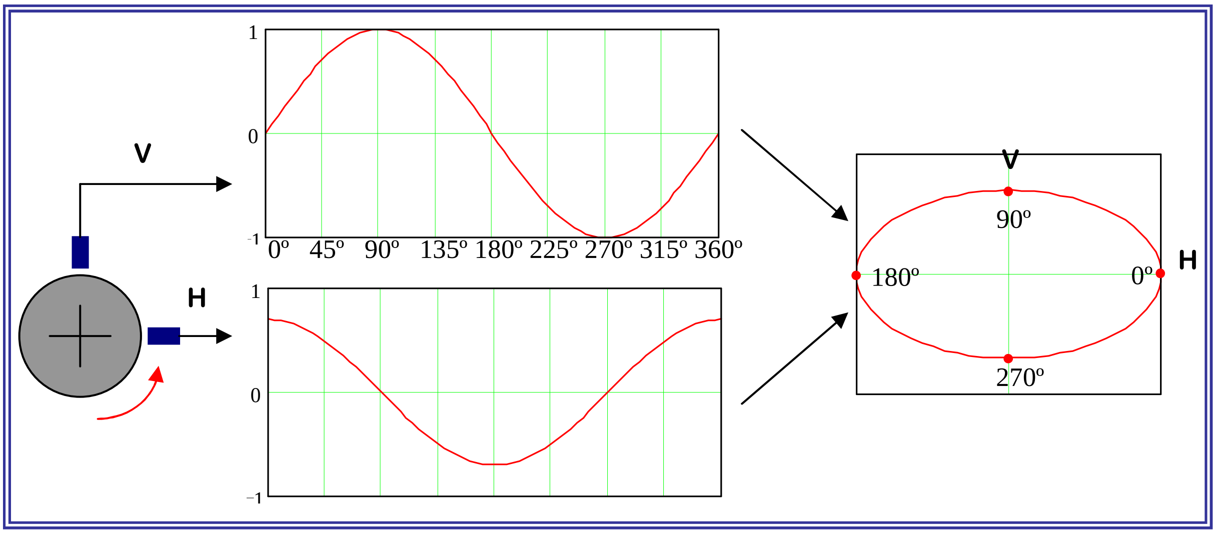


Fig. 2.15 - Principle of Orbit Creation

Note: In physics, you may have met Lissajous curves when worked with an oscilloscope. These curves were more complex because the signals that created them had different frequencies. The pattern shown in Fig. 2.15 has one serious drawback - it is not obvious whether it was created:

- when the motion around the equilibrium position was done in the same direction as the shaft rotation - this case is called *forward precession*
- or when the motion around the equilibrium position was done in the opposite direction as the shaft rotation - this case is called *backward precession.*

It is quite important both from mechanical and diagnostical point of view, particularly with regard to cyclic loading of the shaft, to the critical speed, etc. Therefore, the measurement system is complemented by an external signal derived from the rotating shaft. The measurement is triggered by an external trigger - signal from phase reference (see section 2.5.4.3). Usually, this signal is marked as the point which the orbit begins from, and before its end there is an empty space (see Fig. 2.16). Note: the term *Keyphasor* is commonly used for this mark, even though it is a protected trademark of Bently Nevada.

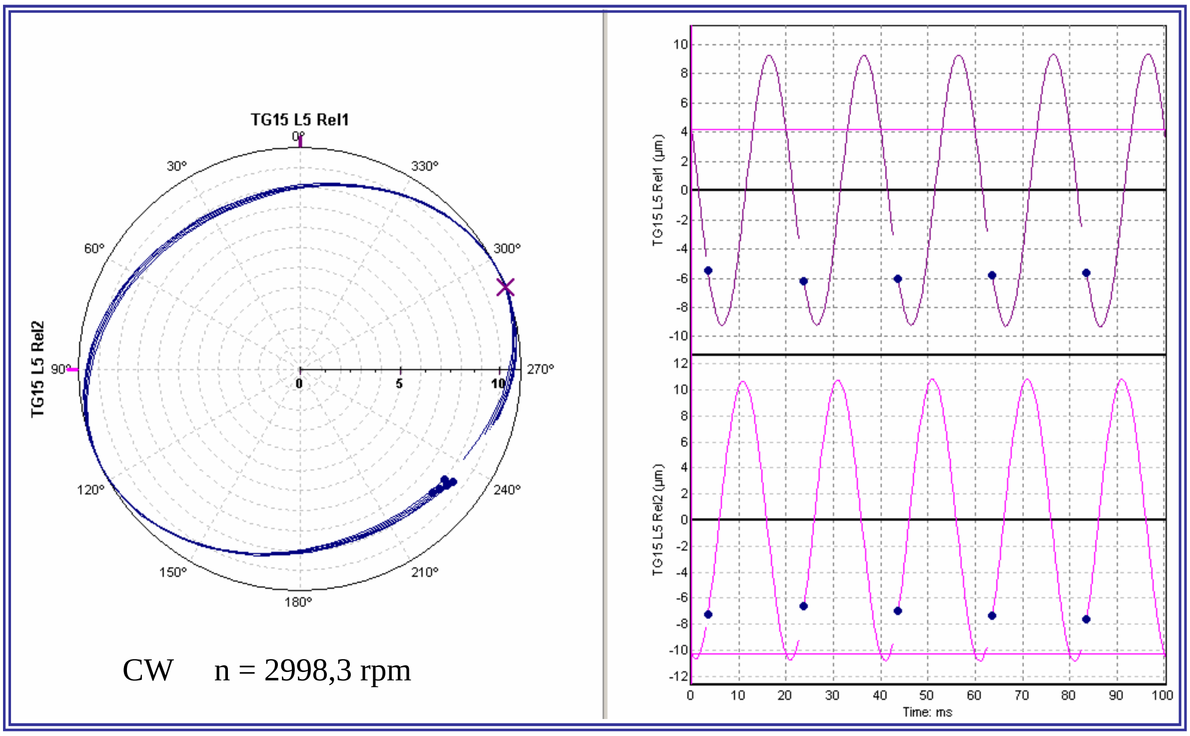


Fig. 2.16 - Orbit with the Keyphasor

3 VIBRATION ANALYSIS (FREQUENCY ANALYSIS)

Previous chapters have dealt mainly with vibration waveform. Evaluation of the waveform by comparing a single number (rms value) with the standard or by monitoring the development (trends) was described. There are methods for analysing the waveform as such and they are quite effective for some failures. However, it is not the most common way how to analyse vibration. More often, so called *frequency analysis* is used.

Basic consideration about more detailed analysis of vibration is shown in Fig. 3.1. Each waveform consists of contributions from individual vibrating parts, usually having different frequencies. Frequency analysis is a tool that is capable to identify these individual contributions directly. It can be compared with the music: an orchestra plays at correct volume (≅ vibration overall value is within normal limits), but some instrument may play out of tune (≅ some part of a machine has a problem). Even if it does not affect the overall intensity, in music it can be identified by hearing, especially if you know the composition. Likewise, the diagnostician can identify the machine fault on the base of frequency components that occur in the vibration *spectrum*. How to get the spectrum from the waveform is discussed in the following chapter.

Frequency analysis is performed by Fourier transform (by its decomposition into Fourier series). All procedures related to the frequency analysis, which will be discussed below, are implemented in analyzers that are used in vibration diagnostics. There are different types of analyzers - operational or laboratory, with one or more channels - but the principle of their operation is always the same.

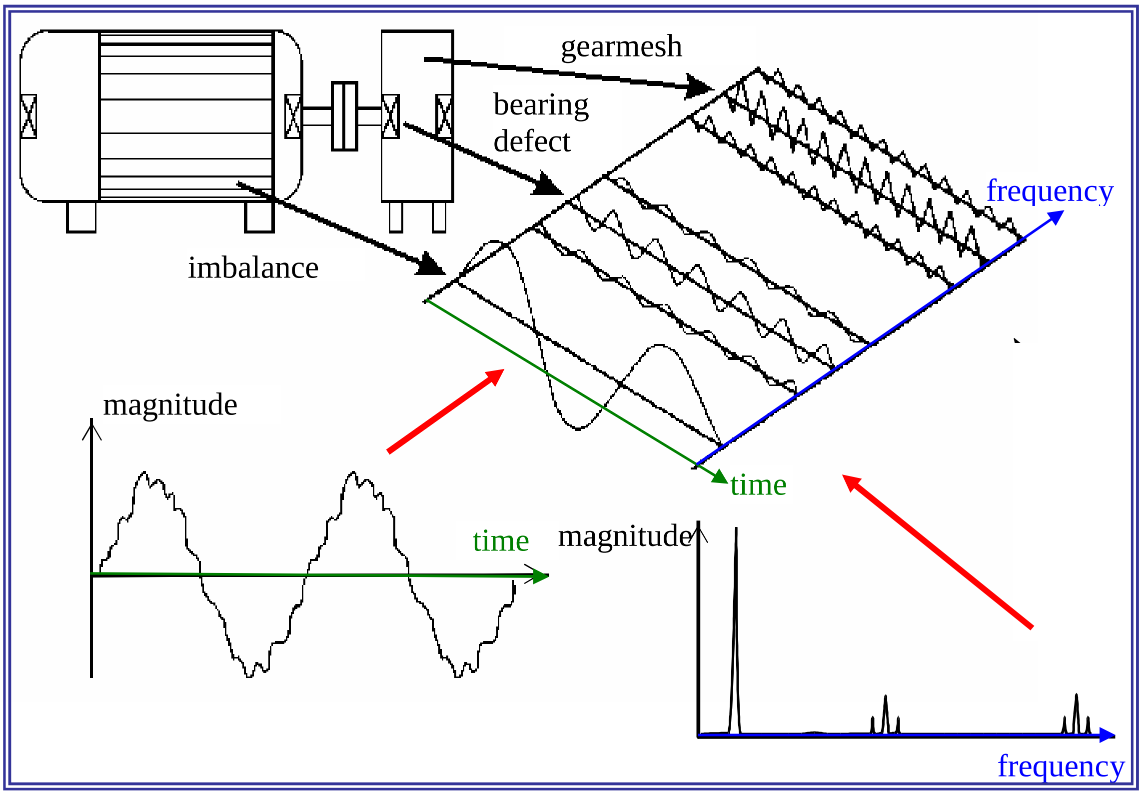


Fig. 3.1 - Principle of Frequency Analysis

3.1 Fourier Transform

Function **x(t)**, periodical in time **T**, can be expressed as an infinite progression:

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cdot \cos(n\omega t) + b_n \cdot \sin(n\omega t) \right] \qquad ; \qquad \omega = \frac{2\pi}{T} \qquad (3.1)$$

This expression means that the original function **x(t)** can be composed from (infinite) number of sinusoids of different amplitudes, frequencies of which are multiplies of the fundamental frequency **ω**.

Coefficients $\mathbf{a_n}$ and $\mathbf{b_n}$ are *Fourier* or *spectral* coefficients of the function **x(t)** and can be computed using expressions:

$$a_n = \frac{2}{T} \cdot \int_0^T x(t) \cdot \cos(n\omega t) dt \qquad ; \qquad b_n = \frac{2}{T} \cdot \int_0^T x(t) \cdot \sin(n\omega t) dt \qquad (3.2)$$

When working with measured vibration signals, a function is considered as periodical in the measured time interval **T**, though it is mostly not the case and additional adjustments of the signal are needed in order not to commit errors (see leakage, chapter 3.3). Current analyzers for processing measured signals do not work with a continuous waveform, but the measured signal passes to the analyzer through the A/D (analog/digital) converter, which records the waveform as a sequence of **N** discrete values with regular spacing in the time interval **T**. This procedure is called *discretization*. Discrete function **x (t)**, which is defined on the set of **N** different instants of time $\mathbf{t_k}$ (k = 1, N), can be written as a finite Fourier series:

$$x_k (= x(t_k)) = \frac{a_0}{2} + \sum_{n=1}^{N/2} \left(a_n \cdot \cos\left(\frac{2\pi n t_k}{T}\right) + b_n \cdot \sin\left(\frac{2\pi n t_k}{T}\right) \right) \quad ; \quad k = 1, N \qquad (3.3)$$

Fourier coefficients are often represented in the form of amplitude $\mathbf{c_n}$ and phase $\boldsymbol{\phi_n}$:

$$c_n (= X_n) = \sqrt{a_n^2 + b_n^2} \qquad \text{and} \qquad \phi_n = \operatorname{arctg}\left(-\frac{b_n}{a_n}\right) \qquad (3.4)$$

Then, the finite Fourier series can be written as:

$$x_k (= x(t_k)) = \frac{a_0}{2} + \sum_{n=1}^{N/2} \left(c_n \cdot \cos\left(\frac{2\pi n t_k}{T} + \varphi_n\right) \right) \qquad (3.5)$$

This form of Fourier transform is called the *discrete Fourier transform* (DFT). The resulting Fourier series, a set of sinusoids from which the original waveform can be composed, is called a *frequency spectrum*. By Fourier transform, the original information about vibration in the time domain, where the individual events are mixed, is transformed into the frequency domain, in which each physical phenomenon (imbalance, damaged teeth, etc.) is represented by a single sine wave of the corresponding frequency (i.e. by a *frequency* or *spectral* line).

There is a basic relationship between the length of the sample **T**, the number of discrete values **N**, *sampling* (or *capture*) *frequency* $\mathbf{f_s}$, frequency range and *spectral* (or *frequency*) *resolution*. Spectrum frequency range is 0–f_{max}, where $\mathbf{f_{max}}$ is the *Nyquist frequency* and **Δf** is the frequency resolution (spacing between frequency lines):

$$\Delta f = \frac{1}{T} = \frac{f_S}{N} \tag{3.6}$$

$$f_{max} = \frac{f_s}{2} = \frac{1}{2} \cdot \frac{N}{T} \tag{3.7}$$

An algorithm called Fast Fourier Transform (FFT) is used in up to date analyzers, where **N** is an integer power of number 2. In fact, the upper frequency spectral limit $\mathbf{f_{max}}$ is even more reduced in comparison with the theoretical value (e.g. for $N=2^{11}=2048$, only 800 frequency lines are used rather than 1024), which will be explained in the following chapter.

Digital Fourier Transform has many features which, if not properly treated, can lead to erroneous results. They occur in the train of a) discretization (aliasing error) and b) the need to reduce the length of the time signal, or, in other words, by an unrealistic assumption that the measured sample of the signal of the length **T** repeats periodically (leakage error).

3.2 Aliasing Error (Stroboscopic Effect)

The principle of the stroboscopic effect is clearly seen in Fig. 3.2. It can be seen for example in a movie in which the wheels (it is best seen on carriage wheels) rotate unrealistically slowly or even reversely with respect to the corresponding direction. It occurs when the sampling frequency is too small to realistically capture a fast action.

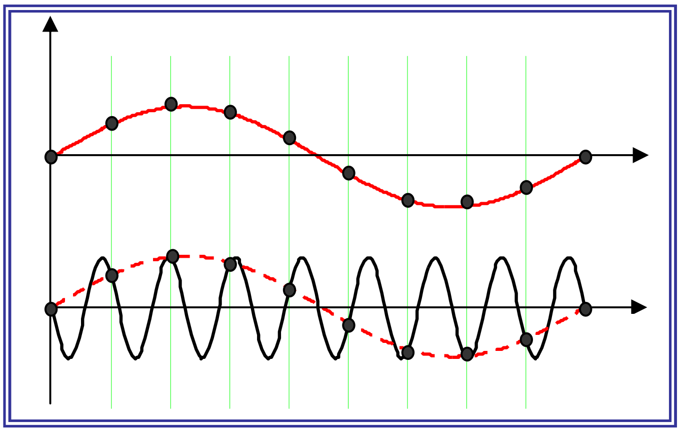


Fig. 3.2 - Principle of the Aliasing Error

If the sampling frequency is $\mathbf{f_s}$, then the signal of frequency **f** and the signal of frequency $(\mathbf{f_s}\text{-}\mathbf{f})$ are indistinguishable after discretization. The highest frequency $\mathbf{f_{max}}$, which may be contained in the spectrum, is $\mathbf{f_s/2}$. The part of the signal which has frequency components above $\mathbf{f_s/2}$ appears reflected in the range $\mathbf{0–f_s/2}$. These high frequency components, thus, act as if they were (*alias*) of low-frequency and they create an indistinguishable mixture with the real low-frequency components.

Solution to this problem is to use an anti-aliasing filter (a low pass filter with a steep falling edge), which removes the components higher than half the sampling frequency of the original signal. The filter is an analogue device that is an integral part of up to date analyzers. It should be inserted before the signal is discretized. Characteristics of the filter is steep, but not perfectly perpendicular, and therefore, the upper part of the spectrum is also removed (typically, the frequency range from $0.8 \cdot f_s/2$ to $f_s/2$ is removed). For this reason, the 2048-

point transform does not result in the full 1024-line spectrum, but only the first 800 lines are used. The requirement for the correct sampling frequency can also be expressed as:

$$f_s = 2.56 \times f_{max} \tag{2.8}$$

Fig. 3.3 shows four signals and their corresponding spectral lines when the sampling frequency is f_s = 6 Hz. Spectral lines appropriate to the actual frequency of the signal are displayed in blue, while red lines are lines that would appear in the spectrum after Fourier transform as a result of aliasing, if anti-aliasing filter is not used. From the bottom up:

- direct signal (DC)
- signal of frequency f = 2 Hz
- signal of frequency equal to the sampling frequency - this signal appears in the spectrum as if it was a DC signal
- signal of frequency f = 4 Hz, which is more than half the sampling frequency. It appears in the spectrum as a signal of frequency f_s-f (6-4 = 2 Hz).

Note: Frequencies in Fig. 3.3 are chosen in order to illustrate the sampling frequency of 6 Hz (vertical grid) and understand the process of aliasing. But it is clear that such a low sampling frequency is not the case of vibration measurements. Typically, sampling frequencies are in thousands Hz.

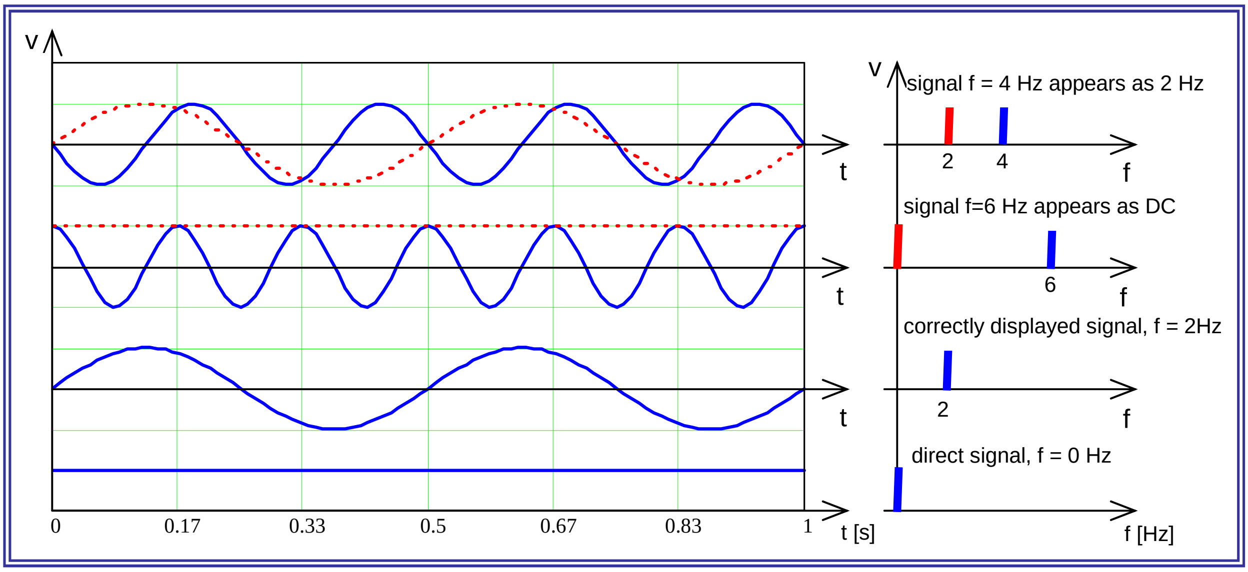


Fig. 3.3 - Manifestation of the Aliasing Error in the Spectrum

3.3 Leakage Error

Fourier transform assumes a periodic function. But real signals from vibration often do not have the period significantly marked, neither the course of the signal in different periods is exactly the same. Consequently, the spectrum of such a signal can be distorted (this distortion is called *leakage*), unless the signal is appropriately adjusted. To prevent a leakage error, weighting windows are used. The original signal is treated as if through a window of an appropriate shape (mathematically this means convolution of the original signal and the weighting function in the time domain and multiplication of the original signal and weighting function in the frequency domain).

Fig. 3.4 shows the principle of the leakage error and a window suitable for its suppression. If the signal entering the Fourier transform is periodic (i.e. periodic in the measurement time T, which in Fig. 3.4 left means that an integer number of sine waves enters the transform), there is no need to use any weighting window (in analyzers, RECTANGULAR window is set, which actually means no window). In case of a simple sine wave, the result of the transformation is a single spectral line, which corresponds to reality. Another situation occurs when the signal entering the Fourier transform is not periodic (Fig. 3.4 right). It can be seen that the signal is actually the same as in the case of the left, but the measuring time is different, and this produces a non-periodicity. In this case, the transform algorithm tries to model the resulting discontinuity, and the only way it can do this is by (infinitely) many other sine waves. Result of the transformation here is not the only spectral line with the corresponding frequency, but the energy "leaked" into many other spectral lines. It is obvious that such a spectrum does not testify much about the original signal. In addition, when there are more "spread" lines, important diagnostic information (e.g. a weak signal from rolling bearings) can be hidden in the increased threshold.

To prevent the leakage error, *Hanning window* is used when measuring steady signals. The point is that the signal was suppressed at their ends to zero, thereby removing the signal discontinuities and creating a signal that is closer to the real periodic signal. The result of using Hanning window is shown in Fig. 3.4 below. In the case where the signal was periodic, the result got worse (instead of one spectral line there are three lines), but in case of non-periodic signal the result is substantially improved - there are just a few spectral lines.

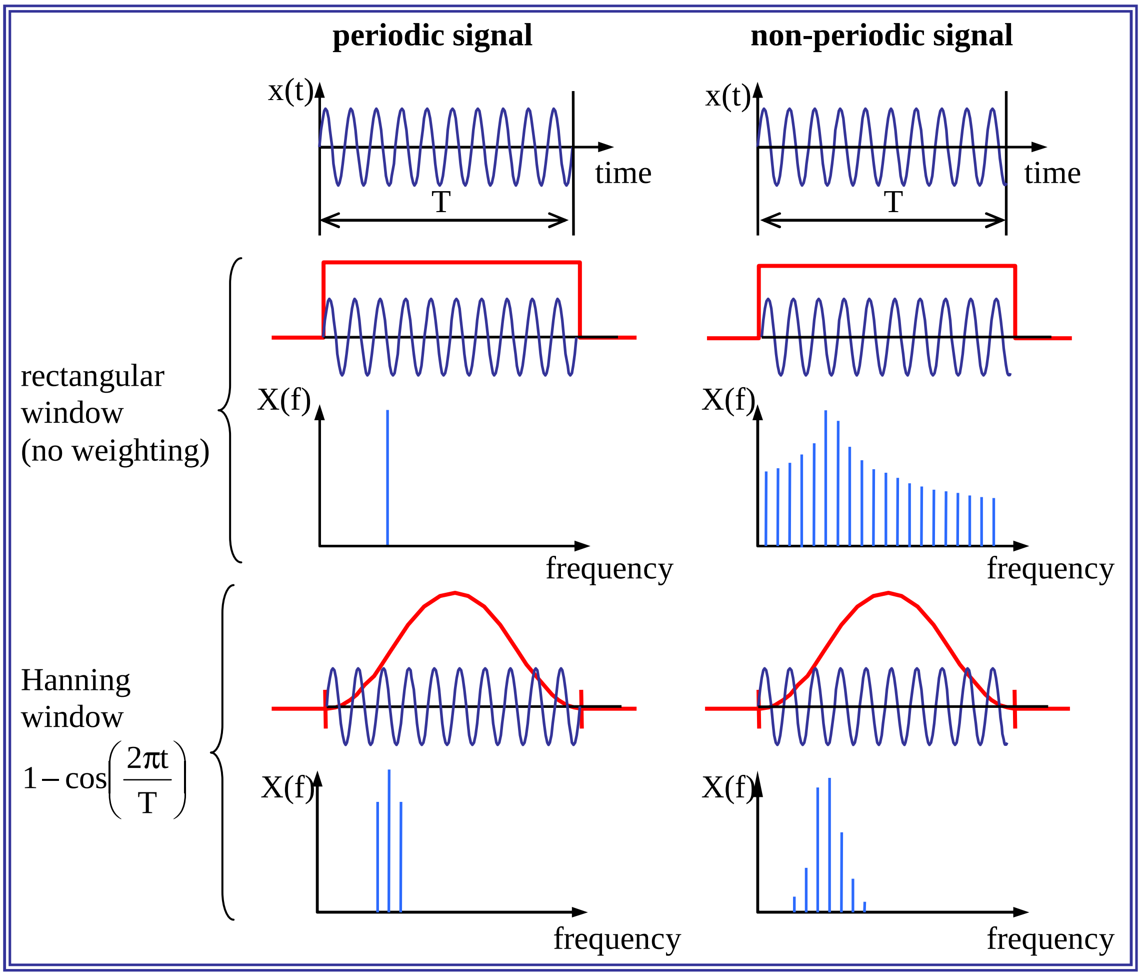


Fig. 3.4 - Influence of Signal Periodicity and Weighting Windows to the Leakage Error

In addition to the most often used Hanning and rectangular windows, other types of windows are used in diagnostics as well:

- FLAT TOP window - It is used for transducer calibration; it does not distort the signal magnitude (Hanning window would decrease the signal magnitude a bit, which does not matter in operational measurements).
- TRANSIENT window - It is a rectangular window but shorter in time than the measurement period T. It is used for shortly acting signals, e.g. for impact excitation signal.
- EXPONENTIAL window - It is used for measuring the response to impact excitation.

3.4 Setting the Analyzer

Fast Fourier Transform algorithm is implemented in all up to date vibration analyzers, eliminating the need of the user to have a good grasp of it. However, it is necessary to know basic principles to be able to avoid mistakes and to set efficiently transformation parameters and other measurement parameters, namely:

- frequency range (0 to f_{max}, or zoom f_{min} to f_{max} resp.)
- number of spectral lines
- number of averages
- averaging type
- overlap
- trigger type

The frequency range has been discussed earlier. Now, other parameters of FFT will be discussed. Remember that the measured data are sampled by A/D converter and then stored into a buffer, from which the data are taken by the FFT processor when needed (see analyzer scheme in Fig. 1.2).

3.4.1 Number of Spectral Lines

Setting the number of spectral lines affects the frequency resolution, i.e. resolution of any frequencies close to each other. It is preferable to consider the frequency lines as "columns" - everything that fits into the given column contributes to its size; but the individual frequency components cannot be distinguished within this column. In Fig. 3.5, the same frequency spectrum is displayed with finer (top) and coarse (bottom) frequency resolution.

It can be seen that with coarse frequency resolution, two distinct frequency components at 99 Hz (twice the rotational frequency) and 100 Hz (electrical forces) will not be displayed correctly. When deciding about the frequency resolution when setting the FFT parameters, it should be taken into account how close the frequencies in the spectrum could be in order to distinguish between them.

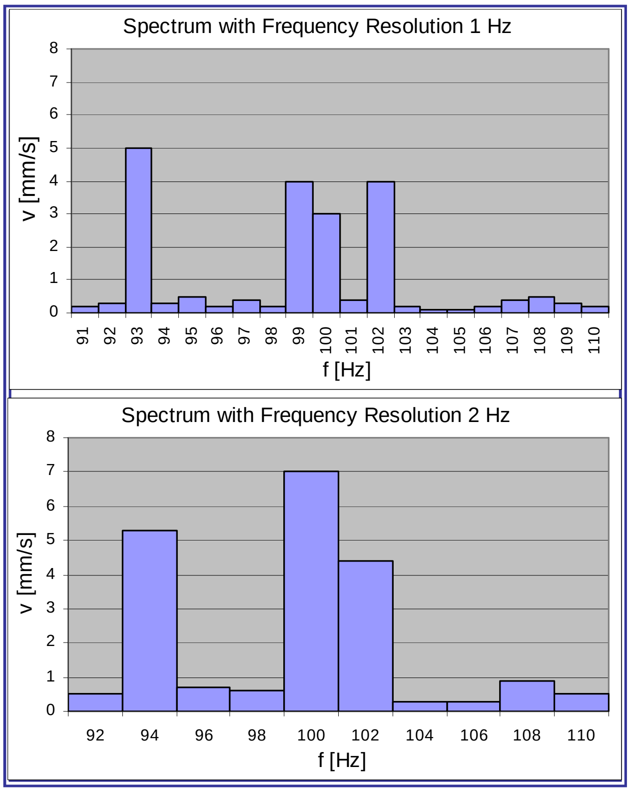


Fig. 3.5 - Effect of the Frequency Resolution to the Displayed Spectrum

3.4.2 Number of Averages and Averaging

Given that the measured vibration signal is not deterministic but its character is probabilistic, the signal should be averaged in order to base the diagnostic assessment on reliable data. Generally, the more volatile the signal is and the more contaminated with noise, the more samples of the signal should be used.

Most applications in vibration diagnostics use averaging *in the frequency domain*. The principle is that spectra obtained by Fourier transform of the measured time period of signal are averaged. The analyzer usually offers several types of such averaging:

- *linear (arithmetic)* where all measurements have the same weight
- *exponential* where new measurements are of greater weight
- *peak hold* – no averaging is performed, but maximum peak values of the individual spectral lines are stored.

Another parameter that is set in analyzer is the number of averages. Four to eight averages are typically set for steady-state, but it is advisable to set more averages at first and observe, since when the spectrum obtained is stable and unchanged. Depending on this, the number of averages should be set. For transient actions, e.g. run-up or coast-down, the peak-hold option and a large number of averages can be used to catch the entire transient process.

Further, the averaging can be set in analyzer either to

- CONTINUOUS where the measurement is still in progress and the actual spectrum is computed from the given number of averages; the older measurements are forgotten and replaced by newer ones; or to

- FINITE where, after starting the measurement, the finite number of averages is captured and the spectrum is computed and then the measurement is stopped.

The last parameter concerning averaging that is set in analyzer is the amount of OVERLAP. It is the way in which the samples of the continuous time signal are input to the analysis. If the time needed to process the signal sample in the analyzer is less than the length of the sample, averaging with the overlap is possible. Overlapping means that a part of the previous signal sample together with a new portion of the signal are taken for the FFT processing rather than the entire new signal sample. Historically, it led to time reduction of the averaged FFT computing (the processors were slow). Nowadays, the consideration shown in Fig. 3.6 prevails. It means that when Hanning Window is applied, the significant part of the signal at the beginning and end of the measurement period is omitted. Overlapping, thus, leads to a more credible use of the measured signal rather than decreasing the statistical relevance of the result. In Fig. 3.6, the actual signal is displayed in a weak line, Hanning window in a blue dashed line and the signal entering the transformation in a red line. Averaging without overlap is shown above, averaging with 50% overlap below. It can be seen that the signal that enters the transform is closer to the actual signal when overlapping is used. As to measurement time, when the length of one sample is T=1s, to use 6 averages would take 6 seconds without overlap and only 3½ seconds with 50% overlap.

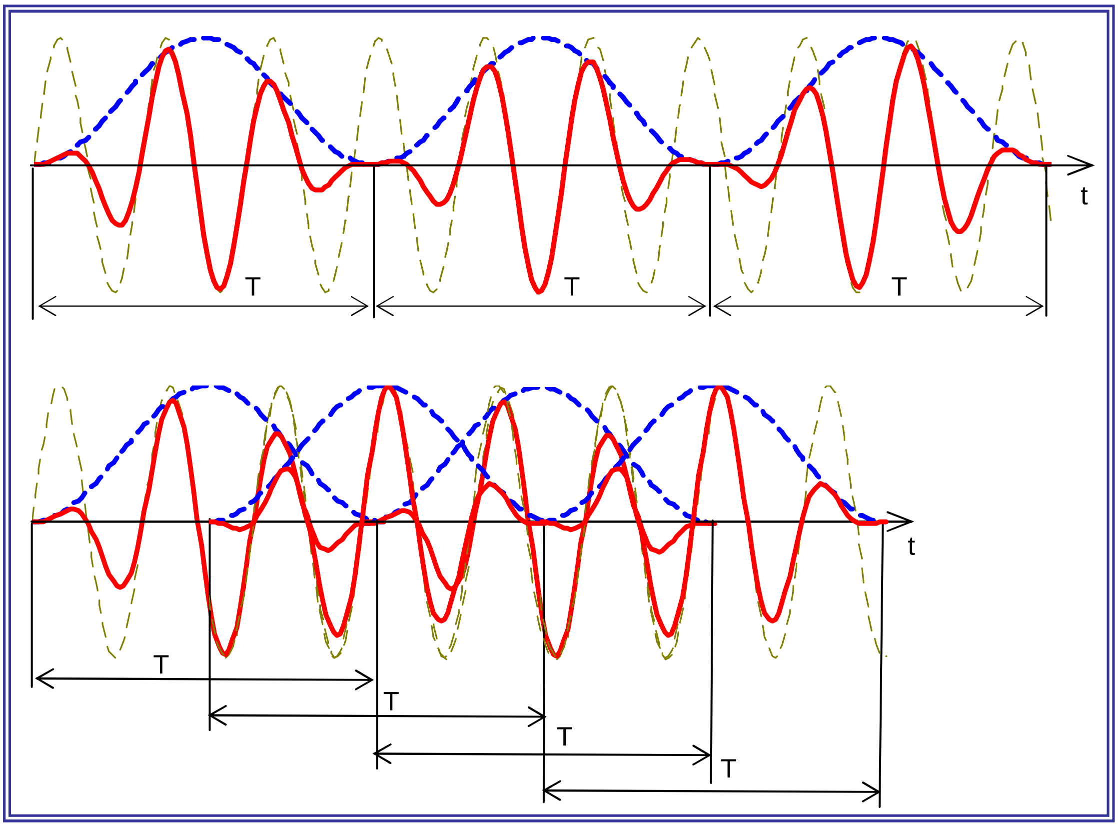


Fig. 3.6 - Averaging without Overlapping (above) and with 50 % Overlap (below)

Averaging can also be done in time domain, i.e. before FFT is performed. Chapter 3.4.4 deals with this type of averaging. To be able to deal with it, the way of starting measurements, so called triggering, should be explained at first.

3.4.3 Starting Measurements

When setting the parameters, the way how the analyzer recognizes that a signal sample should be captured should also be set. Three ways are possible:

- FREE RUN (without trigger) - The first sample of the signal is captured immediately after starting the measurement and the next one immediately after the previous one is processed (or depending on overlap setting).
- INTERNAL trigger - The course of the signal determines when capturing of a signal sample starts (see Fig. 3.7 above). The following information should be set:
  - level of the signal that determines starting a measurement (e.g. 10% of the maximum level)
  - if this level should be achieved on leading or declining edge of the signal
  - time delay. This means that when the required level of the signal is detected, signal sample that begins few milliseconds before this event is taken into processing (typically, TIME DELAY = -20 ms). This is enabled by the fact that the measured signal is continuously saved to the buffer.
- EXTERNAL trigger - Starting of capturing a signal sample is determined by some external event (see Fig. 3.7 below). It may be manually triggered, but most often the signal from the so called *key phasor* is used. In this case, there is a mark on the shaft that is sensed by a *phase reference probe* (e.g. by optical probe). Sensing this mark ensures that the measurement always starts in the same position of the shaft, namely in the very moment when the phase reference probe captured the mark.

It should be noted that the thoughts about overlapping during averaging as were presented in chapter 3.4.2 apply only to cases when the analyzer operates without trigger, i.e. when "free run" is set.

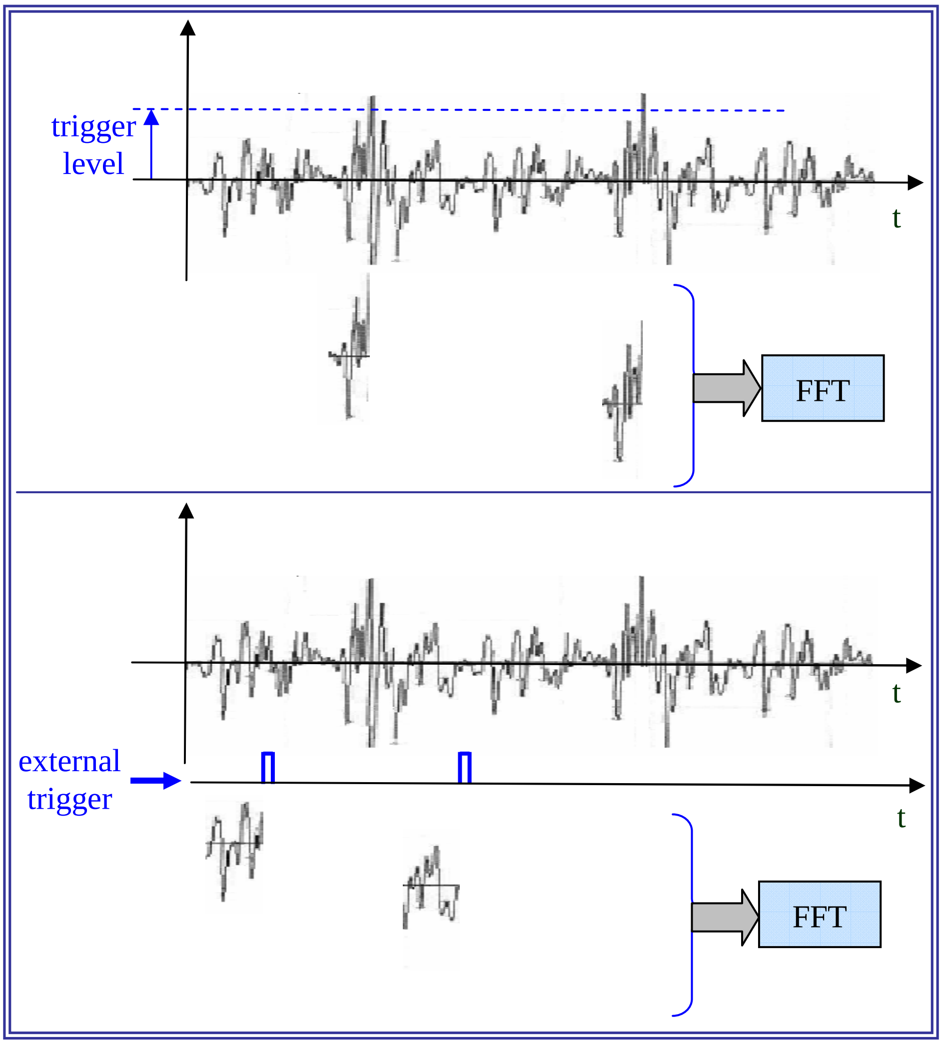


Fig. 3.7 - Internal and External Trigger

3.4.4 Time Synchronous Averaging

Apart from the common averaging in frequency domain, i.e. spectra averaging after Fourier transform of each signal sample is performed, averaging in time domain can be used as well. In this case, a final spectrum is computed after the individual samples are averaged in time domain. This procedure makes sense only when trigger (usually external) is used that is derived from the machine's motion (typically from shaft rotation in rotating machinery).

Great advantage of the time synchronous averaging is that when a sufficient number of averages is used (typically 50 to 100 averages), all non-synchronous parts of the signal disappear from the spectrum. In other words - everything that is not related to the rotation of the measured part of the machine, e.g. signals from adjacent vibrating machines or signals from machine part behind the gearbox, will trickle away. In some cases, this might be a drawback - when there are non-synchronous processes on the measured machine, e.g. self-excited vibration of a journal bearing, rotating stall in compressors etc. Therefore, time synchronous averaging should be used for the detailed analysis only and the spectrum obtained like this should be compared to the common spectrum obtained by spectral averaging.

3.5 Methods of Spectra Analysis

The aim of diagnostics is to assess the machinery condition and, based on deviations from the expected state, to presume emerging machine faults. For this purpose, spectra obtained by Fourier transform of the time record of vibration are commonly used. In addition to vibration measurements and monitoring of the measured values, additional information about the machine is necessary:

- basic technical data (power, operational rotational speed, number of blades, number of teeth ...)
- operation history (downtimes, operating modes ...)
- maintenance history (routine repairs, overhauls, lubrication ...)
- faults in the past

3.5.1 Significant Frequencies

To be able to read various information about a machine from a spectrum, it is worth to know which frequencies would likely occur in such a spectrum (rotational, gear-mesh, blade-pass, from bearings, etc.). The spectrum usually contains a number of discrete lines and sometimes areas of the increased noise. It is appropriate to divide the spectrum into areas, in which various symptoms occur:

- Usually, the spectra evaluation process starts with identification of the frequency pertinent to the rotational speed of the shaft - rotational frequency for which the notation **1X** is used.
- Further, the integral multiplies of this rotational frequency (2X, 3X ...) are identified - they are also called *harmonics*.
- Spectrum is divided into three main areas (see Fig. 3.8):
  - area below rotational frequency - This area is called *subsynchronous* and if any peaks occur in it, they tend to be dangerous (e.g. journal bearing oil whirl).
  - area from rotational frequency up to ten times of it - the area of low-frequency events related to the rotation. Symptoms of all the basic mechanical faults (unbalance, misalignment, looseness, etc.) usually occur in this area.
  - area above 10X - the area of high-frequency events. Symptoms of roller bearing defects, gear faults, etc. occur in this area.

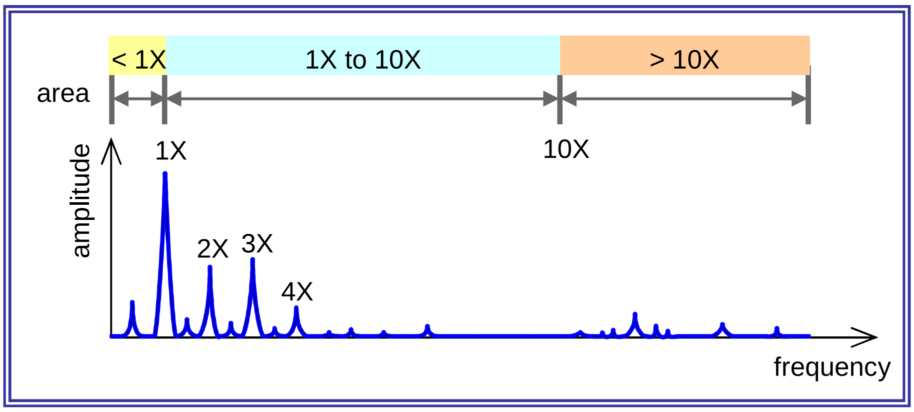


Fig.3.8 - Dividing the Spectrum into Significant Areas

3.5.2 Reference Spectrum and Monitoring of Changes

Common way of evaluating spectra is that at first the machine in good state is measured, i.e. when the machine is new or after a repair. *Reference (baseline) spectra* are measured in all the measurement points that should be monitored. Then, during subsequent measurements, the same points are measured and the measured spectra are compared to the baseline spectra.

In Fig. 3.9 above, there is an example of a baseline spectrum, beneath there is a situation when the basic rotational frequency 1X has changed and is significantly higher than normal. This indicates that the vibration signal is periodically changed once per shaft rotation. A typical cause for such behaviour use is either unbalance or misalignment. The last spectrum below shows that in addition to increased 1X, peaks from roller bearing defects have increased as well, which indicates that the problem on 1X gave rise to bearing damage.

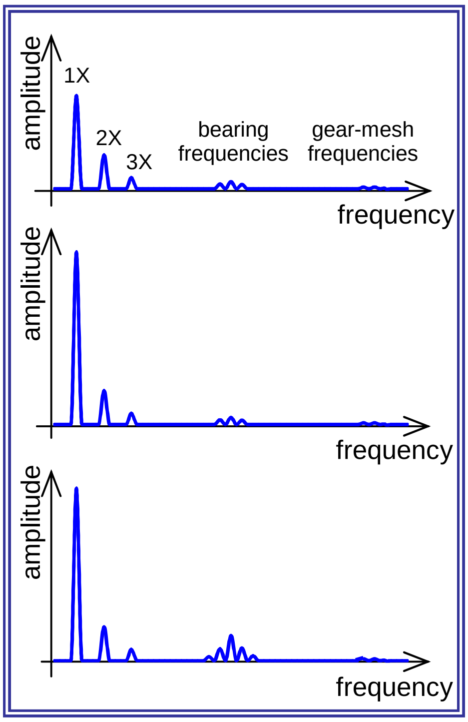


Fig. 3.9 -Baseline Spectrum and Changes in Spectra while Fault Development

3.5.3 Special Types of Cursors

To facilitate the work with the spectra, software for processing the measured data offer various diagnostic tools (e.g. special types of cursors) and different ways how to display the measured spectra.

Harmonic cursor can be used to identify harmonic multiples. It marks all multiples of the frequency that is designated by the basic cursor. Harmonic cursor is mostly used to indicate multiples of rotational frequency (see Fig. 3.10 above).

In the case that two components with different frequencies are contained in the vibration signal which originate from different sources (for example, rotation frequency and frequency of bearing defects), their modulation takes place. Modulation in a spectrum appears as a series of side bands (see Fig. 3.10 below) which can be determined using the SIDEBAND cursor.

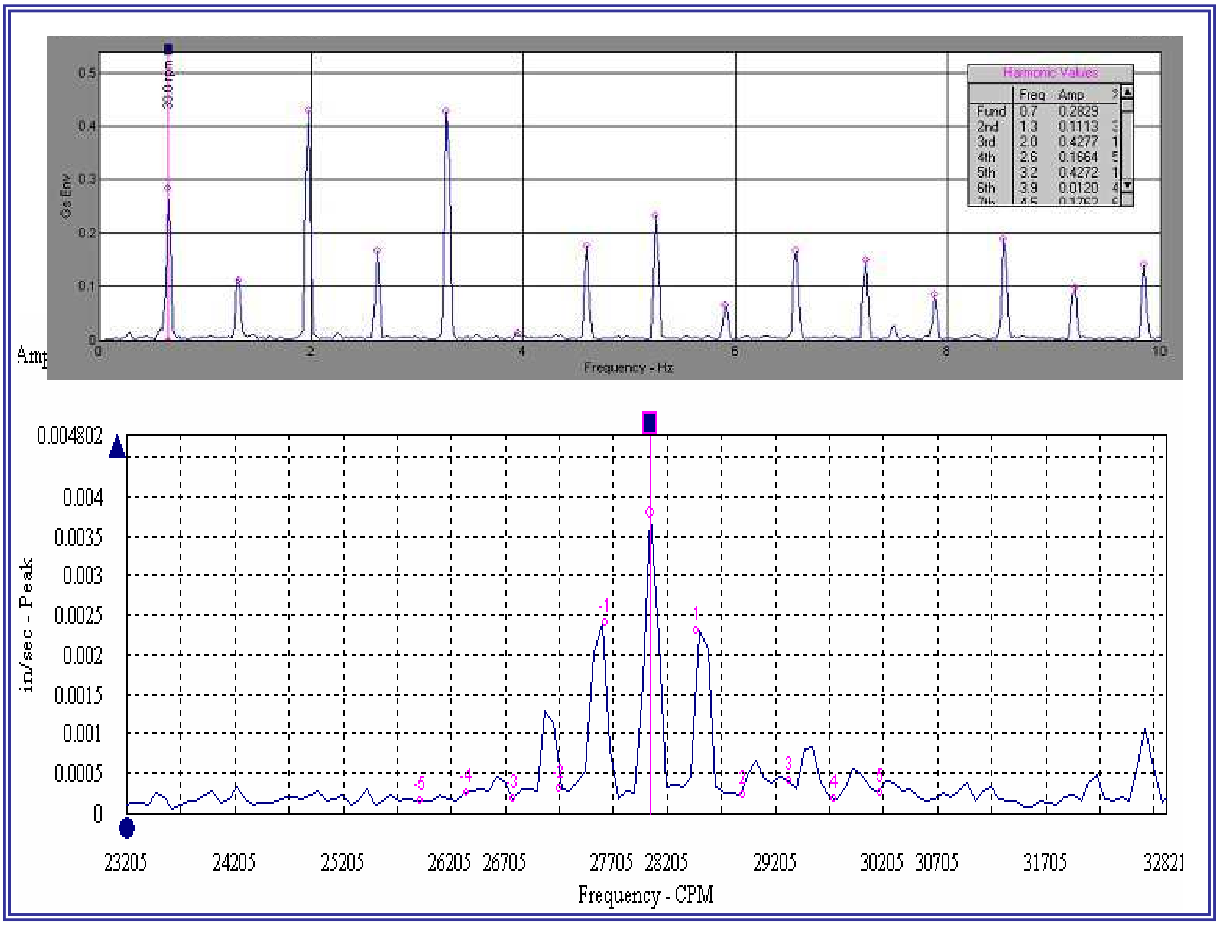


Fig. 3.10 -Example of Using Harmonic Cursor (above) and Sideband Cursor (below)

3.5.4 Waterfall Diagrams

Another diagnostic tool for quick visual comparison of spectra is *waterfall (cascade) diagrams*. They are used in two ways:

- To compare spectra from steady operation during a longer period of time (see Fig. 3.11 above). Gradual or sudden change in some frequency components is clearly visible.
- To compare spectra during transient process - usually during a run-up or coast-down of the machine or during a significant change in load (see Fig. 3.12 below). Resonant frequencies of the machinery can be detected in this way (especially those of a rotor).

3.5.5 Phase

As described in previous chapters, mechanical vibration can be described using frequencies and amplitudes at these frequencies, always in a given measurement point. Now, another variable will be introduced, which allows an assessment of the mutual movement of different points in time. This quantity is called the *phase*. It is a very important quantity, which is mainly used for detailed analysis of the detected problem, i.e. when actual diagnostic considerations take place.

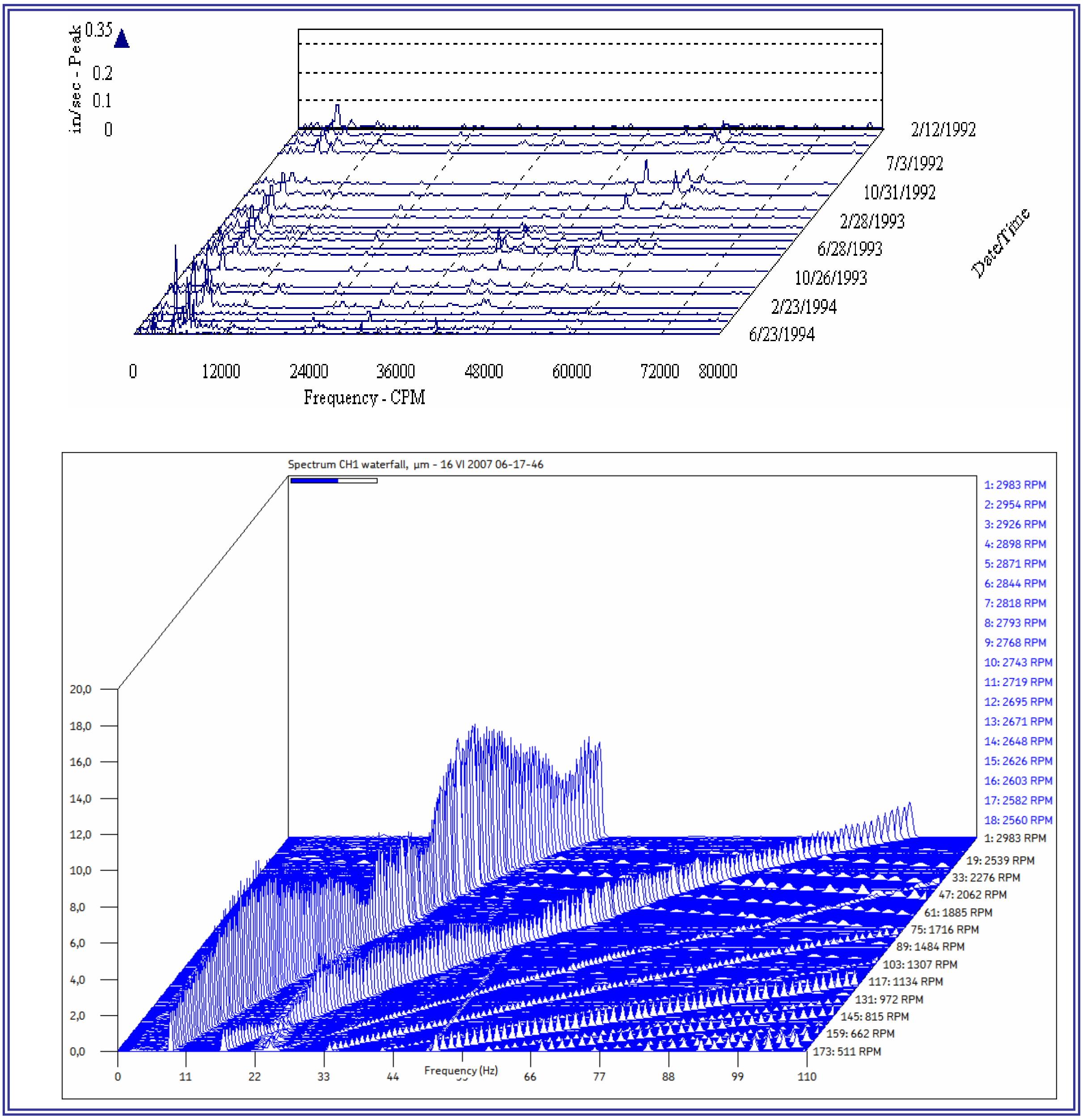


Fig. 3.11 - Waterfall Diagrams (above steady operation, below coast-down)

However, phase in diagnostics has different meaning from those used in dynamics. In dynamics, the phase means the *phase delay* between displacement and the excitation force (remember the amplitude-phase characteristics of forced oscillation).

In diagnostics, phase is determined with respect to some external signal (trigger) and it has no significance in itself. Diagnostically significant is the difference between two phases - either between two different points on the machine or between the same points measured at different operating conditions.

In order to assess the relative motion between the individual points, the best solution is to add the reference mark, to which the information about the phase shift will be related. This mark is the same for all measured points. It was discussed already in chapter 3.4.3 dealing with the external trigger.

Thus, the *phase shift* (shortly *phase*) indicates in which part of the cycle from capturing the trigger signal the level of vibration reaches maximum in the given measured point (see Fig. 3.12); it is given in degrees.

Note: In some special cases when such external signal is not available, signal from one of the measured points, in which the course of the signal is sufficiently good (without noise or harmonics interference), could serve as trigger.

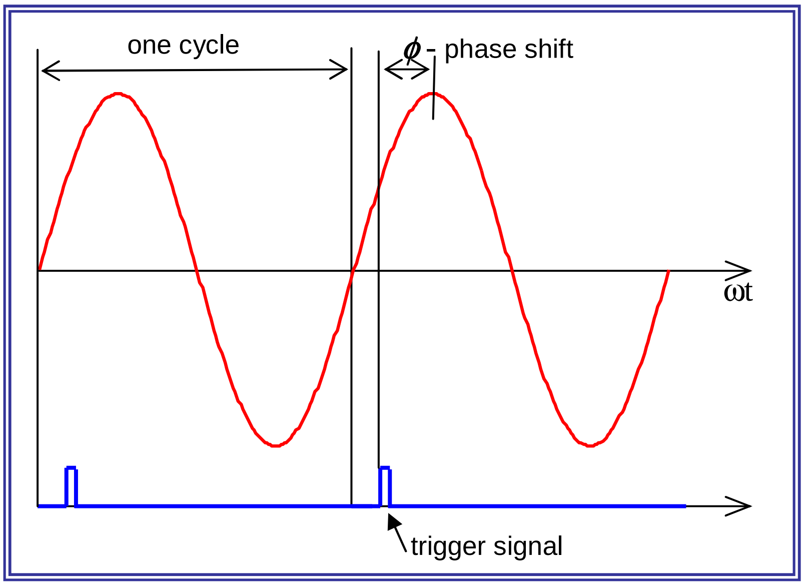


Fig. 3.12 - Depicting of Phase Shift on the Waveform

A common method of obtaining the reference signal is the use of *optical probe*, which is a device that transmits a light beam and receives the reflected one that is converted into an electrical impulse. This approach requires placing of a reflective mark on the machine shaft. Thus, the machine should be stopped and the mark (strip of a reflective tape) should be placed prior to performing measurements. Usually, it cannot be assumed that the mark will last long because it is gradually clogged with dust or oil (usually both).

Optical probes are manufactured in various designs and prices. Cheaper ones use an infrared beam and are limited by a small range (sometimes only approximately 0.2 m), lower sensitivity and low resistance to interfering effects. More expensive laser probes have substantially longer range (several meters), which is advantageous especially for not easily accessible machines.

Note: Optical probe should be aligned before each measurement. As a result of inaccuracy of alignment, it is unrealistic to expect good reliability of such a phase measurement for monitoring purposes (detection of deviations). Therefore, this method is more suitable for single diagnostic measurement. This method is often used for field balancing.

Another way to obtain a reference signal is to use *non-contact proximity probe* based on eddy currents, which is normally used to measure shaft vibration, in combination with a groove or protrusion on the shaft. Usually, this probe is permanently installed on the machine, which somewhat raises the cost. On the other hand, this method is the most suitable for monitoring purposes, as the measurement is not affected with random error caused by repeated attaching of the probe, as is the previous case.

Phase information is often displayed as a clock hand on the clock face in machine schemes, which is often more illustrative than its number representation. A machine scheme showing four measurement points and the measured values is shown in Fig. 3.13. Vibration velocity values in mm/s are placed in circles, phases values are indicated with short lines. This example indicates that the static unbalance was diagnosed on the machine (see Chap. 4.1.3) - phase values for both bearings are the same, horizontal and vertical direction has a phase difference of approximately 90º.

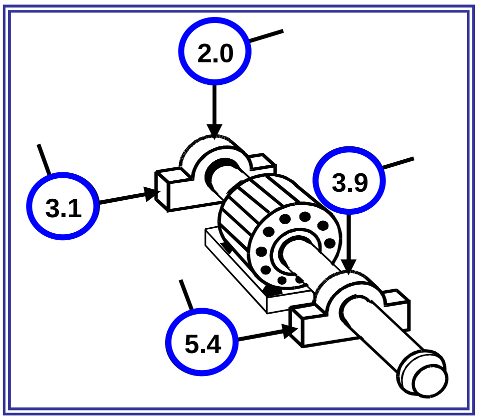


Fig. 3.13 - Measured Vibration Values Including Phase Indication

The use of phase information significantly helps to diagnose different types of machine faults. Individual faults including phase relations will be discussed in detail in Chapter 4.

3.5.6 Applying Spectra and Phase Analysis to the Diagnostics of Machine Faults

If vibration exceeds the preset alarm limits for overall vibration or for spectral bands during condition monitoring, it is necessary to diagnose the arising problems from these data. A lot of valuable information is contained in the vibration spectra or phase relations in different parts of the machine. In many cases, it is appropriate to assess the shape of signal waveform as well.

Various machine faults or problems will be described in Chapter 4 together with the symptoms that enable to diagnose them from the measured vibration data. Examples of typical vibration spectra will be introduced for all the mentioned problems and phase relations on various parts of a machine for some of them.

A *typical spectrum* shows important information about the problem source. While evaluating spectra, the following questions are worth to answer:

1. Which frequencies are present in the spectrum and how are they related to the rotational speed of the machine? (This means: Are the peaks 1X, 2X, 3X ... 5.28X or which ones present in the spectrum?)
2. What are the amplitudes of these peaks?
3. What is the relationship between the individual peaks? For instance: Is the 2X much greater than 1X? Do high 6.43X peak exist? Are there a lot of harmonics of rotational speed? Are there high sidebands along the gear-mesh frequency?, etc.

4. What is the <u>source of peaks</u> with significant amplitudes? For instance: Is the 6.43X peak a bearing defect frequency? Is the 7X peak in accordance with the number of the pump blades?, etc.

As the designation *typical spectrum* suggests, it is a representative spectrum with respect to the problem. Anyway, these spectra do not include everything. Although the spectrum characteristic for the problem exhibits high values of some peaks, it is not unusual for any of these peaks to be strongly dominating. In addition, it is not unusual that the machine has two or more problems simultaneously. For example, if a machine has a mechanical looseness problem and a rotor unbalance problem at the same time, then each of the problems is involved in the frequencies in the spectrum thus leading to both high value of 1X and its harmonics.

Generally, it can be said:

- *Amplitude* indicates *how much* a machine vibrating is.
- *Frequency* indicates the *number of vibration cycles per time unit*. (Taking time of one revolution as the time unit, it leads to *nX* multiplies.)
- *Phase* completes the illustration with *how* the machine vibrates.

Phase is of great importance and it is a powerful tool that allows to distinguish which of several possible machine problems dominates. For example, many problems generate frequencies 1X and 2X. Using of phase indicates how the individual parts of the machine vibrate one to each other, thus enabling to focus on the real, existing problem.

Unless the phase measurement is used, diagnostic possibilities are strongly limited. However, it is not practical to measure phase on all machines in the scope of regular monitoring. Utilizing phase measurements is important especially for diagnostics of machines on which high vibration has developed on 1X, 2X or 3X and thus it is clear that the machine has a problem. The analysis of phase relations enables to diagnose the dominant cause of the problem before the corrective action is performed. Typical examples of the use of phase information are as follows:

- Phase measurements enable to distinguish between the individual types of unbalanced rotor and thus choose the appropriate balancing method.
- Comparing phases around the coupling enables to find out whether there is a misalignment during operation and which type of the misalignment is dominant.
- Comparing phases on the individual parts of machine and the pedestal enables to indicate the presumable location of the mechanical looseness.
- Change in phase during change of rotation speed indicates the presence of resonance or critical speed respectively.

For more complex cases, it is often worth to assess also the signal waveform, which often helps to identify the problem more exactly.

4 DIAGNOSTICS OF COMMON ROTATING MACHINERY FAULTS

In previous chapters, tools for processing and understanding the measured data were discussed. This chapter is focused on describing common faults that are diagnosed in rotating machinery and on describing their relevant symptoms and possible cure.

When evaluating the condition of rotating machinery using vibration analysis, it is important to consider their substantive characteristics, both structural and operational. These characteristics can be divided into several groups:

1. The first group of characteristics is related to the fact that these machines are mechanical devices and that mechanical forces associated with rotation, especially **centrifugal force**, are acting on them. Response on rotational frequency and possibly on its multiples corresponds to this force. Considered that no machine is operating alone, but it is always a machine set consisting of at least one driving and one driven machine, forces generated by the connection of these machines (usually by a coupling) and corresponding responses are also involved. Other forces and specific responses may be associated with contact between a rotor and a stator, etc.
2. The second group of characteristics is related to the fact that each machine has its **operational function** to which operating forces and their specific responses correspond. For example, operating forces acting in the pumps or in electrical machines are quite different and the corresponding frequency components in spectra are different as well.
3. The third group of characteristics is associated with **specific components** of individual machines. Forces acting in journal or rolling bearings are different, different forces are characteristic for gears or for belt drives. Typical responses for different types of forces differ as well.

The above mentioned characteristics and responses can be greatly affected by **machine support**. Also, retuning as a consequence of decreased stiffness of the attachment, caused e.g. by mechanical looseness or by a crack, can modify mechanical characteristics of machines significantly.

The above mentioned characteristics may be related. A large centrifugal force from unbalance can cause an excessive bearing load, and this will result into the vibration response of bearings; excessively worn teeth can cause significant excitation of machine shaft natural frequencies, etc. In particular, change in operating conditions - change in load and thus in corresponding operating forces and their responses - can significantly affect both the mechanical characteristics, particularly the effect of connection of rotors, and specific characteristics such as occurrence of self-excited vibration of rotors in journal bearings in certain operating modes.

In all the above mentioned groups of characteristics it should be considered that non-linear characteristics that are practically negligible in the new machine may be manifested with the deteriorating state of the machine.

As this text is intended for mechanical engineering students, it will be focused on the first and partly on the third group of characteristics, and machine bedding will be taken into account. The influence of operating conditions strongly depends on the machine type and it will be mentioned only marginally.

4.1 Unbalance

The centrifugal force is one of the basic excitation forces in rotating machinery. It originates from the fact that with a real product (rotor), it is not possible to achieve the centre of gravity to be exactly on the axis of rotation and this axis of rotation to coincide with the principal axis of inertia (see Figure 4.1.1). The causes of this condition are:

- in design (some parts may be not perfectly symmetrical)
- technological (non-homogenous material)
- manufacturing (everything is produced in some tolerances, rotating parts exhibit runout)
- in mounting (namely with mounted rotors)

Since the centrifugal force is mostly harmful and causes increased load of the individual parts of rotor, bearings and supporting structures, various actions for its minimization are carried out. Rotors are balanced - either in the manufacturer's plant or in some service organization on balancing machines (see Chap. 4.1.2) or in field (see Chap. 4.1.4).

Note: There are some special exceptions when the centrifugal force is desirable and used in the manufacturing process, but these are not of interest of this text.

4.1.1 Types of Unbalance

There are three types of unbalances depending on how the mass is distributed on the rotor and how it will affect the position of the principal axis of inertia with respect to the axis of rotation: static, couple and dynamic unbalance. Various ways of balancing are derived for these various types of unbalance. Static unbalance is the easiest to eliminate (balancing in one plane is sufficient, see chapter 4.1.4.5.1). In practice, almost always a dynamic unbalance is the case, yet even in these cases methods for solving static unbalance are often used because of their simplicity - they are not able to eliminate the problem, but still can reduce its severity.

4.1.1.1 Static (Force) Unbalance

A rotor is statically unbalanced when the principal axis of inertia **t** (red) is parallel to the axis of rotation **o** (blue) (Figure 4.1.1 above). Static unbalance can be understood as two equal masses (*unbalances*) $\mathbf{m_n}$, placed symmetrically in relation to the plane perpendicular to the axis of rotation and passing through the centre of gravity. This unbalance is called static because it manifests itself even when the rotor is not rotating - the rotor tends to take the rest position with the centre of gravity below the axis of rotation (the wheel of the bicycle would take its position with the valve at the bottom). To eliminate the static unbalance, a mass should be simply added or removed in a single trim plane, so that the centre of gravity is shifted back to the axis of rotation.

It should be noted that the static unbalance shown in Figure 4.1.1 above is rather theoretical case. To remove it by placing a single balance mass, this mass should be put to a plane perpendicular to the axis of rotation and passing through the centre of gravity. In practice, the unbalance is considered as static when the shaft diameter **d** is much larger than its length **l** (the ratio $d / l \geq 10/1$ - so called short shafts) or when a single disc is placed on a long shaft. Then, of course, the unbalance results mainly from this disc and a balancing mass is added also to this disc. In other cases, the shaft unbalance is almost always dynamic.

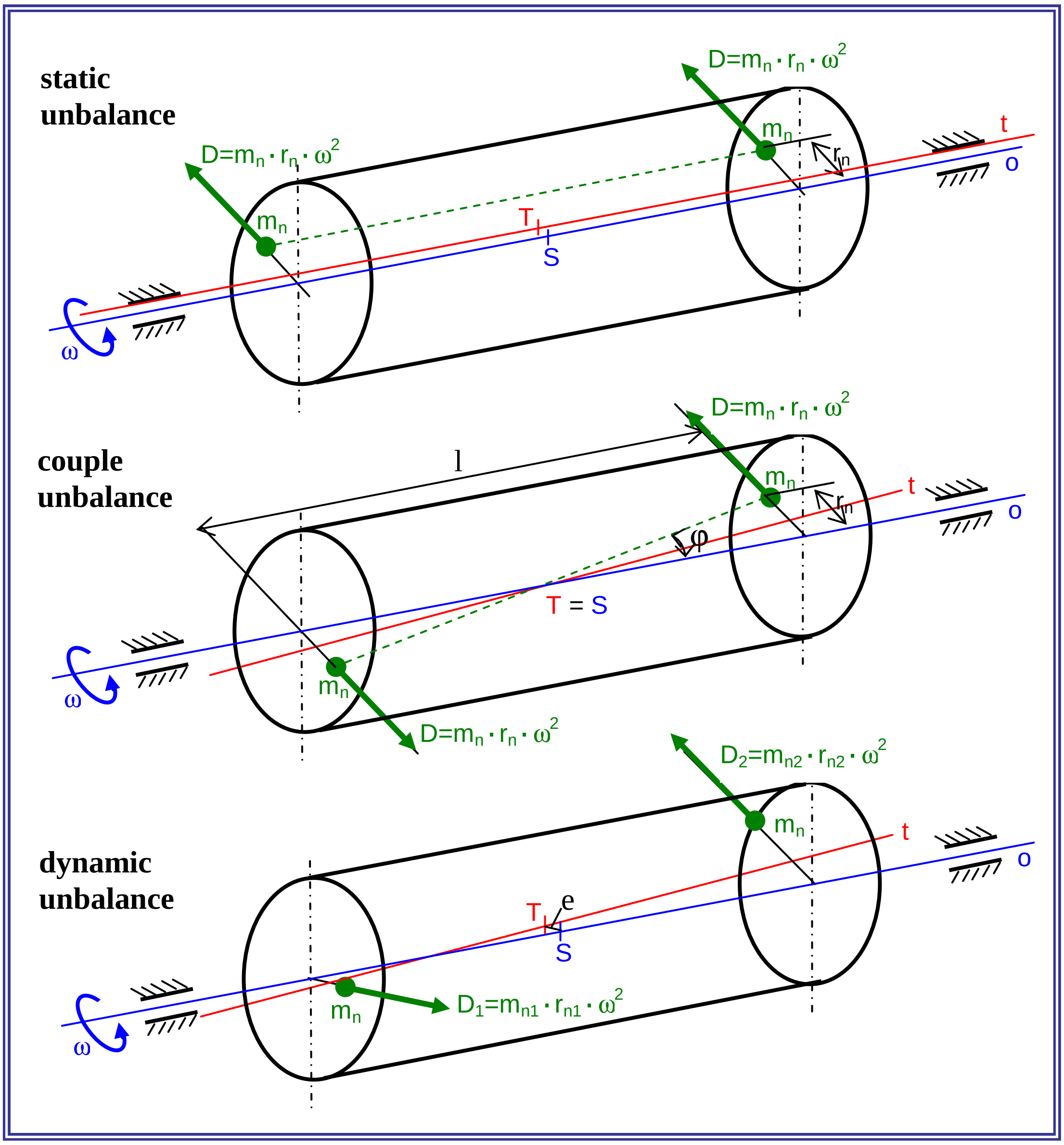


Fig. 4.1.1 - Static, Couple and Dynamic Unbalance

4.1.1.2 Couple Unbalance

The couple unbalance means that centre of gravity of the rotor is on the axis of rotation and the principal axis of inertia is intersecting with the axis of rotation. This unbalance can be understood as two equal masses $\mathbf{m_n}$ placed on the radius $\mathbf{r_n}$, placed symmetrically (Fig. 4.1.1 middle). If a rotor with couple unbalance is at rest, the unbalance has no effect and the rotor will remain at rest in any position. Couple unbalance will not manifest itself unless the rotor is rotating, as the centrifugal forces from unbalanced masses created couple that deflects the principal axis of inertia.

4.1.1.3 Dynamic (General) Unbalance

Dynamic unbalance is the most common type of unbalance and is essentially a combination of static and couple unbalances. It occurs when the principal axis of inertia and the axis of rotation are skew lines. This unbalance can be understood as two different masses $\mathbf{m_{n1}}$ and $\mathbf{m_{n2}}$, placed arbitrary on the rotor (Figure 4.1.1 below). Two centrifugal forces $\mathbf{D_1}$

and $\mathbf{D_2}$ arise due to these masses during rotation. Their effect can be compensated by two weights that are added or removed in two trim planes. The size and position of balancing weights on the rotor are determined by the balancing procedure described in Chap. 4.1.2.

4.1.2 Balancing on Balancing Machines

An important measure to minimize the unbalance of a new rotor (or a rotor after an overhaul) is balancing on balancing machines, which is performed by the manufacturer before delivery to the user. Either the entire rotor is balanced or its individual parts before the final assembly and then a final balance is carried out.

Balancing methods are highly variable depending on whether the rotor runs well below the first critical speed during operation, or it operates above the first and often even above the second critical speed. (Note: The critical speed is approximately equal to the bending natural frequency of the shaft.) Following terminology is used:

- Rigid rotor balancing = below critical speed
- Flexible rotor balancing = above critical speed

This distinction is very important and its principle is that:

- *Rigid rotor* is not deformed during rotation - its axis remains unchanging up to a maximum operating speed.
- *Flexible rotor* bends and rotates around its deflection curve (see Fig 4.1.5) and the shape of this curve strongly depends on the instantaneous operating speed.

This issue is described in more detail in the standard ISO 1925 *Mechanical vibration - Balancing - Vocabulary*. From the above mentioned distinction, following conclusions arise:

- If the rotor is rigid, it can be balanced even at low speed, e.g. from 600 to 800 rpm (there is no need for balancing at the operating speed). This seems to be insignificant, but it is not just about saving energy to spin the rotor for balancing. This is particularly about the fact that most rotors have some protrusions from the smooth surface - working blades, cooling fans blades, various pins, etc., which cause air turbulence with increasing rotational speed. When balancing at lower speed than the operating speed is, there will be less windage, the power needed for rotation and vibration generated will be smaller, and last but not least, working conditions and occupational safety will be better. Such balancing is significantly cheaper (example: CZK 50 000 for a smaller type of turbine rotor).
- If the rotor is flexible, it must be balanced at the actual operating speed (when the deflection shape has its real shape, influenced by internal unbalanced moments). Due to the above mentioned ventilation problems, most of these rotors have to be balanced in an environment with reduced air pressure. A balancing machine is placed in a concrete bunker called balancing tunnel which is equipped with several vacuum pumps. As balancing is an iterative process, the air should be pumped off several times from this large space. Not only for this reason, such balancing is significantly more expensive (about CZK 500 000 for a turbine rotor). This may lead to an effort for savings, particularly during repairs, and sometimes such a rotor is balanced at low speed (for little money), which may result in large problems with vibration during operation.

The issue of balancing, which is often crucial in terms of smooth operation for rotating machinery, is described in series of standards and special publications (see list of references).

Basic principle of two-plane rigid rotors balancing (according to [14]) is shown in Fig. 4.1.2. With regard to the construction of the rotor, suitable planes labelled I and II - so called *balancing planes* - should be selected for placement of balancing weights marked $\mathbf{Q_I}$ and $\mathbf{Q_{II}}$. Each of them affects vibration in both places **A** and **B** where vibration is measured. This effect is assessed by influence coefficients $\boldsymbol{\alpha_{ij}}$, where the first index indicates the plane of measurement and the second index indicates the location of the plane where a balancing weight is placed:

$$X_A = \alpha_{AI} \cdot Q_I + \alpha_{AII} \cdot Q_{II} \qquad (4.1a)$$

$$X_B = \alpha_{BI} \cdot Q_I + \alpha_{BII} \cdot Q_{II} \qquad (4.1b)$$

where X_A , X_B ... vibration in the measurement planes

Balancing machines are special machines designed so that influence coefficients (which apply permanently) are known and their values are inserted directly into the computer program so that both size and location of balancing weights can be set after single measurement of vibration at points **A** and **B**. A typical well-known example is a balancing machine used for car wheels balancing. Occasionally (usually due to non-linearities), further trim runs are required to achieve the required balance quality.

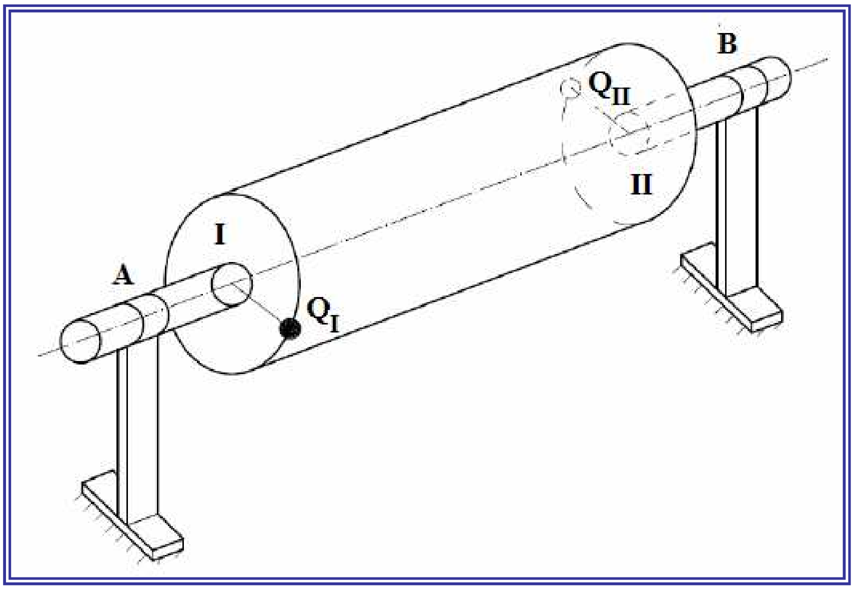
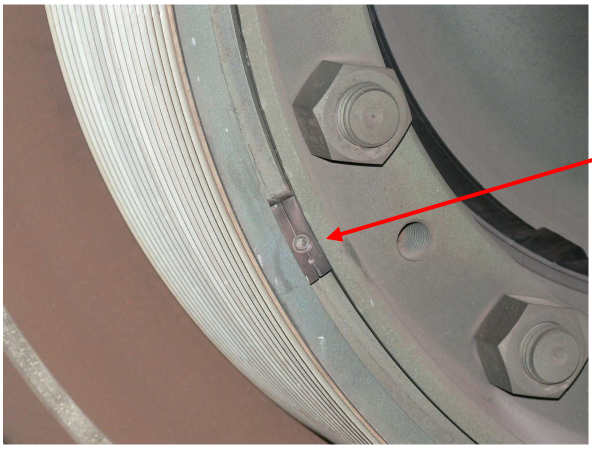


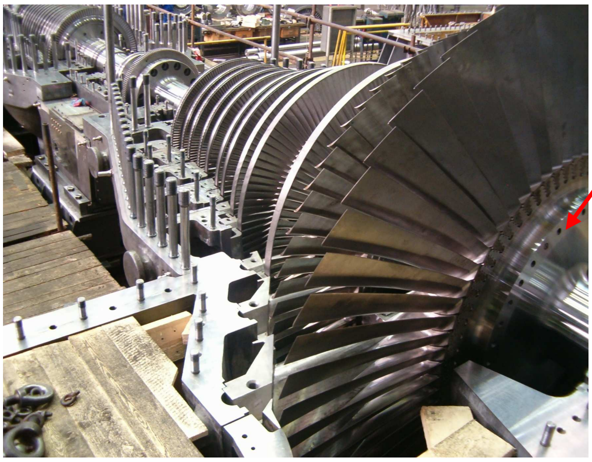
Fig. 4.1.2 - Two-plane Balancing

One of the important designer's tasks is to determine appropriate planes for balancing. These planes should be accessible without significant disassembly of the machine housing, which cannot be always managed. In addition, it is necessary especially for flexible rotors to choose the planes to allow good balancing with regard to operational deflection shape. Designer's task is also to select the appropriate way of attaching the balancing weights, or to provide adequate places for material removal during balancing.

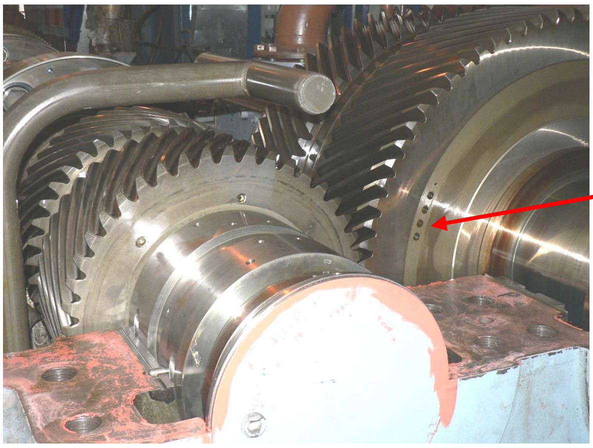
During a repair, the diagnostician should check the condition of balancing weights and compare their position with manufacturer/repair company documentation.



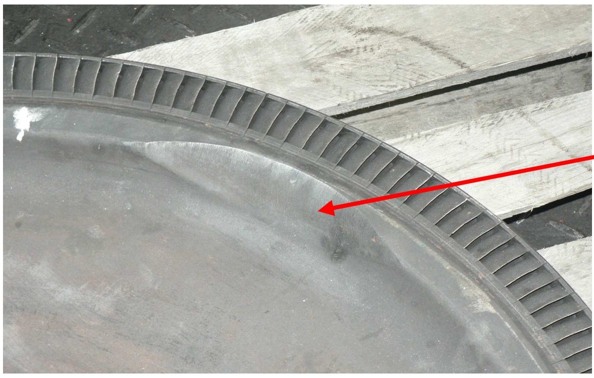
1. Inserting a suitable body into a dovetail groove. It is important to fix the body in the groove and prevent its shifting or dropping out. Here, the body is segmented and fixed with a screw.



2. Sometimes, threaded holes of sufficient depth are used instead of grooves. Balancing bodies are then screwed into these holes. This applies mostly when there is no place for a sufficiently dimensioned groove. This example is a steam turbine rotor. This method is often used for electric generators as well.



3. In cases where balancing weight cannot be attached, removal of material is carried out, e.g. by drilling off. The balancing procedure requires special care here, as the operation is irreversible. The example is from a high-speed power gearbox.



4. Drilling off is not possible for wheels of centrifugal pumps and compressors which have thin disks. Then, the material can be grinded off. In this example, there is a turbine disc. Considerable grinding off can be dangerous for strength reasons. Therefore, a designer can create a special collar for removing the material from the disc at positions where the stresses in the material are not critical.

Fig. 4.1.3 - Practical Examples of Balancing Weights Placement or Removal

It should be noted that it is impossible to balance the rotor perfectly, i.e. to achieve a zero unbalance. There always remains some residual unbalance.

In Fig. 4.1.3, some practical examples of balancing weights and their attachment are shown.

Balancing of flexible rotors state is a much more demanding operation. A flexible rotor does not rotate around its geometric axis any more, but rotates around its deflection curve (see Fig. 4.1.4), shape of which depends on the relation between operating speed and critical speed (which modal shape is close to the operational one and thus dominates in the deflection curve). Fig 4.1.4 (left) also highlights the fact that the deflection curve is not a plane curve (even if close to the first bending mode shape).

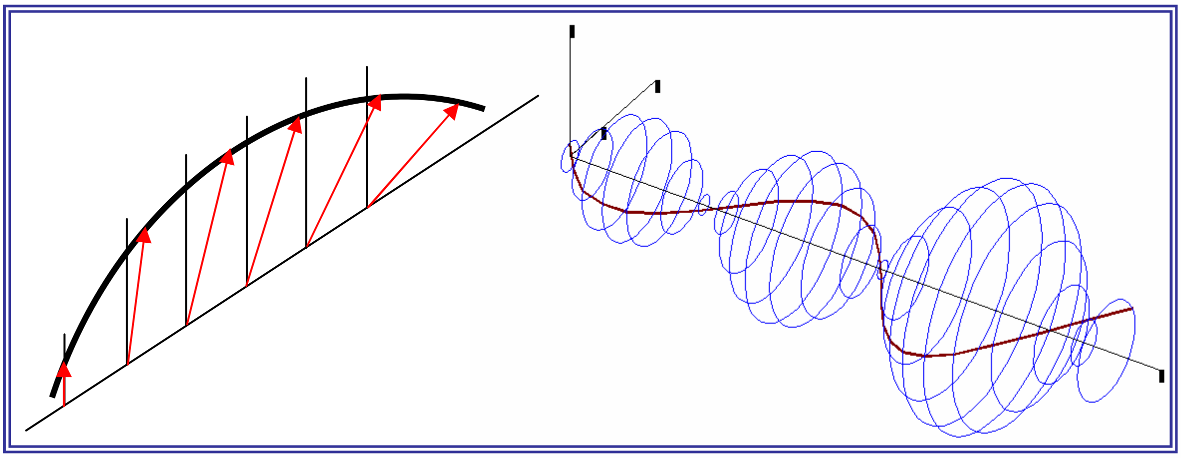


Fig. 4.1.4 - Deflection Curve of a Flexible Rotor

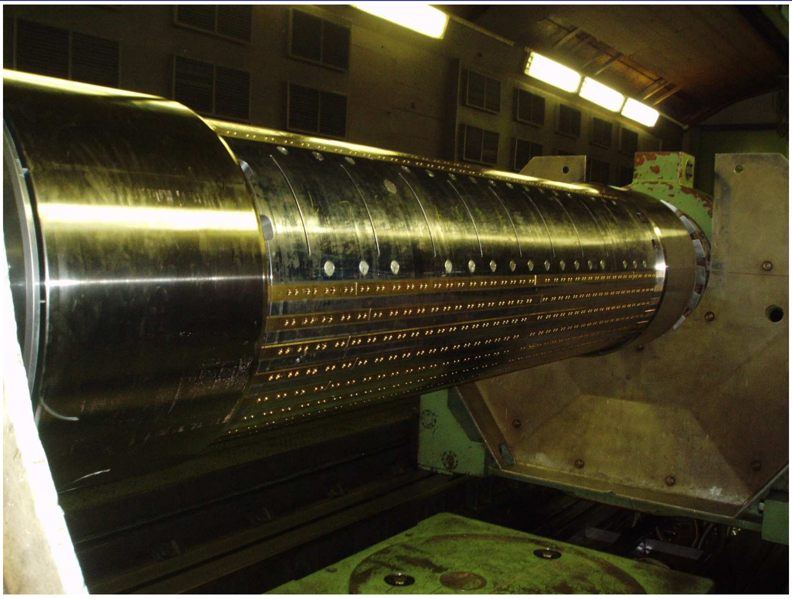
Not only unbalanced masses themselves, but also internal moments caused by these unbalances are involved in the flexible rotor unbalance. Therefore, flexible rotors should be balanced at operating speed. It might be a very serious problem - particularly in the case of turbomachine rotors with blades which cause large ventilation during rotation. Therefore, such rotors are balanced in balancing tunnels from which air is pumped off to a very low pressure.

Also, the number of balancing planes for flexible rotor balancing is larger than that for rigid rotor, for which two-plane balancing is usually enough. A rule of thumb says that the number of balancing planes must increase by at least one for each crossing of a critical speed. When a rotor operates above the first critical speed, then 2 +1, i.e. at least three balancing planes should be used, etc.

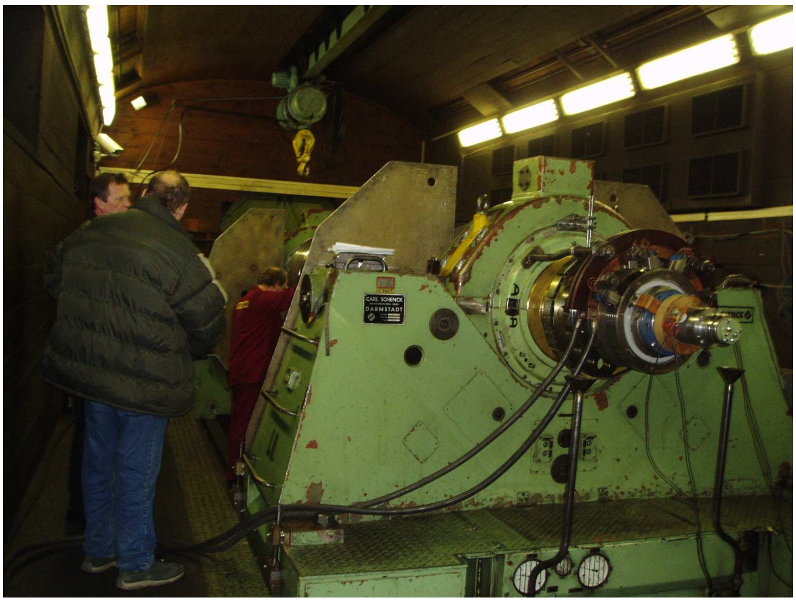
Balancing method is strongly dependent not only on running speed and critical speed, but also on the configuration of the rotor. The ISO 11342 standard defines and further describes seven different rotor configurations and pertinent rotor balancing procedures. When balancing, a rotor should be supported in the same bearings as during operation, which is a question of stiffness and damping and their influence on the critical speed and deflection curve shape. Practical examples of balancing on balancing machines are shown in Fig. 4.1.5.

The final balance is evaluated by the rms vibration velocity values measured on stands. It should be noted that when the actual running speed differs from the operating speed, vibration may be higher.

Note: High-speed rotors usually undergo centrifugation at a prescribed higher running speed "once in a lifetime" in order to ensure that they will not be destructed due to internal defects and also in order to have the individual parts "settled" (such as blades in hinges, etc.).



rotor of a generator (approx. 70 MW) on a balancing machine - a lot of balancing bodies and locations of final balancing are visible



high-speed balancing machine stand

Fig. 4.1.5 - Balancing on a Balancing Machine

4.1.3 Diagnostics of an Unbalance

In this chapter, diagnostic symptoms of unbalance will be introduced - manifestation of an unbalance in the measured vibration spectrum and phase values related with an unbalance.

1. Important symptom of an unbalance in a spectrum is **high vibration amplitude in radial directions at the rotational component** (1X). The reason is that the centrifugal force caused by the unbalance rotates with the rotational frequency and causes forced vibrations with the same frequency. A typical spectrum of unbalance is shown in Figure 4.1.6.

Harmonic cursor (vertical lines) indicates rotational frequency multiples (harmonics). Rotational frequency is 40 Hz here.

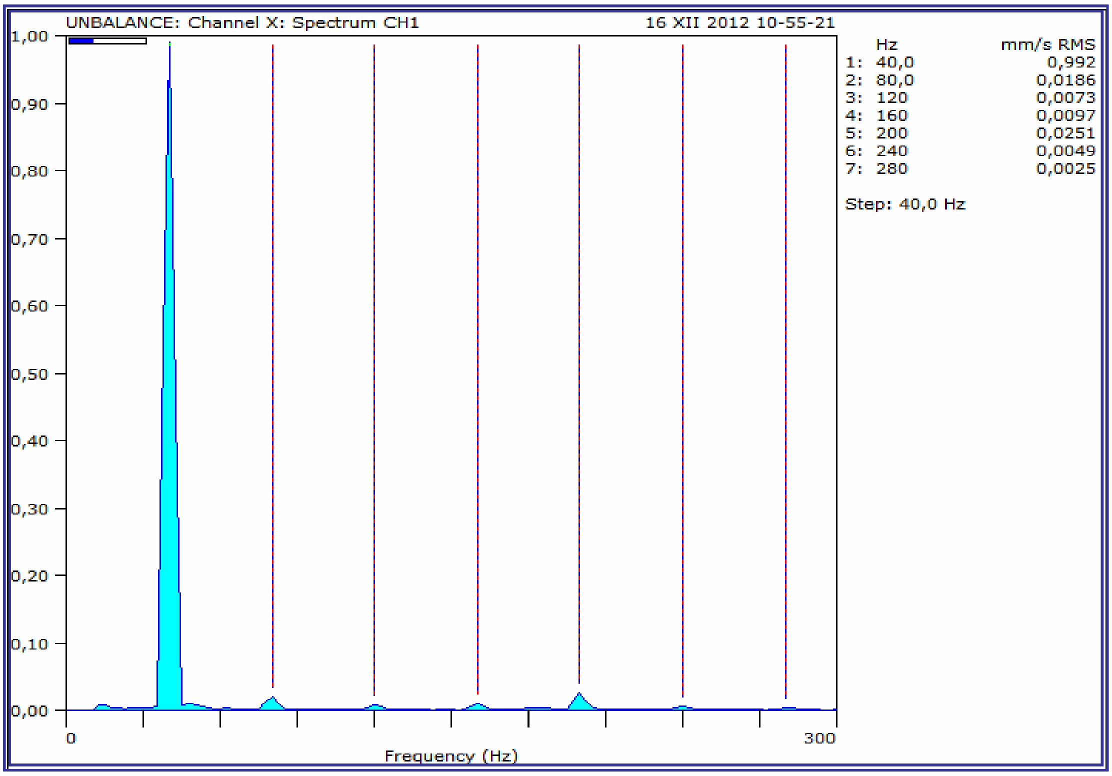


Fig. 4.1.6 - Typical Spectrum of an Unbalance

2. If it is not a very unique design of a rotor support in its housing, it is very important to compare the magnitudes of 1X components in horizontal and vertical directions in the given plane (on the given bearing stand).

   This ratio should not differ significantly from 1:1 ratio. However, as bearing stands are not equally rigid in these directions, many authors stated that **this ratio may be close to the value of 3:1** (regardless of which component is larger). If the ratio is significantly different, some other major cause of vibration is indicated (usually looseness or resonance - see below).

3. Except for machines with an overhung rotor (overhanging rotor end with a disc), there is no reason for high vibration in the axial direction in the case of unbalance ("high" here means comparable to the radial vibration). If so, this is again the influence of some other defect.

4. Given that the symptom of high radial vibration at the frequency of rotation occurs also with other mechanical machine faults, it is necessary to use further important sign for more precise diagnostics - **the mutual ratio/relationship of phase on essential parts of a machine**, usually on bearing stands or similar locations.

   Regardless of the type of unbalance, it is always important for the **phase difference to be about 90° between horizontal and vertical measurements** on the measured bearing stand (in the given radial plane). The word "about" in practice means that the angle can vary about ±30° from 90°. Otherwise, there is again some other major defect. (If the angle is close to 0° or to 180°, looseness or resonance is most likely the cause of problems.)

In addition, phase relations enable to distinguish between the typical cases of unbalance, which simplifies the selection of the appropriate balancing method of field balancing.

The individual types of unbalance have been described in section 4.1.1.1 and are shown in Fig 4.1.1. Now, it will be explained how they manifest themselves in spectra and what their phase relations are.

Static unbalance (shifted centre of gravity) manifests itself by the same phase being at both rotor bearings. There is approximately zero phase difference between vibration in horizontal direction measured on both bearings and the same applies to the vertical direction (see Fig. 4.1.7 bottom left). The term that "vibration is in-phase" is used. Furthermore, the phase difference between horizontal and vertical vibration on each bearing should be approximately 90° (±30°). To place (or remove) one balancing weight in a single balancing plane is sufficient to correct this unbalance.

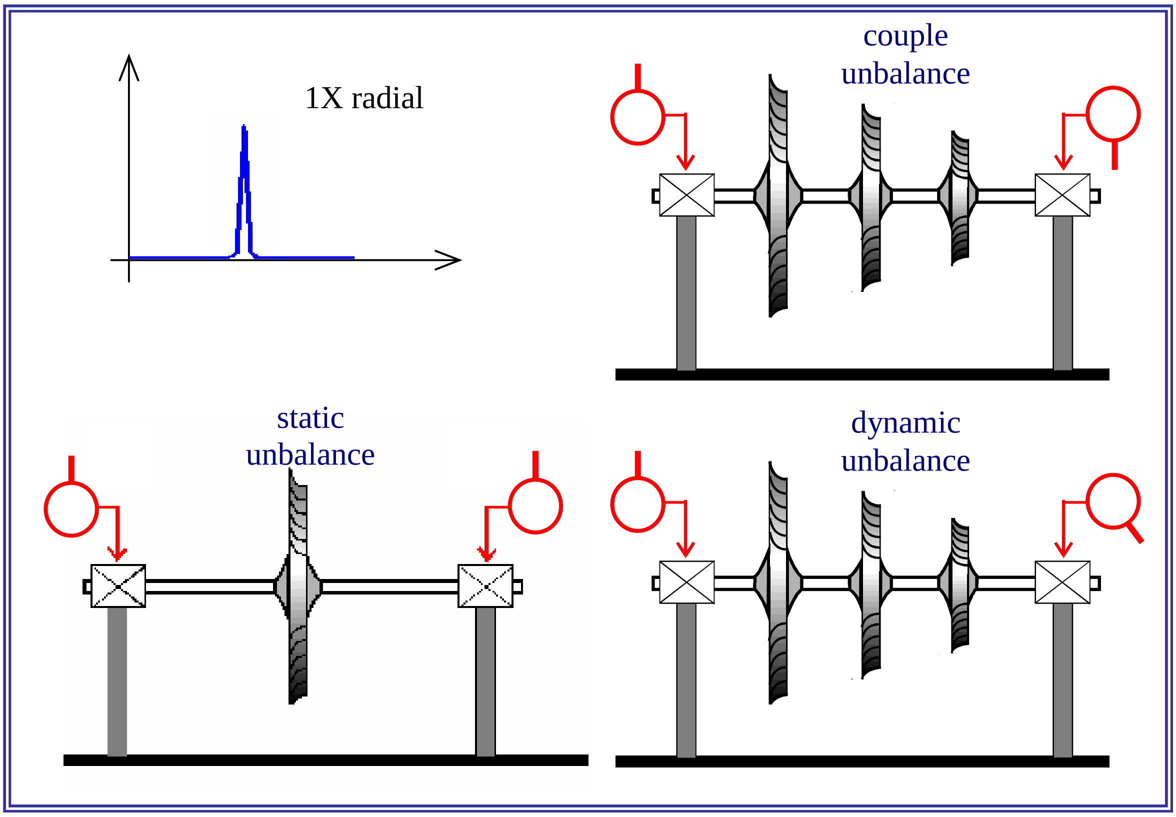


Fig. 4.1.7 - Unbalance - Typical Spectrum and Phase Relations

Couple unbalance manifests itself in such a way that there is approximately 180° phase difference between vibration in horizontal direction on the two bearings and the same applies to the vertical direction (see Fig. 4.1.7 top right). The term that "vibration is out-of-phase" is used. Furthermore, the phase difference between horizontal and vertical vibration on each bearing should be again approximately 90° (±30°). The 1X component is always dominant in the spectrum. Unlike the static unbalance, axial vibration can be increased as well. To correct the couple unbalance, balancing weights should be placed (or removed) in at least two balancing planes.

Dynamic unbalance is the predominant type of unbalance and is a combination of static and couple unbalance. The 1X component is dominant in the spectrum; axial vibration can be again increased. The phase difference between horizontal vibrations on the two bearings can be any from 0° to 180° (see Fig 4.1.7 bottom right) and the same applies to the vertical direction. However, when comparing measurements on the two bearings, the phase difference in horizontal direction should be about the same as the phase difference in vertical direction (± 30°). As with static and couple unbalance, phase difference between horizontal and vertical vibration on each bearing is about 90° (±30°). This applies if the unbalance is the prevailing problem. To correct the dynamic unbalance, balancing weights should be placed (or removed) in at least two balancing planes.

Overhung rotor unbalance causes high vibration with 1X component in radial directions, as with previous types of unbalance. Moreover, there is high vibration with 1X component also in the axial direction. Axial values tend to be in-phase (see Fig 4.1.8) while phase readings for radial directions may be unstable. However, the phase difference in the horizontal direction generally agrees with the phase difference in the vertical direction (±30°). Overhung rotors have both static and couple unbalance and the correction therefore requires placement of balancing weights in two planes.

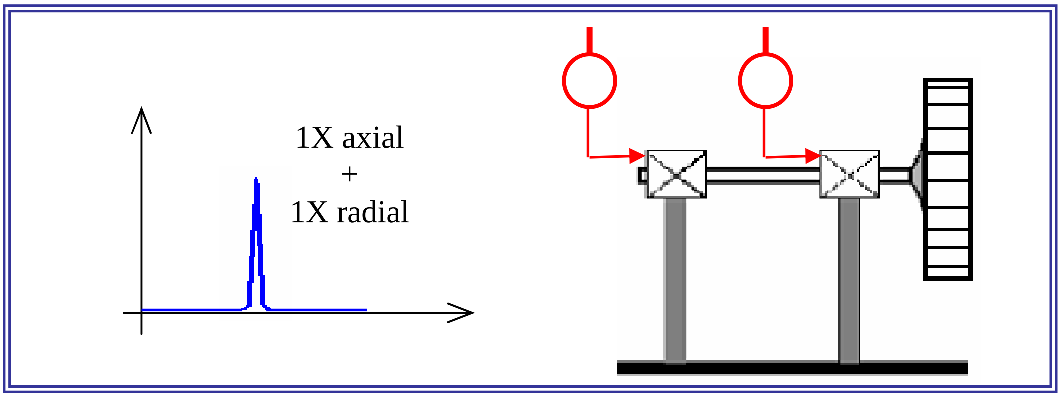


Fig. 4.1.8 - Overhung Rotor Unbalance - Typical Spectrum and Phase Relations

4.1.4 Field Balancing

In chapter 4.1.2 the issue of balancing on balancing machines was introduced. This chapter will focus on so-called field balancing. After the diagnostician identifies, on the base of diagnostic symptoms listed in the previous chapter that the monitored rotor is unbalanced, the usual solution is its balancing. Field balancing is used in various cases:

- When a **new machinery** (or **machinery after repair**) **is set into operation**, it is sometimes necessary to reduce vibration to an acceptable level. It happens even in cases when rotors were previously balanced on a balancing machine. This work is usually carried out by a manufacturer specialist who knows details about the machinery design and dynamics.
- It is also used in cases where the **device operates in an unfavourable environment** or in contact with corrosive substances, which cause relatively rapid wear of machine parts or change the balance due to deposits in combination with wear (flue gas fans, especially in the steel industry, coal mill fans in power plants, etc.).
- Another example are **machines, extended shutdown of which** (for dismantling, waiting for the balanced rotor, reassembly) **means a significant reduction in**

production for an extended period of time (usually several days). In normal cases, the field balancing procedure is planned for one shift - usually at night.

For these reasons, field balancing represents significant evidence that the diagnosis is meaningful and that it can help in increasing the efficiency of the entire production process.

In comparison with balancing on balancing machines, the field balancing has its advantages and disadvantages:

+ Cost savings needed otherwise for an extended shutdown of equipment, dismantling and transport to the balancing machine and back (a suitable equipment is often not in the location where the machine is).
+ Rotor is balanced in its own bearings and stands (housing).
- The need of multiple run-up and shutdown of the machine for runs with trial and balancing masses.
- Shutting-down the machine from the manufacturing process during the time of field balancing.

4.1.4.1 Technical Preparation

First of all, it is necessary to verify that the unbalance is really the cause of identified increased vibration, which in practice is quite underestimated. This relates to the fact that unbalance is always present to some extent, and thus it is estimated that up to 80% of vibration problems is just due to unbalance.

This problem is dealt with in the IRD text (see [18]] which discusses cases when balancing is not advisable and describes a procedure for exclusion of cases where balancing does not lead to a successful outcome. Following steps are recommended:

- **Problem analysis** - To determine what was happening with the machine (replacement of its parts, alignment, accidents, etc.). For example, sudden change in vibration level indicates that some machine part has broken off, etc.
- **Assessment of vibration** (for more detail see the diagnostic table [19])
  - See the table - a lot of faults are manifested on 1X, not only unbalance
  - Beware of resonance - at running speeds close to resonance, balancing is not possible
  - Basic rules:
    - vibration ratio vertical/horizontal has to be 1:3 (sometimes up to 1:5)
    - phase difference vertical-horizontal has to be about 90° (on the measured stand)
- High **axial vibration** is associated with a bent shaft or improper alignment rather than with unbalance (unless an overhung rotor).
- High **harmonics** are usually associated with mechanical looseness.
- If there is a **minor phase difference** between horizontal and vertical vibration (the so called "uni-directional vibration"), the problem is most likely eccentricity, mechanical looseness, resonance or bent shaft.
- Some **electrical problems of motors** are also manifested on 1X component or in its vicinity.

Some purely practical problems may occur during field balancing, such as access to the rotor. Not all machines have **access to the rotor**, e.g. via blind holes, through which weights may be attached. This often limits even the use of balancing method - for example only single-plane balancing method is used, although the type of unbalance would require two-plane balancing. Access to the rotor is required also from the viewpoint of cleaning a wheel, removing sludge and any old weights, check of runout and blades wear, etc.

Attaching weights may also be a problem (see Fig 4.1.3). On some machines, welding is used, which introduces two risks:

- Insufficient weld can mean weight detachment during rotation, while thorough attaching, particularly of the test weight, significantly hinders its removal before final balancing.
- In the case that the welding apparatus ground cable is not placed on the rotor to which the weight is attached, it means that the electric circuit is closed through the bearings (usually rolling bearings), and thus in a short time after balancing was performed, condition of some bearing might get worse to such extent that it needs replacement.

Before balancing procedure is started, it is advisable to have prepared the appropriate weights, at best sets of different weights. Prepared **weight sets** can greatly facilitate the balancing work. Size of the weights should correspond to the balanced machine (few grams for small machines to hundreds of grams for larger ones).

It is necessary to ensure the occupational safety as well - the average surface speed when the shaft diameter is of 800 mm and its running speed 1000 rpm is about 151 km/hr !

4.1.4.2 Operational Preparation

Unlike balancing on a balancing machine, on which the size and position of the balancing weight can be determined after one run only (thanks to an electric frame and known influence coefficients), the field balancing procedure usually requires several runs:

- **reference** run - measurement of overall values, 1X component and phase on all important locations on the machine (i.e. on bearings in all three directions)
- one or several (depending on number of balancing planes) **trial** runs when a trial weight is attached to the rotor to find out the sensitivity of the machine to unbalances
- **check** run when the calculated balancing weights are placed on the rotor (measurements at all the measurement points are performed again)
- possible **trim** run for further refining of the balance

All of this takes time. It is necessary to ensure the possibility of multiple starting off and stopping the machine, which involve timeouts. It is necessary to wait for:

- running-up the machine to operational running speed
- steady operation (temperature equalization, etc.)
- coasting-down the machine (it may take several minutes)

Care must be taken to limitations of machine starts - bigger asynchronous motors cannot be started unless the motor is cooled down. Limitations of starts of induction motors is due to the fact that rotor currents are much larger during start-up than during operation and that the motor must be cooled before the next start. Usually, 30 minutes are prescribed for cooling, but it is advisable to check the time before balancing to avoid accidental damage of expensive equipment.

4.1.4.3 Confirmation of Unbalance as the Main Cause of Problems

Given the fundamental importance of verifying that the increase in machine vibration is caused by an unbalance, all the steps that are necessary to perform before balancing procedure is started will be repeated here once again summarized:

- Vibration measurements of the entire machine set are carried out.
- The following should be confirmed for the part that is considered for field balancing:
    - Vibration on 1X component is dominant.
    - Vibration in horizontal direction is usually higher than that of vertical direction (due to support stiffness).
    - Phase difference between horizontal and vertical vibration is approximately 90° (±30°).
    - Horizontal to vertical vibration ratio is not greater than 3:1! (It may also apply inversely - for V/H).
    - Amplitudes of harmonics - 2X, 3X etc. are minor.

Note on measurements: It is recommended to add at least an overall vibration measurement on machine feet and a frame (pedestal) to which the machine is fixed to the standard measurements (3 directions at bearing housing). In some cases this will easily detect mechanical looseness that would otherwise complicate the balancing process. In such cases, vibration values on machine feet are the same or even higher than at standard measuring points.

4.1.4.4 Decision about Balancing Method

Once it was verified that machine problems are caused by an unbalance, suitable field balancing method can be chosen. Several methods exist:

- single-plane balancing without measuring phase
- single-plane balancing utilizing phase measurements
- two-plane or multi-plane balancing

As it was already stated, number of balancing planes is determined by:

- **type of unbalance** - When static unbalance prevails, one plane is enough; in case of dynamic unbalance two planes are necessary and even more for flexible rotors.

- **accessibility** to the rotor - When there is only one window for access to the rotor, then only single-plane balancing is applicable. The problem of dynamic unbalance cannot be solved completely by this, but still the unbalance can be reduced.

4.1.4.5 A Brief Overview of the Most Common Methods of Balancing

Basic principles of the most common methods are stated in this chapter.

4.1.4.5.1 Single-Plane Balancing

4.1.4.5.1.1 Three-Circle Method (Siebert Construction)

This method is applied particularly in cases when there is no indication of phase, for example, when the shaft is neither visible during operation (photo-probe cannot be used) nor phase sensor based on eddy currents is installed. However, this method is used in practice, e.g. for centrifuges, large cooling tower fans or other similar machines. In this case it is particularly important to take advantage of comparison of vibration in different directions and places before a decision for balancing is made. Only overall vibration measurements can be used (or, if spectrum is measured, the 1X component). The procedure requires a total of 4 measurements - one reference run for determining the vibration level with the original unbalance and 3 test runs. During each test run, the same test weight is placed to different positions on the rotor. The best results are achieved if the test weight is evenly placed by 120º. This is neither always possible nor necessary. Figure 4.1.9 shows an example where test weights were placed by 90°. (This would be e.g. a fan with 4 blades.) From the measured data, the position to which the balancing weight should be placed (and its mass) is then identified using the three-circle method.

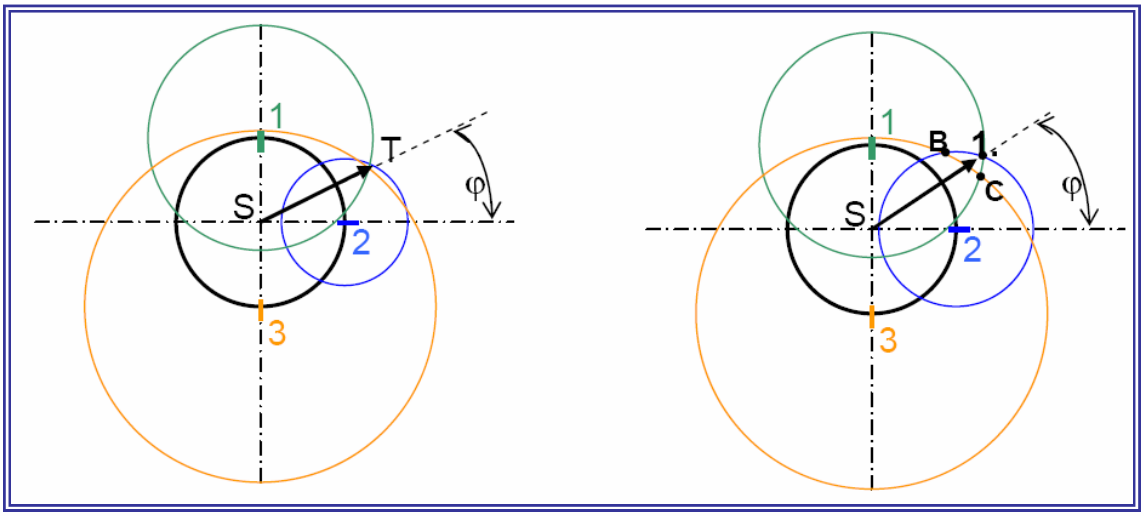


Fig. 4.1.9 - Three-Circle Method (Siebert Construction)

4.1.4.5.1.2 Vector Method (Utilizing Phase Measurements)

Vector method is the most common method of single-plane balancing used in practice. Influence factors are not known in field balancing. Therefore, to identify the mass and position of the balancing weight correctly, test of machine sensitivity to the unbalance must be carried out. This is the principle of all balancing procedures that use the measured values of 1X component of vibration and its phase. The procedure is as follows:

A test weight of known mass is attached to the rotor (i.e. that its position is known as well). From the known responses to the original an additional unbalance, the correct mass and position of balancing weight is determined. Thus, the balancing process requires two runs -

one to determine the initial unbalance, the other to determine the effect of the additional test weight. It is important to understand two things:

- The method assumes that the **balanced system is linear**, which simply put means that a small change in unbalance causes a small (proportional) change in vibration. This may not be true, especially when the machine is working relatively close to the natural frequency (or critical speed), when there is mechanical looseness, etc.
- It is further assumed that the measured amplitude of the **1X component is only due to unbalance**, but sometimes it is not true - there is always also some misalignment, etc.

When using the vector method principle which is in Fig 4.1.10, it is necessary to understand what each vector means and not to confuse the influence of the test weight $\mathbf{v_T}$ with the aggregate influence of both the initial unbalance and test weight $\mathbf{v_1}$. The aim is to determine the size and direction of the vector $\mathbf{v_v}$, which cancels the original unbalance vector $\mathbf{v_0}$. It should be realized that the angle $\boldsymbol{\varphi}$ which determines the position of this vector has not only size but also direction. Here is the most common source of errors during balancing: The entire procedure is performed, mass and position of the balancing weight is determined correctly, but finally the weight is simply placed to the opposite side with respect to the position of the test weight than it should be. This results in a fatal failure - the machine vibrates even more than at the beginning.

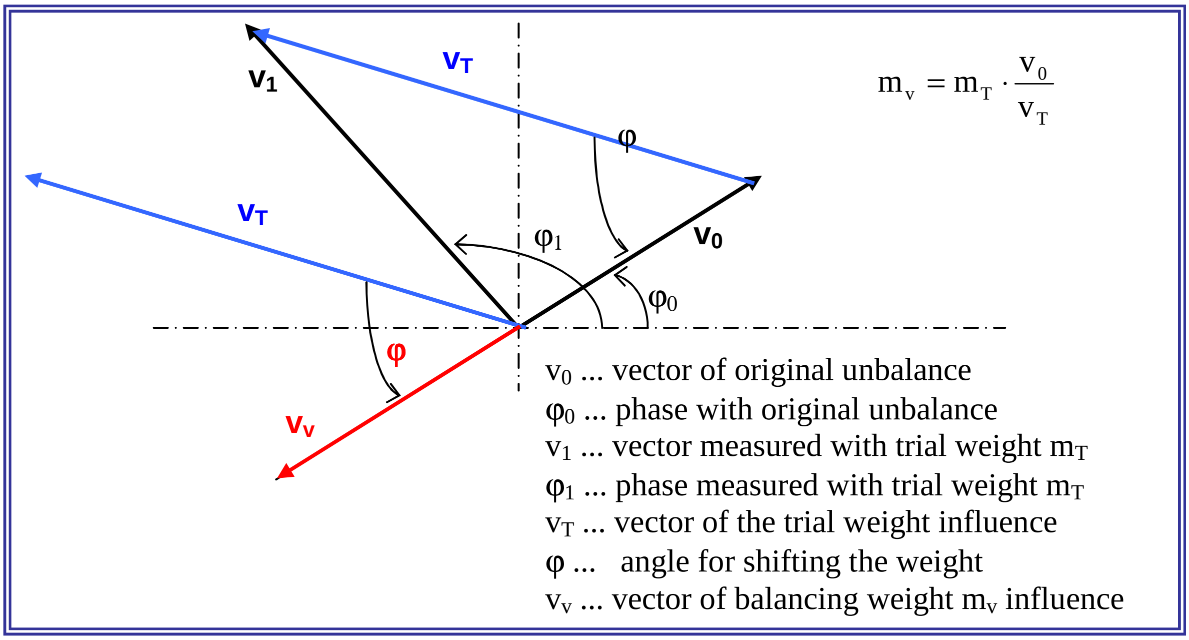


Fig. 4.1.10 - Vector Method

The procedure for placement of balancing weight must be such that for positive direction of rotation from $\mathbf{v_T}$ to $\mathbf{v_v}$ (i.e. counter-clockwise), the balancing weight is shifted with respect to the test weight in such a way that the rotor is rotated by the proper angle in the direction of its normal rotation and the balancing weight is placed to the resulting position. (Imagine that we observe the test weight on the rotor through a peephole. Then we rotate the rotor by the angle $\boldsymbol{\varphi}$ in the direction of its normal rotation and place the weight). If $\boldsymbol{\varphi}$ angle is negative (i.e. clockwise in the diagram), the rotor should be rotated against its normal rotation.

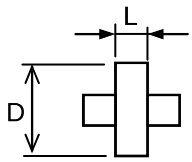
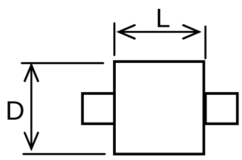
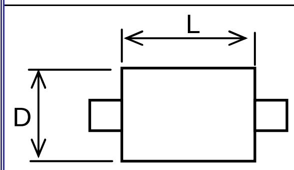
Note: Vector balancing method is commonly implemented in analyzers designed for field measurements, so hand drawing the vector diagram is not necessary.

4.1.4.5.2 Field Two-Plane Balancing

Balancing in two planes, in which four influence coefficients must be identified, is usually done with computer support (using the built-in balancing software in an analyzer). Its principle is the same as the vector method explained in the previous chapter on balancing in a single plane, but due to the fact that the test weight added to one of the balancing planes affects vibration on both bearings, interpretation of vector diagrams is not straightforward any more. Given the availability of modern equipment, vector method is not used directly.

Note: If the rotor is operated with running speed higher than 70% of its critical speed, it is considered as flexible and its balancing usually requires correction in more than two planes (see table 4.1 for details).

Table 4.1 - Number of Balancing Planes with Respect to Rotor Dimensions

	L/D ratio	1 plane	2 planes	more planes
	< 0,5	0 - 1000 rpm	> 1000 rpm	not applicable
	0,5<L/D<2	0 - 150 rpm	150-2000 rpm or >70% first critical speed	> 2000 rpm or >70% first critical speed
	> 2	0 - 100 rpm	100 rpm to 70% first critical speed	>70% first critical speed

4.2 Misalignment

Misalignment is another of the common faults of rotating machinery. It is the result of incorrect alignment of machines, which is necessary to be discussed in the beginning of the chapter focuses on misalignment.

4.2.1 Alignment of Machines

The term *alignment* technically means that the rotors rotate during operation in their cases in correct position so that they can fulfil their function without additional problems. Usually it is assumed that the rotor axis is concentric with the axis of housing and bearings and that rotors connected together have a common axis. It is therefore to achieve alignment.

The concept of *alignment* thus comprises two technically quite distinct categories:

1) alignment of a rotor in its case
2) alignment of the machine set rotors one to each other

It is obvious that basic preconditions for the correct alignment are created by a designer/ manufacturer, in a way that he adequately incorporates elements for adjusting the position of the rotor axis towards the axis of the housing to the construction and also by suitable choice of rotor couplings, including measures to establish their parts one to each other.

Incorrect misalignment can result in:

- worse machine performance (e.g. because of changes in clearances),
- increase of power needed to drive the machine (power is wasted in a coupling and transformed to a heat - see Fig. 1.3),
- increase of wear of bearings, seals and often also in parts of a coupling as a result of extensive forces transferred,
- increase of noise and vibration that again means an excessive load of both the machine and its basement.

A correctly aligned machine set and two types of misalignment - *parallel (radial)* misalignment and *angular* misalignment - are shown in Fig. 4.2.1.

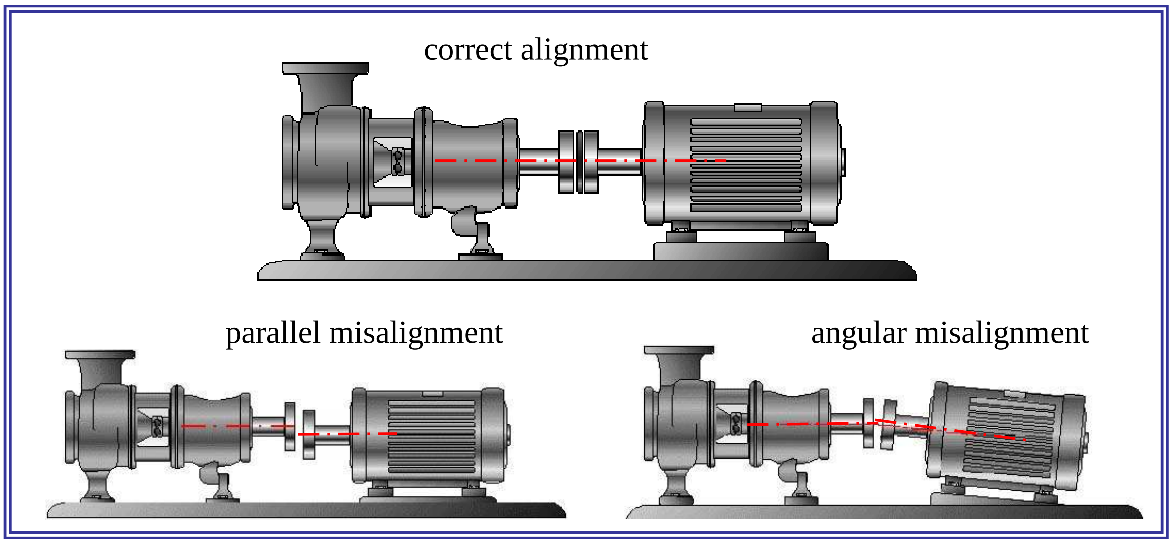


Fig. 4.2.1 - Correct Alignment, Parallel Misalignment and Angular Misalignment

It is obvious that in the case of aligned rotors, only the desired torque without any additional forces/moments is transmitted by a coupling, whereas in a misaligned set, other forces and moments that act on the coupling and the shaft and that must be captured by bearings are generated in the coupling. Therefore, all components that affect this alignment are subjected to strict requirements to achieve the proper alignment; see the following typical examples of tolerances:

- <u>Basement:</u> The recommended tolerance is 0.025 mm for all mounting washers; machine feet should be complanar up to 0.025 mm limit.
- <u>Deformation</u> as a result of pipes or channels: The recommended precision tolerance is limit of shaft deviation up to 0.05 mm in the coupling axis in any direction if pipe or channel is connected.
- <u>Coupling:</u> Common recommendations for "cold" machine alignment (e.g. up to 0.05 mm radial), featured in many books, often do not respect the different thermal expansion of the connected machines and can lead to very unpleasant problems.

Equally misleading is the assertion of different coupling manufacturers of how much of radial or angular misalignment their couplings would withstand. They do not say, however, that this may not withstand adjacent bearings to which the load from misalignment is transferred.

Two typical examples of couplings - rigid and flexible (segment) coupling - are shown in Fig. 4.2.2.

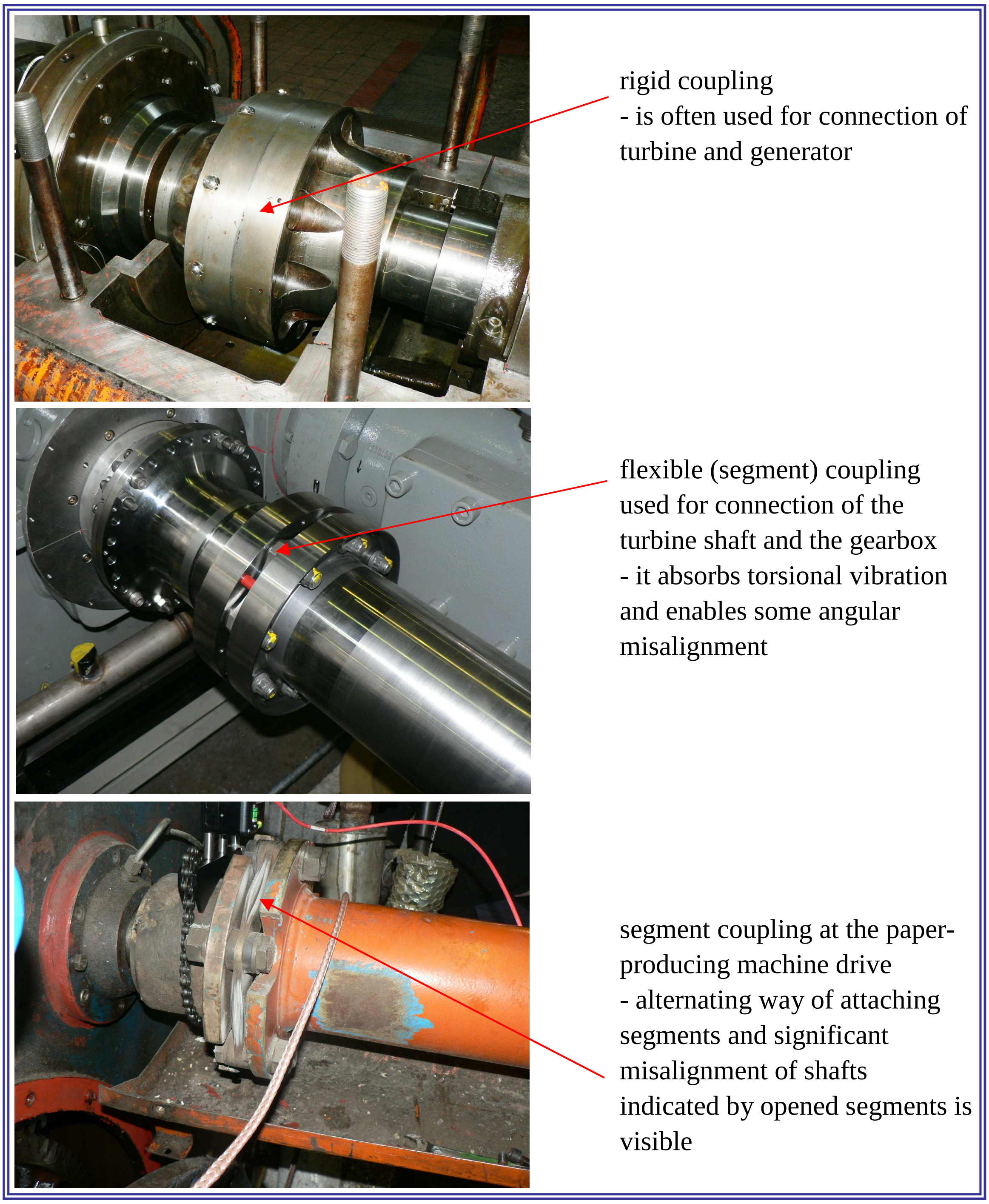


Fig. 4.2.2 - Practical Examples of Couplings

To illustrate the problem with dilatation, coefficients of thermal expansion of different materials are shown in Table 4.2 and an example of the prescribed values for the cold coupling alignment is shown in Fig 4.2.3.

Table 4.2 - Thermal Expansion of Different Materials

Thermal Expansion Coefficient	
Material	mm/ºC/m
Aluminium	0,0234
Brass, cast	0,0187
Carbon steel	0,0113
Cast-iron	0,0106
Nickel steel	0,0131
Stainless steel	0,0173
Concrete (variable)	0,0018

Note: The coefficients are technically correct, but for better understanding of severity of the problem of thermal expansion, the following "example" is perhaps more appropriate:

Steel rod with a length of 1 m is lengthened by 1 mm after 100 °C heating!

Therefore, it is necessary to consider the different thermal expansion when components with significantly different operating temperatures are connected to the sets, for example:

- an electric motor and a pump for high temperature liquid
- an electric motor and a flue gas fan
- a turbine and a compressor with high or low gas temperature
- a turbine with a gearbox that drives a generator

The problem is more complicated when some of the components (or all of them) are supported on journal bearings. As explained in Chapter 4.4 dealing with journal bearings, centre of the shaft is shifted during rotation from the bearing axis to the side in the direction of rotation and upward from the rest position. When couplings are aligned, it is therefore necessary to take into account not only the temperature difference, but also this shift during rotation. The result is that in no case coupling can be aligned "to zero" in the rest state.

An example of the prescription for "cold" alignment for machine set consisted of a turbine, a gearbox and a generator of a power about 40 MW is shown in Fig. 4.2.3. It is evident how much the values of the "cold" alignment can differ from zero to achieve smooth operation of a machine.

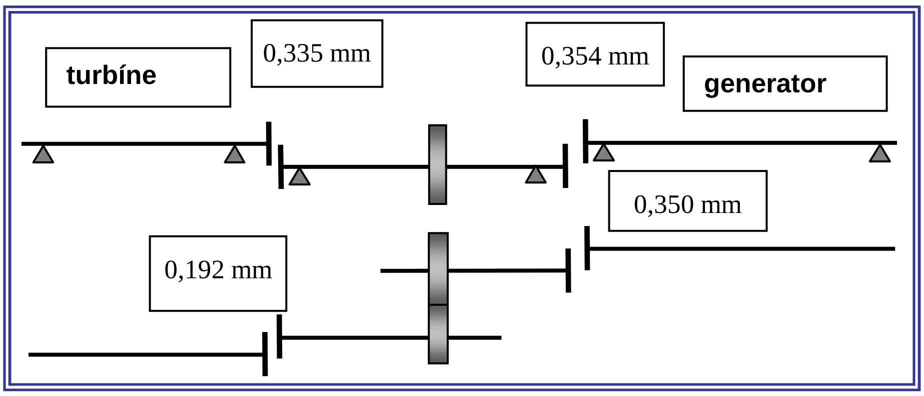


Fig. 4.2.3 - Example of Prescription for "Cold" Alignment
(upper scheme - front view, lower scheme - top view)

4.2.2 Diagnostics of Misalignment

4.2.2.1 Angular Misalignment

Angular misalignment is characterized by large axial vibration, which is out-of-phase, i.e. with the difference 180° over the coupling. In a typical case, large axial vibration is on both 1X and 2X components (see Fig 4.2.5 above). But it is quite common that any of the components 1X, 2X or 3X dominates. These symptoms may also indicate the existence of the problems with the coupling. The considerable angular misalignment may excite many harmonics of rotational frequency.

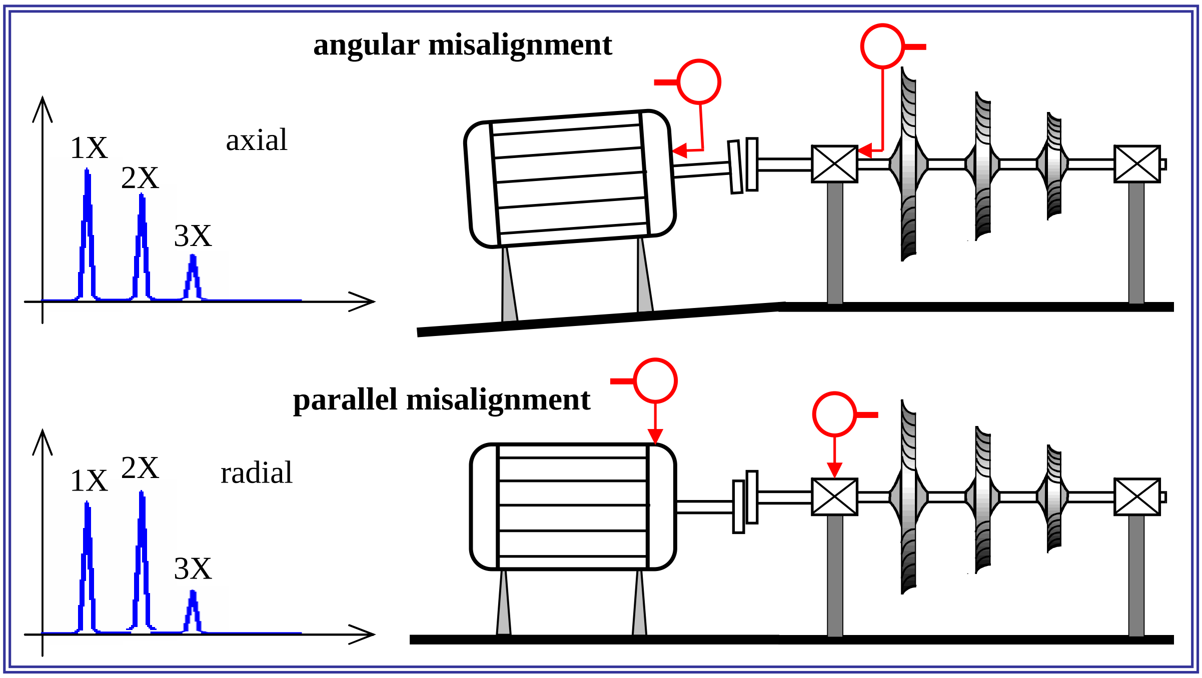


Fig. 4.2.5 - Angular and Parallel Misalignment - Typical Spectrum and Phase Relations

4.2.2.2 Parallel Misalignment

Misalignment, resulting from parallel shifting, has similar symptoms as an angular misalignment, but large vibration is in the radial direction. They are approximately out-of-phase, i.e. shifted by 180° over the coupling. The 2X component is often larger than the 1X, but its size relative to the 1X is often determined by the type and construction of the coupling (see Fig 4.2.5 below). When the angular or parallel misalignment is significant, it can generate either high amplitude peaks at several harmonics (4X to 8X) or even a lot of harmonics to a high frequency, which is similar to the mechanical looseness.

4.3 Diagnostics of Rotor Systems Faults

4.3.1 Rotor Resonance

When the rotor running speed is close to some of its natural frequencies, a phenomenon called resonance occurs. Such speed is called *critical speed*. Given that conventional steel structures have only small internal damping, the increase of the amplitude (gain) in the vicinity of resonance is very high (16 to 20×), especially for the first natural frequency. This means that even a small excitation force can have destructive influence on the structure. Therefore, it is necessary to design a machine so that operating excitation frequencies would be sufficiently apart from its natural frequencies. For example, turbines are usually designed so that the operating speed would be 70% below or at least 15% above the critical speed. Note: Due to the fact that the rotor has more (theoretically infinitely many) natural frequencies, it has also more critical speeds. The *first* critical speed means that rotor vibrates on its first bending natural frequency, etc.

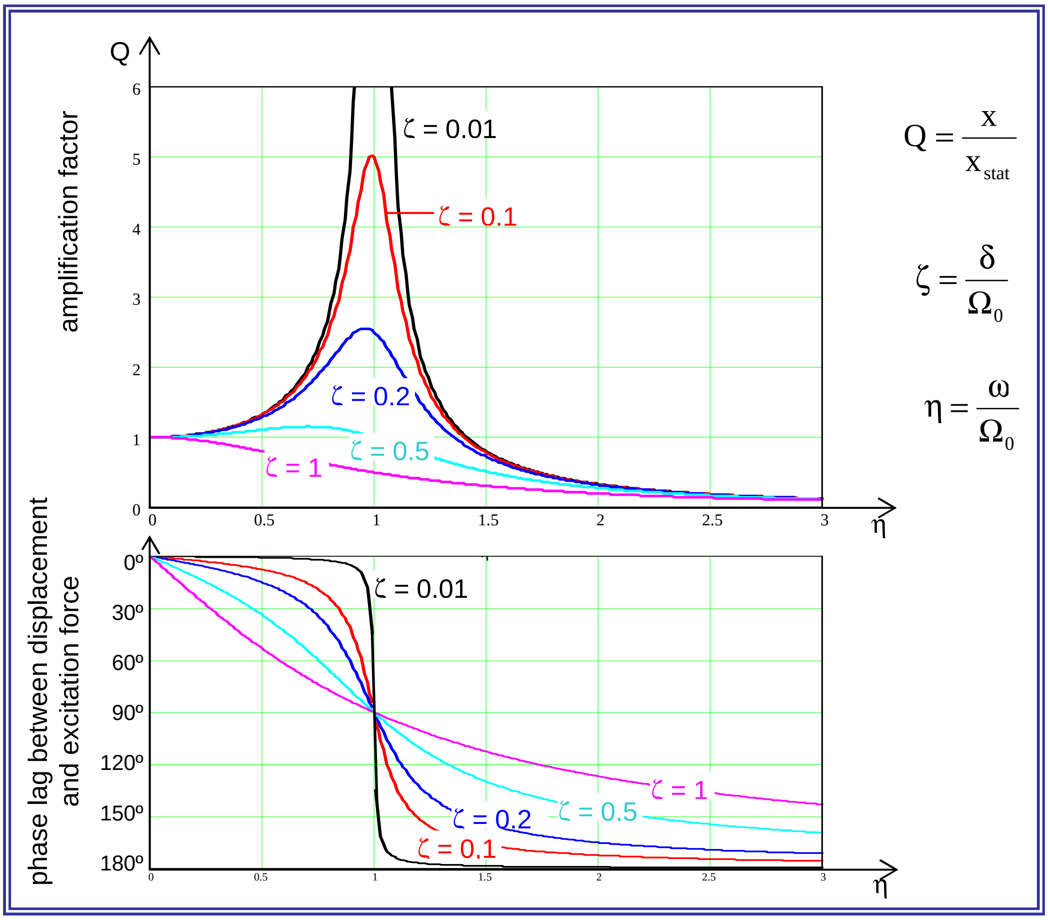


Fig. 4.3.1 - Amplitude-Phase Diagram (Bode Plot) (Dimensionless)

In Fig 4.3.1, so called *amplitude-phase characteristic*s (also called *Bode plot*) is shown. It consists of two plots - amplitude and phase, both as a function of the excitation frequency (i.e. as a function of running speed in the case of rotating machinery). More illustrative form of the amplitude-phase characteristics is its dimensionless form, i.e. with so-called *tuning factor* **η** (ratio of excitation to natural frequency) on the horizontal axis and *amplification factor* **Q** on the vertical axis. Gain factor is a ratio of displacement to static displacement (i.e. indicates how many times the displacement at the given frequency is higher than the static displacement). The characteristic is shown in this form in Fig 4.3.1. Phase here is defined as the delay of the response behind the excitation force. Such a phase could not be routinely measured as excitation forces are not measured during operation. Nevertheless, the difference between phases when operating below and above resonance can be determined by measurements.

It can be seen from the amplitude-phase characteristics that the damping affects not only amplitude of the displacement in the vicinity of resonance but also its phase significantly - when the damping is small (case of rolling bearings), the phase change is rapid, while it is slow with large damping. Journal bearings exhibit greater damping in the oil film and slow phase change, which is often not quite obvious in the measured data.

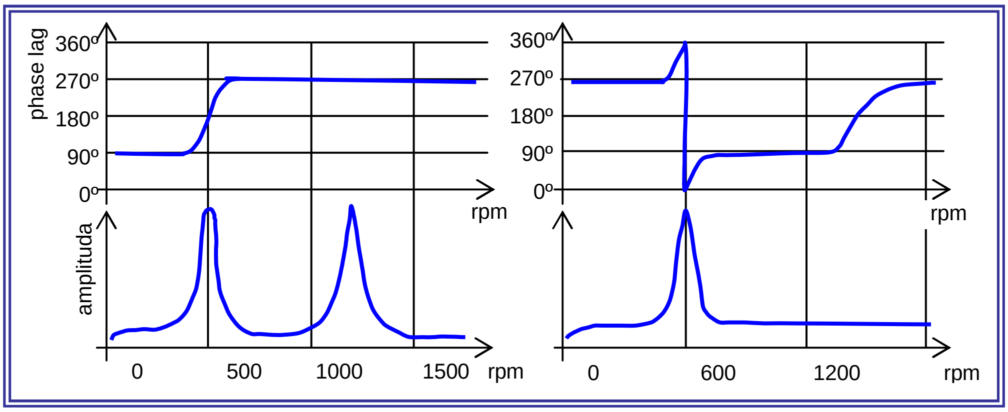


Fig. 4.3.2 - Phase Change in Resonance

In Fig. 4.3.2 there is an idealized example illustrating how to read the phase change in relation with resonance (critical speed) from the measured data. The amplitude-phase characteristic on the left indicates only one critical speed (500 rpm). The second peak amplitude at a speed of 1200 rpm probably corresponds to higher excitation or to transfer from other parts of the system, but there is no phase change, so it is not a critical speed peak.

A characteristic in Fig. 4.3.2 on the right indicates two critical speeds, although there is only one peak in the amplitude plot. Change of both amplitude and phase can be seen at the critical speed 600 rpm, but only phase change can be seen at the critical speed 1400 rpm, the amplitude does not change (measuring point was probably close to the node). Thus it can be concluded that *phase change* is crucial for identifying the critical speed. Note: The phase "jump" from 360° to 0° phase is not a change, but is caused by scaling the vertical axis from 0° to 360°.

The rotor running speed may be (and often is) higher than critical because above the critical speed the rotor displacement decreases up to a value less than the static displacement, as seen from the amplitude characteristics. This is due to the response delayed behind excitation that above the critical speed the rotor's centre of gravity returns to the axis of rotation. This is called *self-centering* of the shaft.

4.3.1.1 Critical Speed of Rotor Sets

Rather than a single machine (single rotor), a machine set is always analyzed in practice, i.e. at least two machines together, of which at least one is driving and one driven.

The issue of critical speed will be illustrated on the example of a machine set that consists of two parts of a turbine (high and low pressure), generator and exciter (see Fig 4.3.3). In this case, the set has four first critical speeds (which is still a simplification). Coupling fields are not considered separately.

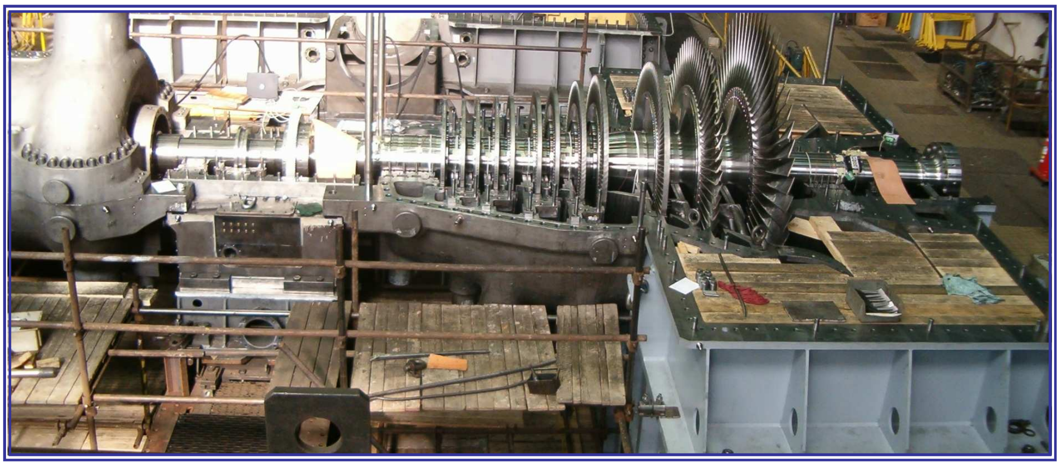


Fig. 4.3.3 - Example of a Part of Machine Set with Four First Critical Speeds

Low pressure turbine rotor (LP) is heavy and less rigid (it is at right in the picture), so it has the lowest critical speed (see Table 4.3). In the LP rotor axis direction (to the left), coupling field can be seen and the high-pressure part of the turbine (HP). Its rotor is lighter and stiffer and has higher critical speed (see Table 4.3). Values of the amplification coefficient Q correspond to the rotor on journal bearings. The intended rotor system has also rotor of the generator and rotor with an overhung exciter. The rotor of the generator is shown in Figure 4.3.7.

Table 4.3 - Computed Values of the First Critical Speeds of a Rotor System

Computed Critical Speed (rpm)	Amplification Coefficient Q	Description of the Modal Shape	Picture of the Modal Shape
1891.1	7.6	LP rotor first bending mode	Fig. 4.3.4 above
1991.5	8.3	gen. rotor first bending mode	Fig. 4.3.4 below
2432	--	HP rotor first bending mode	--

Modal shapes of LP rotor (above) and generator rotor (below) are displayed in Fig 4.3.4. Description of parts of the rotor system that is listed for the top picture applies equally to the picture below. On the free end of the generator shaft there is an overhung exciter, so the vibrations in the free end are high.

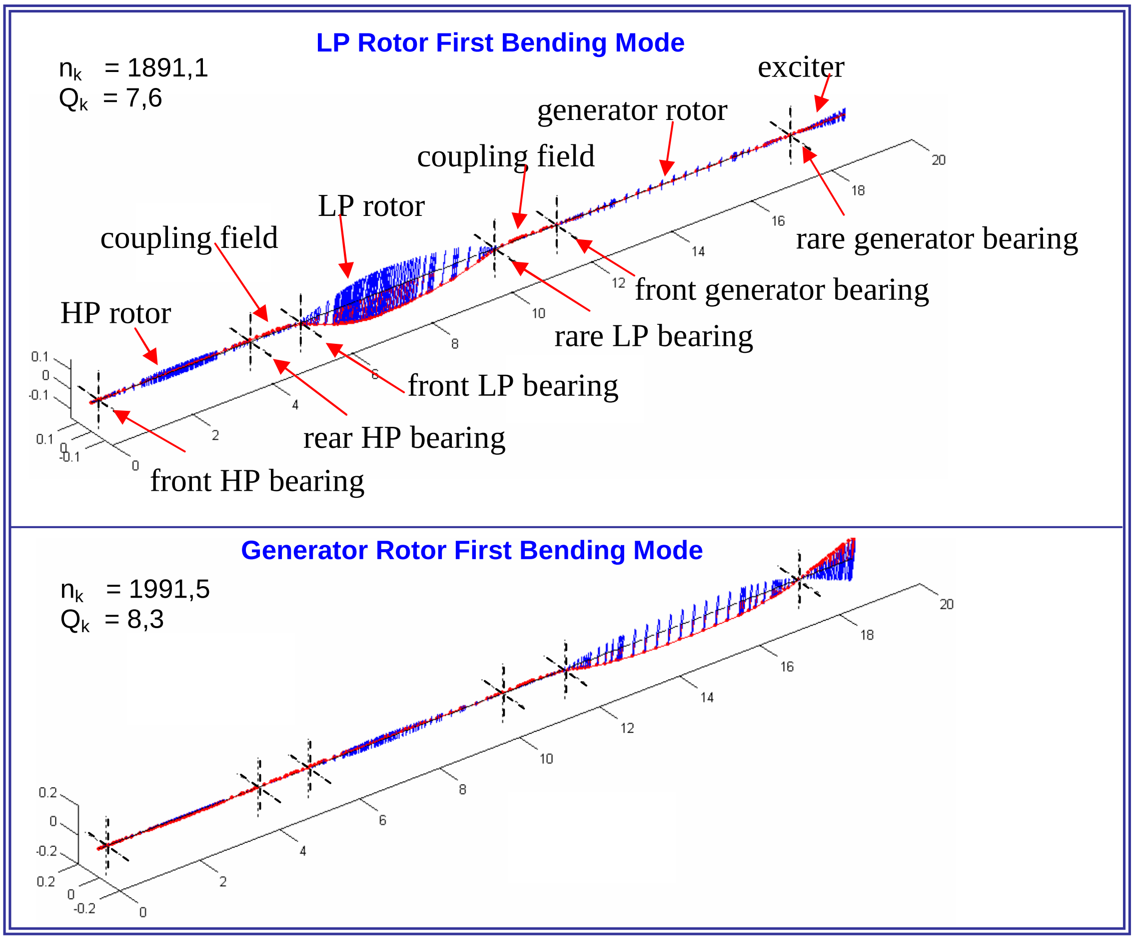


Fig. 4.3.4 - First Bending Modes of the Machinery Set from Fig. 4.3.3

4.3.1.2 Parameters that Affect Critical Speeds

As noted above, each machine (rotor) has more critical speeds. In Fig 4.3.5 it is displayed to which modal shapes correspond the first, second and third critical speeds. Value of critical speed as a function of support stiffness is also shown in that figure.

Along with a decrease of the critical speed value depending on the decrease in stiffness, the shape of vibration changes as well when the stiffness is low (see schematic representation of the vibration shapes on the left and right of the figure). It is seen that the first and second bending mode changed into rigid body modes at very low support stiffness, the third bending mode became the first bending, etc.

The situation is even more complicated as the stiffness of the support also varies with running speed of real machines. This complicates not only the design of machines, but should also be taken into account in diagnosis of the machine.

In fact, resonance and the entire operation are also affected by stiffness of foundations. They are made of concrete for large number of machines and have very low natural frequencies. Foundations made as a combination of reinforced concrete base plate on steel pillars have recently prevailed. Nowadays, combined foundations, to which flexible elements are inserted, sometimes even with a damper, are used. These springs can be tuned after the basements are built and the machine set installed. They are also often used in the construction of bridges and buildings (skyscrapers in seismic areas have these springs controlled by computer to minimize vibration of the building). Figure 4.3.6 shows an example of such system of springs used in the above described machinery. (In technical jargon this system of springs is called (not quite correctly) "gerbs" according to its manufacturer.

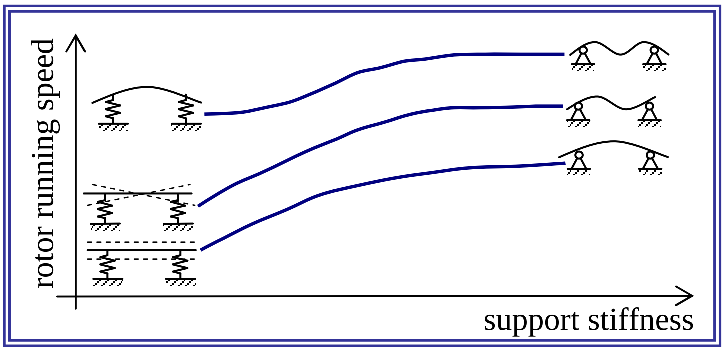


Fig. 4.3.5 - Shape of Vibration as a Function of Support Stiffness

Fig. 4.3.6 - Spring System for Flexible Machine Mounting

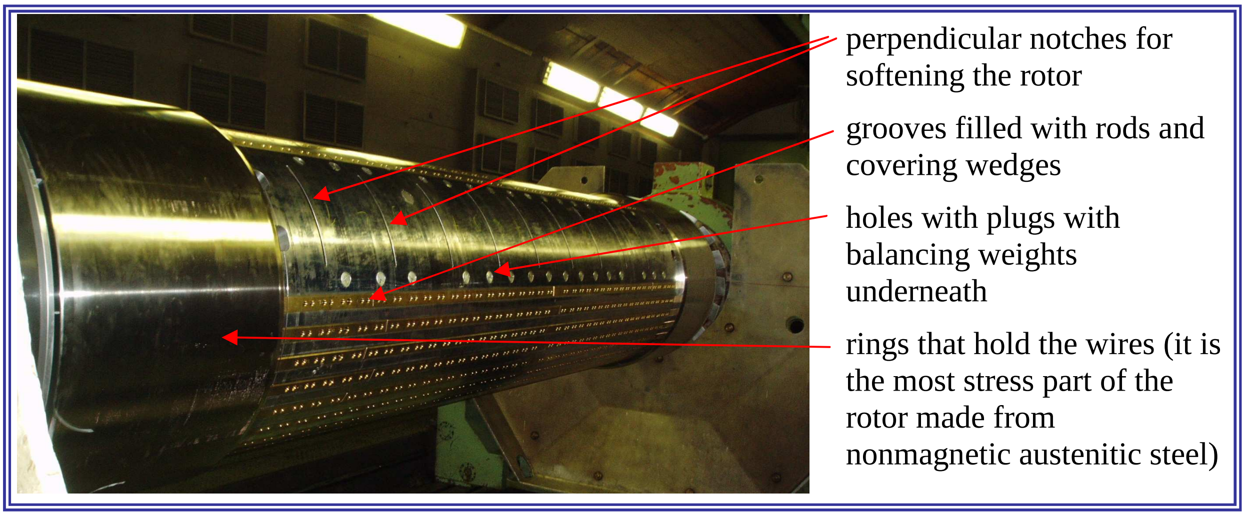


Fig. 4.3.7 - Rotor of Generator

Calculation of critical speed in the example in Table 4.3 and Fig 4.3.4 is, in fact, simplified. The first difference between the computational model and the reality is the fact that the generator rotor in its essence cannot have the same stiffness in two perpendicular transverse directions - over the poles and perpendicular to them. In one direction the rotor is less rigid because there are grooves for wires. This rotor will therefore have two first critical speeds. To prevent them from spreading into a broad frequency band, the stiffness in the direction between rotor poles is deliberately reduced by a number of notches. The rotor generator with transverse relieving notches is shown in Fig 4.3.7.

Idealized plot of the amplitude and phase of a rotor with two first critical speeds (20 and 30 Hz, i.e. 1200 and 1800 rpm) is shown in Fig 4.3.8 above. In fact, a characteristic that significantly differs from the ideal one is often measured - see Fig 4.3.8 below, in which two resonant peaks merge into a single broad peak. Two critical speeds may also result from different cause than the non-uniform stiffness of the rotor is, as described below.

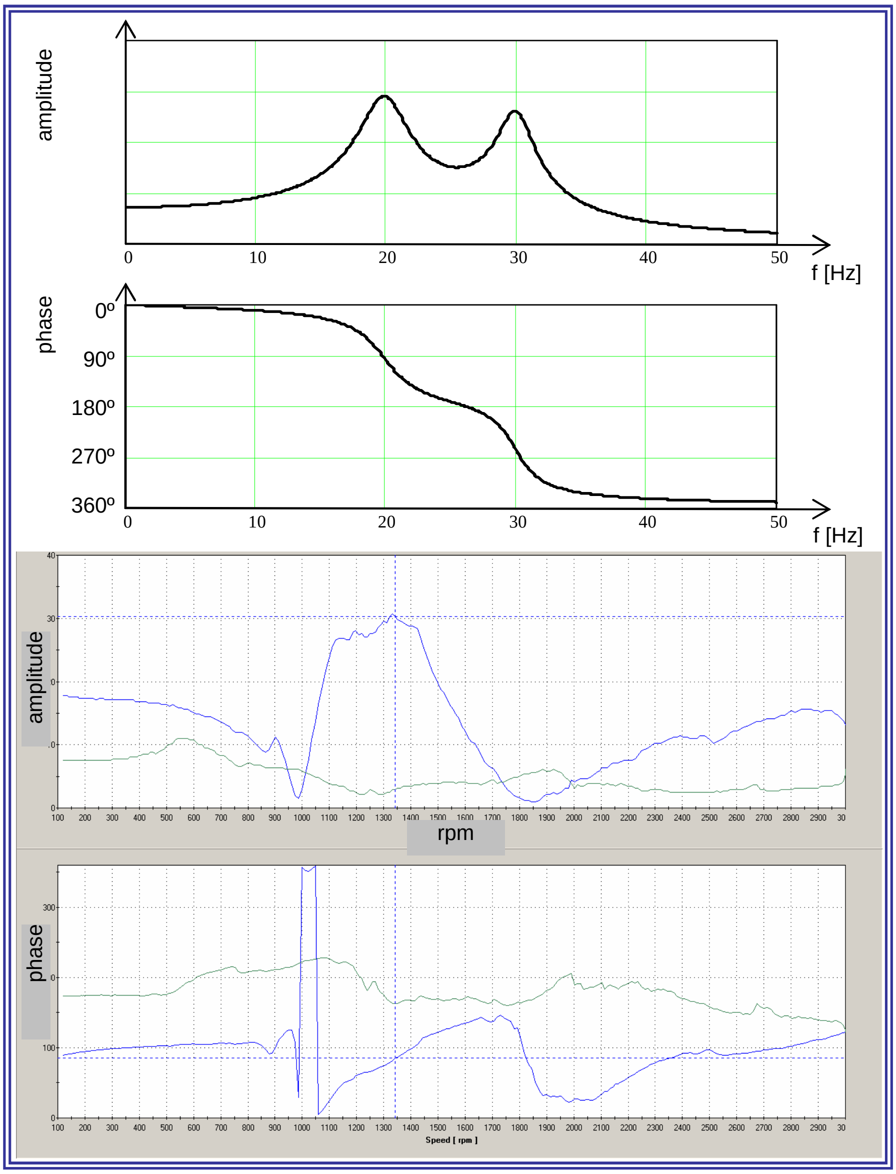


Fig. 4.3.8 - Idealized (above) and real (below) Characteristic of a Rotor with Two First Critical Speeds

The second difference between the computational model placed above and the reality lies in the fact that the supports with bearings do not have the same stiffness in vertical and horizontal directions. This is due to both different stiffness of journal bearings and different stiffness of the stands themselves in these directions.

For example, the 40 MW turbine with the operating speed of about 5500 rpm has the first calculated critical speed 2150 rpm in the vertical direction and 1470 rpm in the horizontal direction when considering different stiffness in vertical and horizontal directions, which is a considerable difference.

Another difference between reality and the computational model is related to different conditions during *run-up* and *coast down* - usually the machine is heated differently, the rate of change of speed and the associated energy storage is different, etc.

All this means that when dealing with a real machine, a certain speed range in which vibrations are increased should be considered as the critical speed. In this range the machine cannot operate for a long time without a risk of damage.

In machine diagnostics, change in critical speed is monitored and evaluated whether the deviation is acceptable, or it may mean developing of a fault - mechanical looseness of parts of the supporting system or a crack in the shaft.

Fig 4.3.9 shows an example of running through the critical speed. The area of critical speed is clearly apparent on the waterfall diagram from the change of 1X component amplitude (excitation by unbalance). It is interesting that there is similar, although much smaller response, also on 2X component (excitation by a small non-linearity) in the same frequency band. This phenomenon is sometimes used to diagnose a possible crack in the rotor (see e.g. [24]).

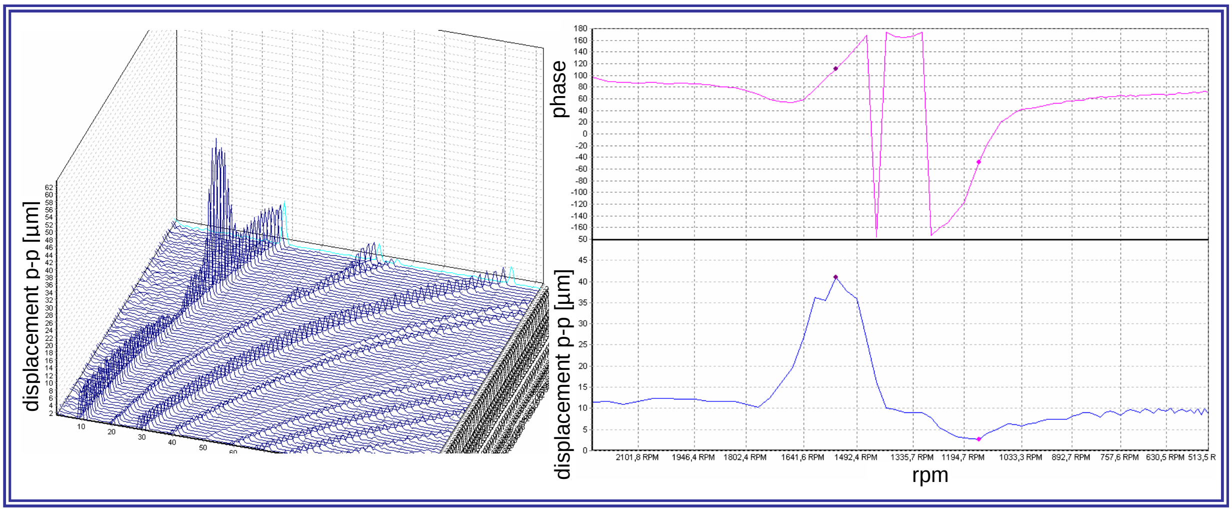


Fig. 4.3.9 - Record of Running through the Critical Speed - Rotor on Journal Bearings

Waterfall diagram provides an overview of the running through the critical speed, but Bode plot is used for easier value readings and for assessment of phase changes. From the above example it is again evident that phase change spreads to a greater extent around the critical speed when a machine is supported on journal bearings. Some "unevenness" of the plot results from the way of software post-processing; it is not record of real changes.

4.3.2 Orbit and Shaft Centerline

Measurements of the relative vibration are primarily used for diagnostics of rotor systems on journal bearings although some faults manifest themselves also in absolute vibration measured on bearing housings.

Chapter 4.3 is focused mainly on symptoms that can be detected by analysis of shaft vibration, namely by analysis of spectra, waterfall diagrams, orbits and centerlines (and sometimes also by analysis of a time waveform).

As mentioned in chapter 2.4.3, the direction of orbit rotation is of diagnostic significance. Under normal circumstances, the orbit is *forward* - the center of the shaft exhibits precession motion around the equilibrium position in the same direction as the shaft rotates.

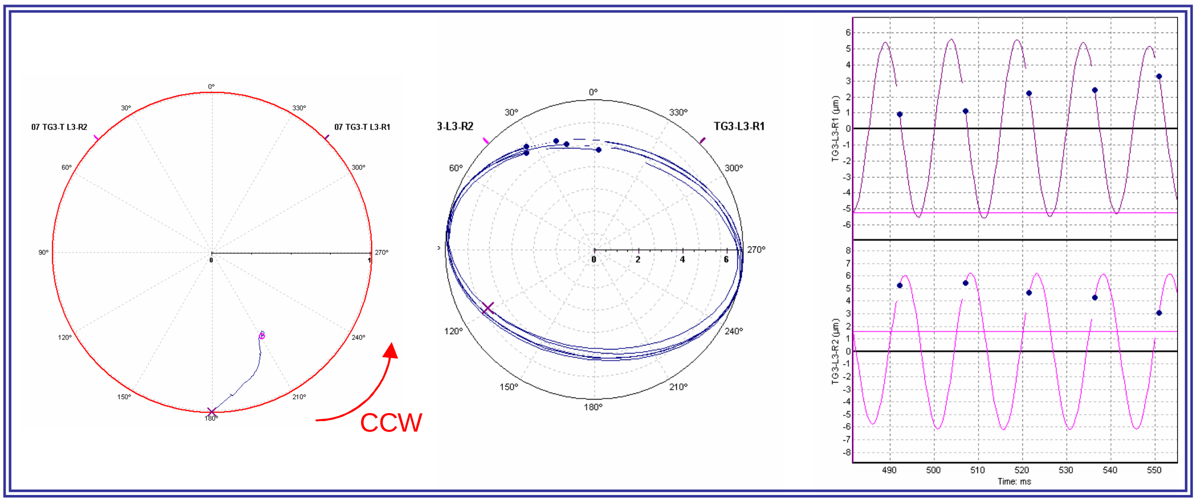


Fig. 4.3.10 - Centerline (left) and Orbit (middle) for Correct Operation

Figure 4.3.10 shows changing position of the center of the shaft during change of running speed - so called *centerline* - and an orbit at operating speed for a system without problems for counter clockwise rotation.

- The center of the shaft is in the stable quadrant - to the right from the vertical axis, which means a well-formed oil wedge.
- The orbit is forward (in this case counter clockwise) and slightly flattened (it is not a circle).

Flattening of the orbit is given by relatively small stabilizing force acting on the rotor. Such load can be either *normal* or *acceptable*: it is the weight of the rotor, the forces from gears, load by process forces (aerodynamic/hydraulic forces), an intentionally designed misalignment in order to compensate deflection of the rotor or to suppress instabilities (whirl), etc. For vertical rotors (especially pumps) such stabilizing force is normally created using the appropriate design of inlet and outlet so that the resultant force would act on the rotor in certain direction.

Load to the rotor may also be *unintentional* and *harmful*. A large lateral force may have various causes: misalignment in the coupling, misaligned machine parts due to thermal expansion, misalignment of bearings, seals, etc. due to improper installation, slight contact of a rotor and a stator, etc. Sometimes the force originates from pipelines connected to the machine case. An excessive force or moment may originate from a pipeline if its flanges are not parallel or if its suspensions are not adjusted correctly, thus causing the pipeline to shift the machine case in radial direction.

Such load is indicated by an *orbit deformation* (significant flattening) and also the *centerline* sometimes moves to a less stable quadrant. Precession (orbit) remains forward. Great change of position of the center of the shaft and significant flattening of the orbit in Fig. 4.3.11 indicates the existence of an excessive loading force. Note: For extra heavy loads, usually when there is significant misalignment of bearings or large misalignment in the coupling, the orbit may be even backward.

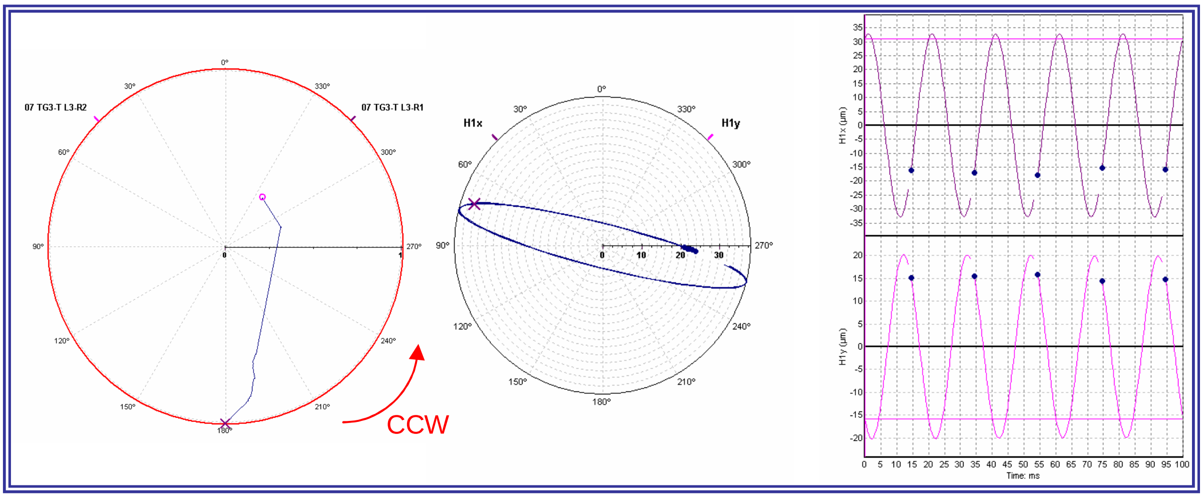


Fig. 4.3.11 - Centerline (left) and Orbit (middle) for Excessive Loading Force

4.3.3 Rotor Rub

A specific machine fault is a slight contact between a rotor and a stator (called *rotor rub*), which usually occurs in seals, especially when introducing a new or repaired machine into operation. Great effort during construction, operation and maintenance of a machine is aimed at preventing rub. There are several types of rub. The shape and precession of the orbit is different in case of operation below and above the critical speed. The orbit may also change from forward to backward. Rotor/stator contact is quite a complex problem that requires detailed knowledge. There are a number of special articles and publications on the subject (e.g. [24]).

An example of a typical orbit and spectrum when a rotor operates with rotor rub is shown in Fig. 4.3.12. In this case, the spectrum usually contains components that correspond to the fact that the motion of the rotor is unilaterally limited. This can be seen at the top waveform that is truncated - the signal is not symmetrical around zero and the top peak is "cut off". As a result, a number of harmonic multiples of rotation frequency will appear in the spectrum and sometimes also so called interharmonic components (1.5X, 2.5X ....) and subsynchronous components (1/2X, 1/3X, 2/3X ...) may appear. Almost any frequencies may occur in the measured spectrum, including resonant frequencies of the rotor and stator.

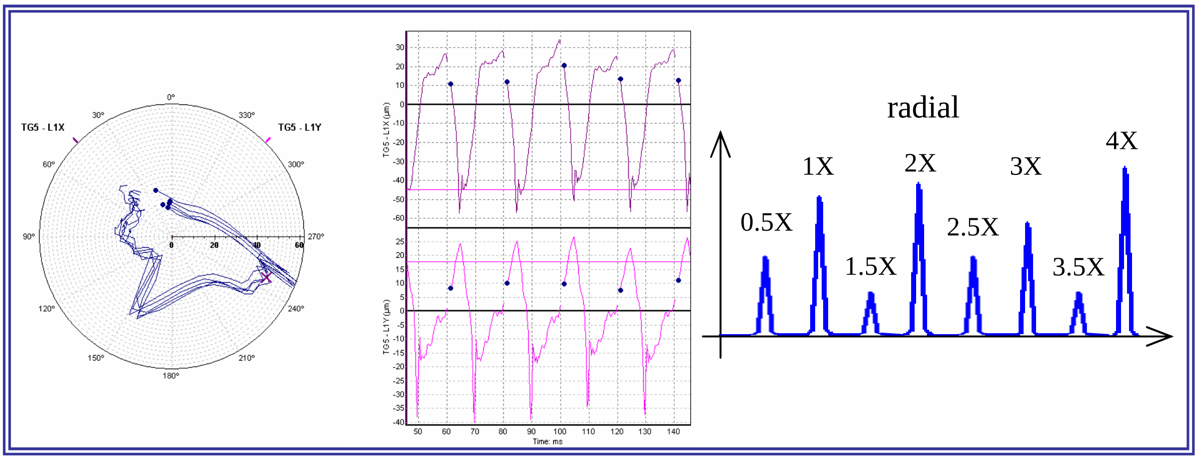


Fig. 4.3.12 - Orbit, Waveforms and Typical Spectrum when Operating with a Slight Rotor Rub

4.4 Journal Bearings

This chapter describes design and operating characteristics of journal bearings that are related to vibration diagnostics of rotating machinery.

Nowadays, design, construction and production of journal bearings is at such a high level (particularly for high-speed machines) that specialized companies are engaged in their production and deliver bearings to machine manufacturers. These companies also publish special publications in which typical journal bearings defects and their causes are described.

4.4.1 Principle of Journal Bearing Operation

Figure 4.4.1 shows the simplest form of a journal bearing. It is called cylindrical bearing. The hole in the bearing is circular and its diameter is larger by the radial clearance. In rest position, when the pin sits in the bottom shell, the side clearance is equal to one half of the vertical clearance. In real constructions, this will be about 1.5 to 2 per mille diameter (for diameter of the pin 200 mm it is about 0.3 to 0.4 mm).

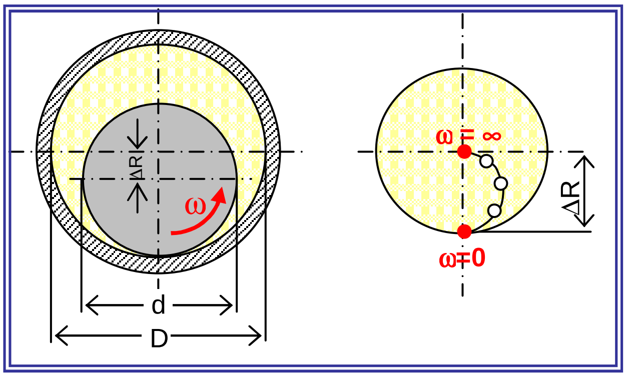


Fig. 4.4.1 - Cylindrical Journal Bearing

Figure 4.4.1 (right side) shows schematically the movement of the pin center in the bearing *clearance* with increasing running speed (be careful - the circle represents the "diametric" bearing clearance only, not the entire bore). Change of position of the center of the pin (centerline) was already discussed in the chapter 4.3.2. Theoretically, the infinite running speed would get the center of the pin onto the axis of the bearing bore. From a diagnostic point of view it is important in this picture that the center of the pin should be located in only one of four quadrants to which the circle can be divided - precisely, when the shaft rotates counter-clockwise it should be the right lower quadrant. This is the quadrant where, according to the theory, rotor operation will be smooth and stable on a well-formed oil wedge.

4.4.2 Principles of the Cylindrical Bearing Construction

- It is necessary to solve the supply and drainage of lubricating oil and prevent its escape "to the side". Supply is usually solved using a wide groove in the upper shell or longitudinal groove in the split line, sealing on the sides is carried out using a special narrow pad with very little clearance.

- Sliding surface cannot be made of steel because the metal contact of a pin and a bearing would cause an excessive heat generation and a damage to both parts by jamming. This is mostly solved by a babbitt (white metal) lining of a certain thickness coated over the steel (or cast iron) bore. Older structures used dovetail grooves for proper connection of the two metals; new structures have finer grooving in the steel liner.

- To reduce losses by friction, the length of cylindrical bearings of older construction is usually greater than their diameter (see Fig 4.4.2). Specific pressure in such bearing is set to be about 0.8 MPa (calculated as the ratio of reaction force from the rotor weight to the projection of the bearing area).

- The bearing itself must be properly fixed in the bearing housing. It has to guarantee transmission of forces while allowing aligning of the bearing to the axis, often with consideration of the rotor deflection.

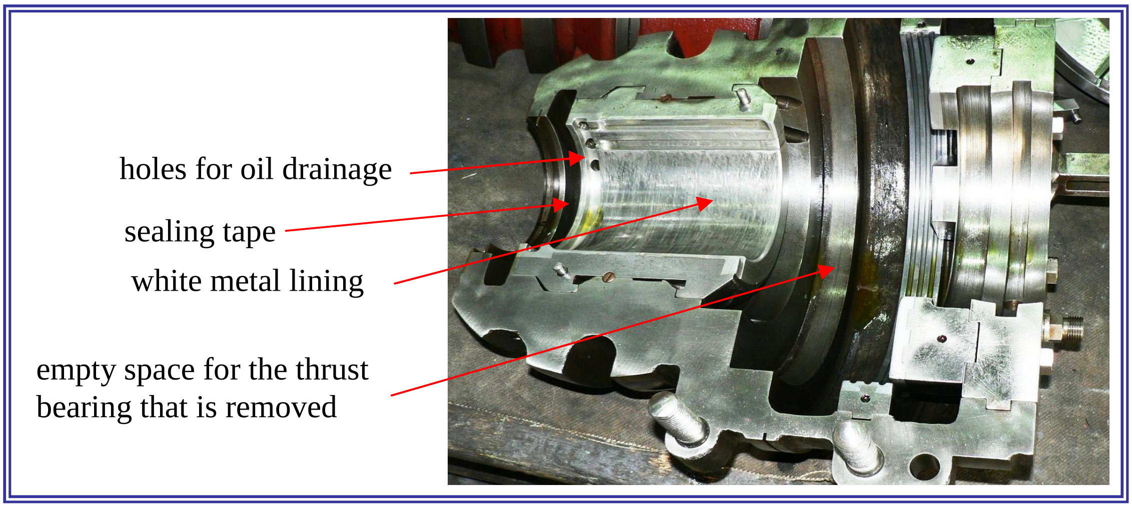


Fig. 4.4.2 - Journal Bearing

Figure 4.4.3 on the left shows a view on the bearing body from Fig 4.4.2 from the outside, so supporting on a ball is visible, which should allow setting of the bearing according to the rotor deflection curve. This method is still used, especially for heavy rotors, when forces transmitted to the pedestal are extensive. An important drawback of this method of bearing support is its very demanding manufacture - the ball must abut on each other in a specified range of the total area - the required contact area according to the manufacturer is more than 80% of the total area. When the contact is not perfect, mechanical looseness occurs that causes increased vibration and is indicated in spectra. Any repair is very difficult. It means repair by milling cut in the split line and scraping the surfaces to fit to each other.

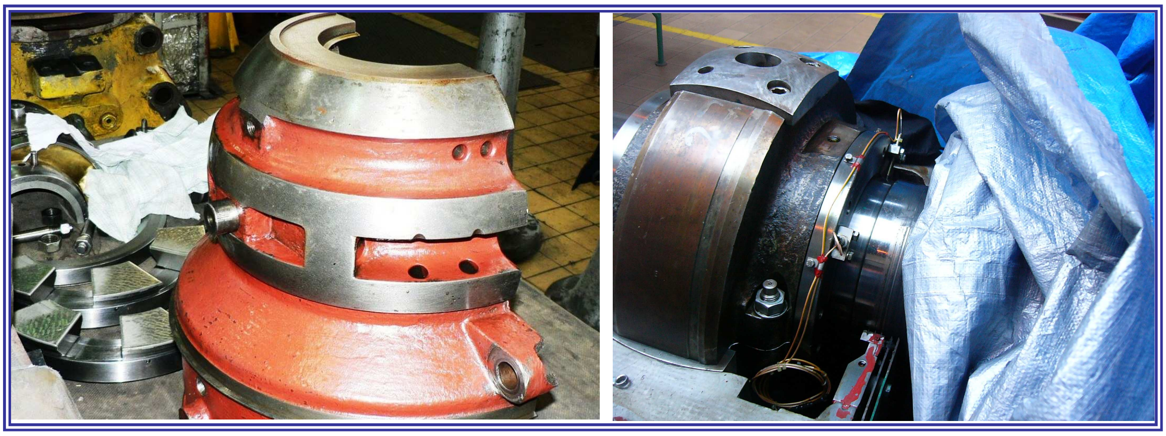


Fig. 4.4.3 - Bearing Support on a Ball (left) and on Stones (right)

Figure 4.4.3 on the right shows another way of bearing support when the sphere surfaces on the bearing are represented only by adjustable parts - *pads* - so that contact between a bearing and a supporting stand is adjusted by metal sheet underlays. Such a method can be applied only for relatively light rotors (to avoid deformation of contact surfaces).

To avoid phenomena associated with mechanical looseness, pre-stress is created between the bearing surfaces of a bearing and a stand - a common interference is about 0.03 to 0.08 mm (prescribed by the manufacturer).

4.4.3 Operational Problems of Cylindrical Bearings and Their Solving

In earlier times, when rotating machinery operated at operating speeds below critical speeds, cylindrical bearings worked well. As the machines were designed to increasingly higher operating speed (in order to improve the process efficiency, size reduction, etc.), it became evident (at circa 1930) that operation at a speed slightly above twice the rotor critical speed is no more possible - significant increase of vibration led to machinery accidents. Extensive research and theoretical analyses of this phenomenon were conducted that resulted in a number of recommendations. For simplicity, commonly accepted criterion containing a combination of design and operating variables will be used for explanation.

The following explanation is focused only on the analysis that can be done by a diagnostician if he encounters the problem of such vibration in a ready-made machine (correct design of a new machine is a matter of the designer). Design and operational parameters were grouped into the so-called *Sommerfeld number* $\mathbf{S_o}$ which includes the following variables:

$$S_0 = \frac{F_{stat} \cdot \Psi^2}{B \cdot D \cdot \eta_{oil} \cdot \omega} \qquad (4.1)$$

where:

F_{stat}	... static load of the bearing
B	... bearing width
$D = 2R$	... bearing diameter
$d = 2r$	... rotor pin diameter
$\Delta R = R - r$	... bearing clearance
$\Psi = \frac{D-d}{d}$	... relative bearing clearance
η_{oil}	... dynamic oil viscosity
ω	... angular rotational speed of the rotor

Essentially, the formula is about:

- The combination of reaction force F_{stat} (force from rotor weight to the bearing) and the geometric dimensions of the bearing B and D, which is nothing else than the specific pressure in the bearing (one characteristic value has been given for cylindrical bearings: about 0.8 MPa).
- Relative clearance in the bearing (the ratio of the actual clearance to the diameter) - it is about 0.15 to 0.20 % of the diameter. It is of great importance as it is in square in the formula.
- Viscosity of oil used for lubrication. Be careful: The viscosity depends not only on oil grade but also on its temperature.
- Angular speed (rpm) of the rotor - for most machines it is a constant quantity.

What is it for and what a diagnostician should be careful about will be explained in combination with the plot shown in Fig 4.4.4. The plot is based on a series of measurements on the basis of which two regions were determined - a region of stable operation (below the limit curve) and a region of unstable operation (above the limit curve). Limit rotor speed (in dimensionless form) is on the vertical axis.

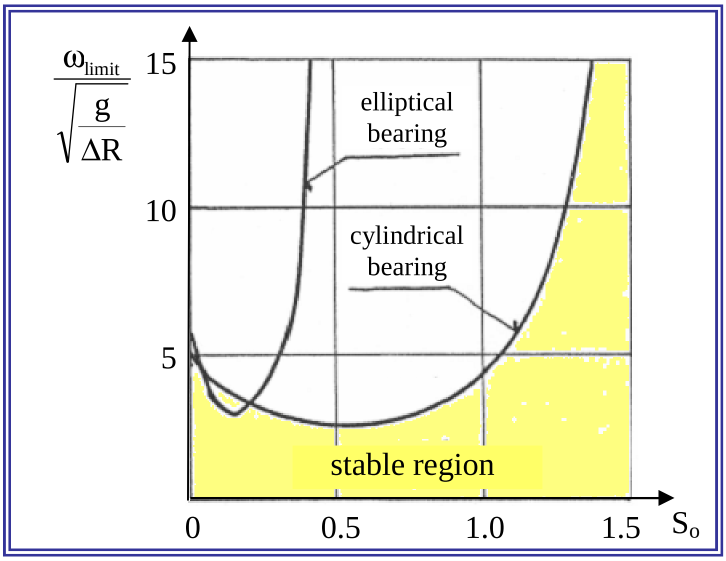


Fig. 4.4.4 - Limit Speed of a Rigid Rotor as a Function of Sommerfeld Number

There are two limit curves in the plot - one for cylindrical bearings and one for elliptical bearings. First, the curve given for the cylindrical bearing will be considered - its stable region is coloured yellow. It can be seen that to achieve stable operation at the given speed, the point representing a combination of Sommerfeld number and the speed should lie in the stable region below the limit curve, which means that the Sommerfeld number should be sufficiently high.

How can a designer achieve this situation?

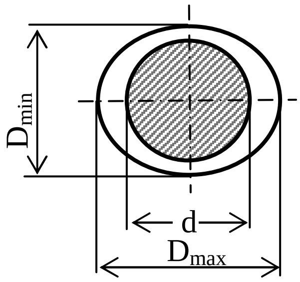
- It is advisable to use greater specific pressure in the bearing (it is usually 1.5 MPa, but may be up to 3.0 MPa in extreme cases) - priority is given to the stable operation rather than to reduction of friction losses. Because the diameter of the pin depends on the design of the entire rotor, such a bearing is usually recognized by being "short" - its length is smaller than its diameter.

And what about reality?

- This specific pressure may seem to be an unchanging quantity. The weight of the rotor is the same (hence the reaction force should be the same), the bearing area is given by construction. But when realized that most machine sets have at least four bearings, which is statically an indeterminate case, then it may happen - when a coupling is not aligned properly - that some of the bearings become unloaded. If a rotor is lifted from one bearing by a coupling, then adjacent bearings take its load. The reaction force and thus the specific pressure decreases in the considered bearing, which may lead to operation close to the stability limit or at worst to its exceeding.
- It is advisable to use larger clearance, but then problems with leakage of oil from the bearings may arise, so the size of the clearance in modern bearings is not significantly different from those previously stated (1.5 - 2 per mile diameter). But a problem may occur during repairs when a technician deliberately sets too small clearance (especially when bearing is repaired) to delay the need for another demanding repair of the bearing during next inspection.
- Viscosity: Given that the viscosity of oil in the Sommerfeld number figures in the denominator, thick (more viscous) oil is "harmful" in terms of stability of operation. If the correct oil is used, the only danger is its low temperature (when the viscosity is greater). Most of operational rules therefore contain the command: *Do not start-up the machine set until the oil temperature reaches (e.g.) 36° C.*
- Note: For outdoor use, this value is sometimes lower. Modern systems include heating equipment for heating oil in a tank.

4.4.4 Elliptical Bearing

There is also a second curve in Fig 4.4.4 which shows that significant increase in stability can be achieved with different bearing design that is called an elliptical bearing (sometimes "lemon" bearing). Its design (see Fig 4.4.5 left) contributed greatly to the possibility of increasing the operational speed of machinery above the critical speed.



$$\Psi_{min} = \frac{D_{min} - d}{d}$$

$$\Psi_{max} = \frac{D_{max} - d}{d}$$

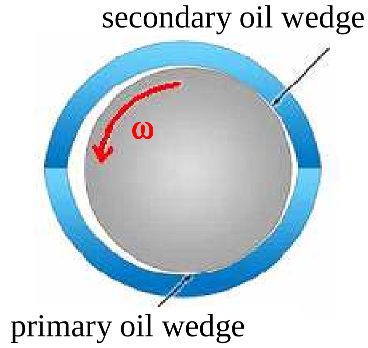
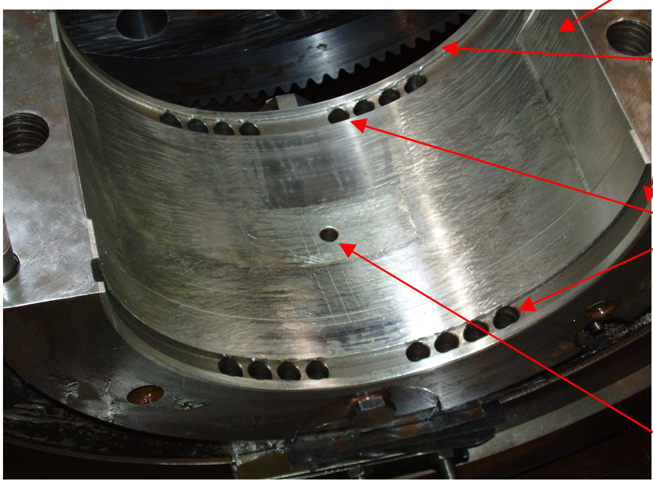


Fig. 4.4.5 - Elliptical Bearing - Design, Constitution of Oil Wedge

- The main difference from the cylindrical bearing lies in the fact that the pin is stabilized not only by the oil wedge that is formed in the lower shell, but also the upper shell is used which in the cylindrical bearings was used only to close and seal the bearing and for oil supply. This bearing works as a two-wedge bearing.

In order to properly constitute an oil wedge towards the upper shell, it is necessary to change its geometry. This is done so that when machining the bore, two shims of thickness specified by the designer are inserted into the split line between two halves of a round sleeve. After round machining, the two shims are removed. It results into elliptical shape of the bearing after installation.

This means that the internal shape is changed - it is not purely circular any more, and also the ratio of lateral to vertical clearances (Ψ_{max} a Ψ_{min}) is changed. This ratio for example 2:1 means that in rest position, clearances on both sides of the pin are as large as the vertical clearance.



"Pouches" for lubricating oil supply in the split line.

There are sealing tapes with reduced clearance at both ends of the bearing (in the axial direction), preventing excessive oil leakage from the bearing.

From these tapes towards the center, there are grooves with holes through which oil is drained away.

In the middle of the lower shell area, there is equipment for preventing the shell from damage caused by a heavy rotor during slow rotation (when the hydrodynamic wedge is not yet created) from the heavier rotor. Through this hole, the so-called lifting oil of high pressure is supplied at that time. This serves only before spinning rotor to the operating speed.

Fig. 4.4.6 - View on a Typical Lower Shell of an Elliptical Bearing

- Another change is the increase of the specific pressure in the bearing - elliptical bearing is visibly shorter than cylindrical bearing with respect to the diameter.

- Another change relates to the use of the upper shell for constituting an oil wedge (see Fig 4.4.5 right). Oil is supplied in the split line, which completes the appearance of a "lemon".

4.4.5 Other Types of Radial Bearings

For the needs of different types of machines and for even higher ratio of operating speed to critical speed, other bearing designs have been developed that are used in practice. They are schematically shown in Fig 4.4.7.

Sometimes "offset" shells of a cylindrical bearing are used (Fig. 4.4.7 left). This bearing only works in one direction of rotation. It is used e.g. for high-speed gearboxes.

When increasing the running speed, more areas on which oil wedge is formed can be used to stabilize the pin. In the middle in Fig 4.4.7 a three-pad bearing is shown. However, bearings with multiple pads are much more difficult to be machined and also their repairs are not easy. They find practical application in small machines with very high speed (about 60000 rpm).

Nowadays, high-speed machines use *tilt-pad bearings* (see Fig 4.4.8 right) which eliminate the problem with machining, but still they are quite complex and expensive. The principle consists in splitting the sliding surfaces into several independent parts that tilt under load in order to create the necessary supporting oil wedge. Four or five-pad bearings are most common.

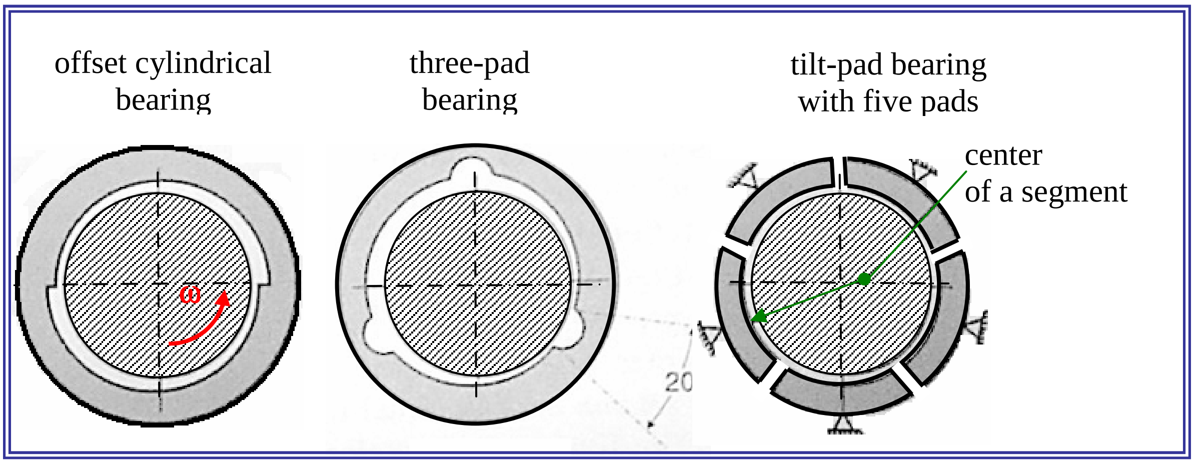


Fig. 4.4.8 - Other Types of Journal Radial Bearings

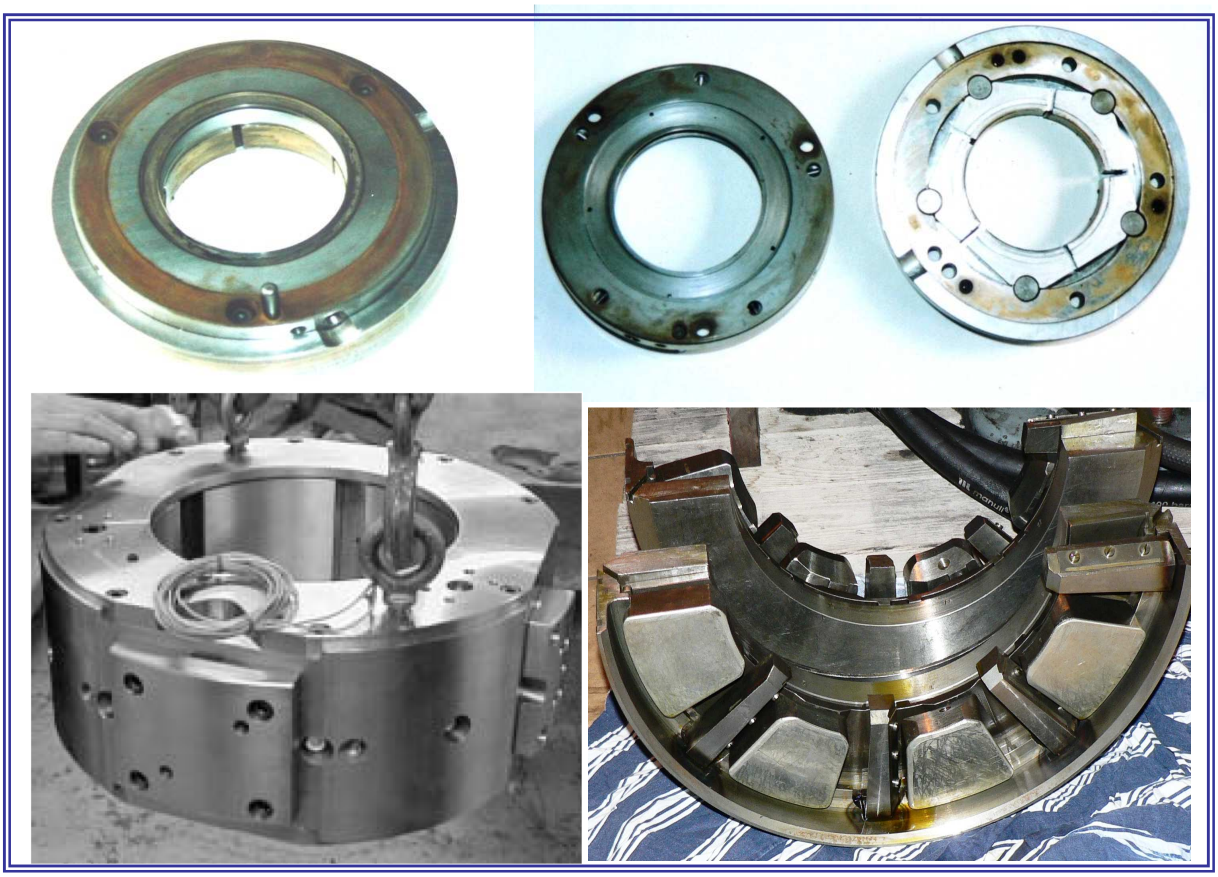


Fig. 4.4.9 - Examples of Tilt-Pad Bearing Designs

Fig 4.4.9 shows three examples of design. The tilt-pad bearing of a high-speed compressor (approximately 16 000 rpm) is on the top. The bearing has 5 pads (segments) supported on pins. Pins are placed out of the symmetry axis of the sliding surface to permit tilting under load and forming the proper wedge. An example of a modern tilt-pad bearing of a turbine is in the left bottom. Moreover, external adjustment "stones" and wires from a thermocouple that monitors the temperature of each segment (which gives information about segment load) are visible.

To complete the issue of bearings, on the bottom right of Fig 4.4.9 a very important machine part is shown - the *thrust bearing* that transmits axial loads. These bearings are almost always constructed with tilting pads (the so called Mitchell's design) and are double-sided, i.e., designed for both directions of axial forces on the rotor. Gearbox bearings are an exception, especially those for gearboxes with arrow teeth. They often use simpler design of thrust bearings (not shown here).

4.4.6 Faults of Journal Bearings - Wear, Excessive Clearance

Wear of the bearing causes an excessive clearance, which is usually reflected by the presence of harmonics of rotational frequency, particularly in the absolute vibration spectrum, see Fig 4.4.10 left. On the contrary, in the relative vibration spectrum, a radial bearing with increased clearance will usually exhibit large 1X amplitude without the existence of harmonics, as shown in Figure 4.4.10 right.

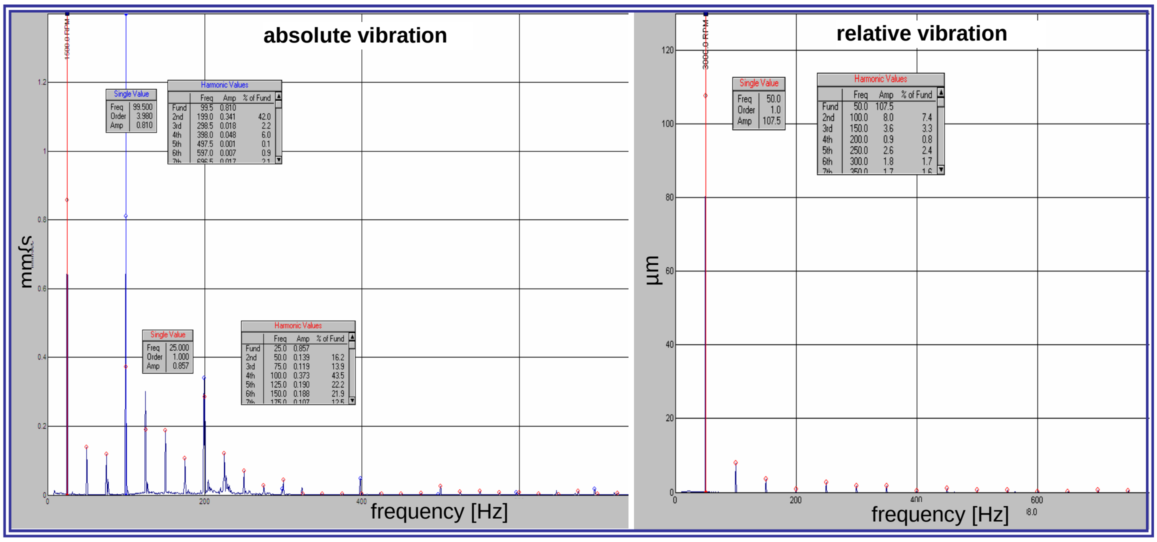


Fig. 4.4.10 - Vibration Velocity Spectrum (left) and Relative Displacement Spectrum (right) of a Journal Bearing with Excessive Clearance

Symptoms of an excessive clearance in the bearing are similar to that of mechanical looseness. Bearing with correct clearance, but with loosed contact to the supporting structure, manifests itself similarly in vibration as a bearing with excessive clearance. Therefore, it is often difficult to distinguish between these two cases.

4.4.7 Operational Problems of Machines Supported on Journal Bearings

Operational problems with vibration of machines supported on journal bearings can arise for various reasons:

- Incorrect assembly (machine alignment) that results in reducing the load on any bearing.
- Inappropriate temperature and, thus, also oil viscosity (it was discussed in relation to Sommerfeld number).
- In some operation modes, forces may occur that cause relieving of a bearing and its moving closer to the stability limit or overrunning this limit. This results in an unstable operation due to *oil whirl*. When the oil whirl frequency is the same as rotor natural frequency, rotor resonance which is called *oil whip* occurs.

4.4.7.1 Oil Whirl

This instability appears at subsynchronous frequency of about 0.40 to 0.48X and is often quite strong. Oil whirl is a case when the oil film causes the subsynchronous precession component of the rotor motion - oil wedge "pushes" the rotor around the shaft in the bearing with a frequency that is lower than the rotational frequency; the precession is forward.

The example in Fig 4.4.14 is from a turbine that approached the stability limit of one of its bearings due to combination of the above mentioned causes and a small internal damage of labyrinth seals. There is a typical subsynchronous component under the half of rotational frequency in the spectrum. When this instability arises, this component is very unstable and it is recommended for on site diagnostics to monitor subsequent instantaneous spectra as this phenomenon can be suppressed by averaging. The waterfall diagram clearly demonstrates this phenomenon. The orbit in this case is unstable and two keyphasors per revolution can be seen both at the orbit and on the waveforms. These keyphasors will gradually shift, suggesting that the subsynchronous frequency is not half, but a little less than half the rotational speed. This phenomenon is affected also by change of oil viscosity (by changing its temperature) and by change of lubrication pressure.

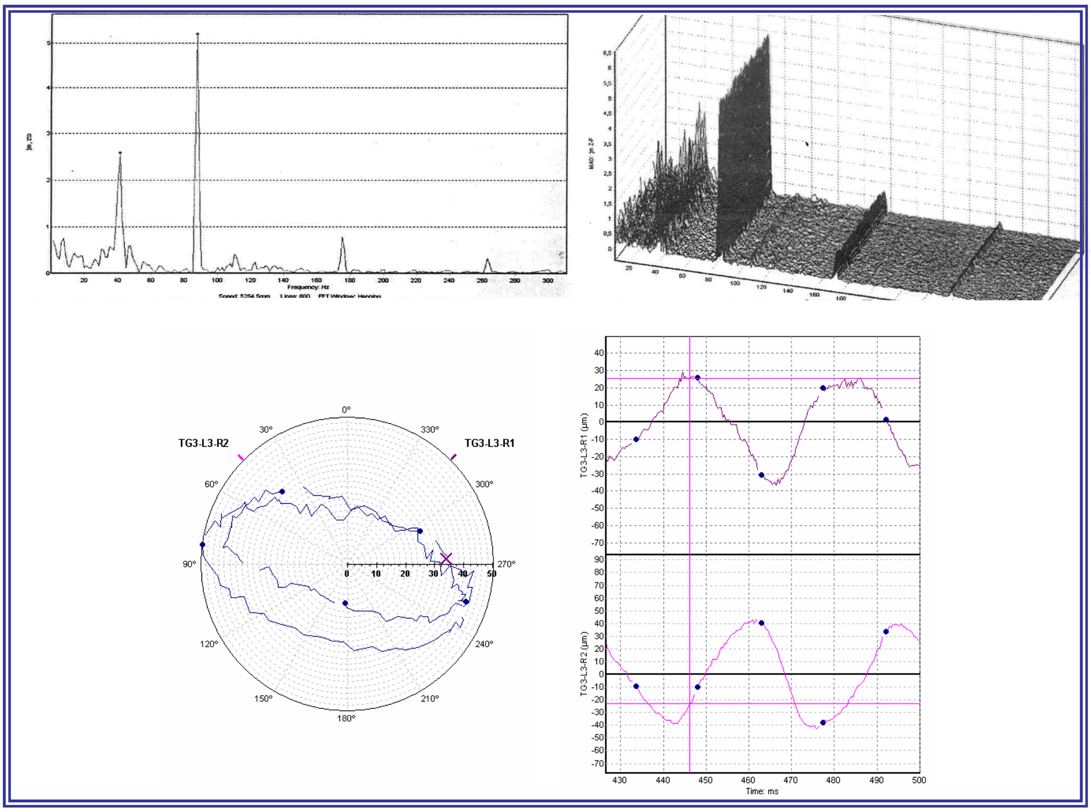


Fig. 4.4.14 - Instantaneous Spectrum, Waterfall Diagram and Orbit while Oil Whirl

4.4.7.2 Oil Whip

This instability can occur when a machine is operated at above twice the rotor critical speed. When the rotor spins to twice the critical speed, the oil whirl frequency can be close to the rotor critical speed and, thus, can excite the resonance and cause excessive vibration. This instability causes transverse subharmonic vibration with frequency equal to the rotor critical speed during forward precession. This is an unstable process that can lead to a catastrophic failure. Instability is in fact "locked up" on the rotor critical speed and persists even when the speed is further increasing. The frequency peak pertinent to oil whip remains in the spectrum and can be easily recognized as it does not change with changing rotational speed (unlike oil whirl frequency); see Fig 4.4.16. If the orbit is measured during this period, a lot of marks wandering around can be seen.

Note: Due to the fact that this instability occurs especially when bearings are relieved from load, the orbit is almost circular before this unstable operation begins. Figure 4.4.15 shows an example of an orbit which should warn the diagnostician about oncoming problems.

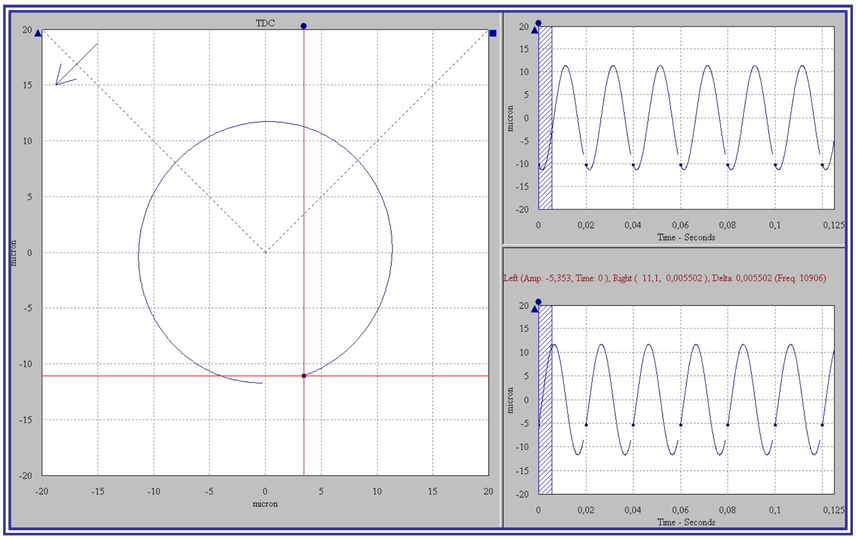


Fig. 4.4.15 - Circular Orbit Indicating the Oncoming Unstable Operation

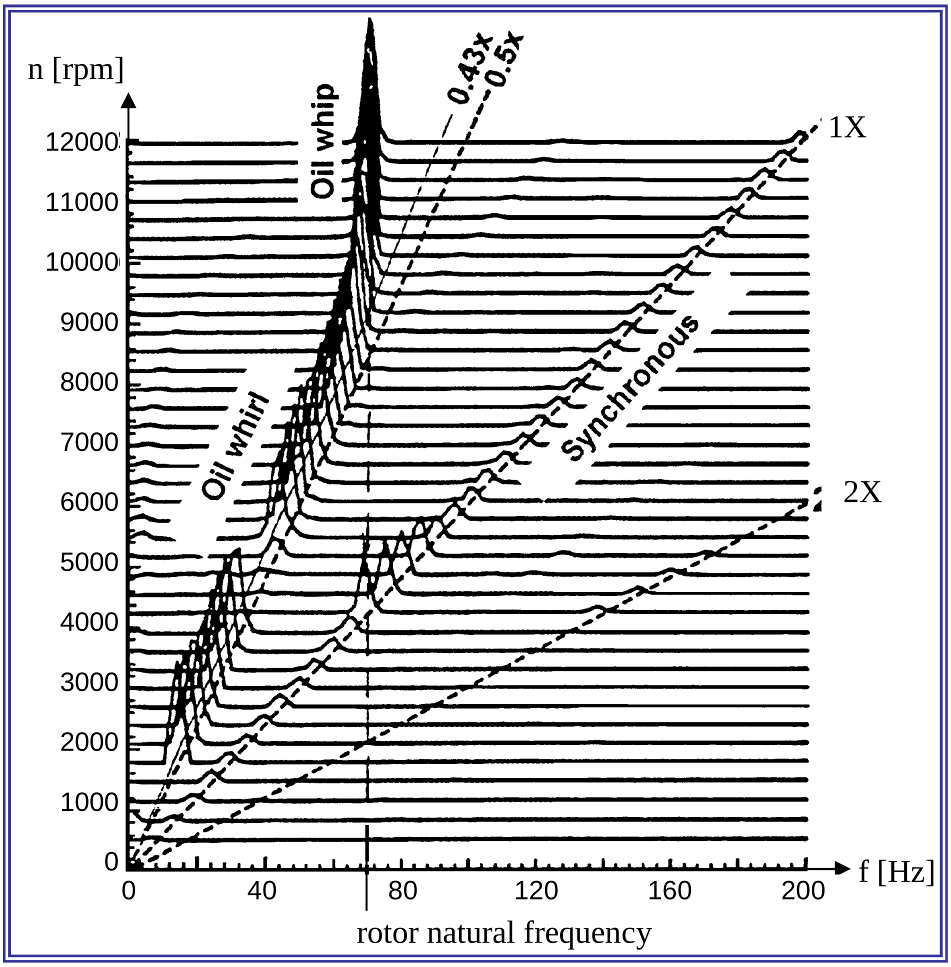


Fig. 4.4.16 - Instabilities of Oil Film

Increasing the operating speed of machines above the critical speed has led to knowledge that the earlier satisfactory concept of an oil film instability using the Sommerfeld number is not valid any more. The issue of unstable behaviour of a rotor on journal bearings is discussed in detail by number of researchers such as Donald Bently and Agnes Muszyńska of Bently Nevada corp. (see [24]). Fig 4.4.16 shows the result of their long-term experimental analyses. These studies have proved that the instability must be addressed for the entire system of bearings with the rotor and that its symptoms vary depending on operating speed, critical speed and the influence of unbalance is significant as well.

4.5 Rolling Element Bearings

A great deal of machines is equipped with rolling element bearings. The basic bearing function is to transfer forces from rotating parts to the construction and to reduce friction in the system. In almost all cases, bearings are the most precise machine parts, generally made with tolerances that are ten times lower than tolerances of other machine components. However, only about 10 to 20% of bearings reach their design life because of different factors that reduce their life. These include particularly imperfect lubrication, using improper lubricant, contamination with dirt or other foreign particles, improper storage out of shipping containers, moisture intrusion, false brinelling (imprint of rolling elements to the race) during transport or when the machine is out of operation for a long time, using an inappropriate bearing for the particular purpose, incorrect installation of bearings, etc.

But main causes for early bearing failures are excessive vibration and high dynamic loads which can thus be transmitted to the bearings. Theoretical lifetime of rolling bearings is a function of the cube of the load to which the bearing is exposed. If the care is taken to eliminate the unfavourable external influences, such as unbalance, misalignment, problems of drive belts, soft feet, inadequate lubrication and incorrect installation, then the bearings should have adequate durability.

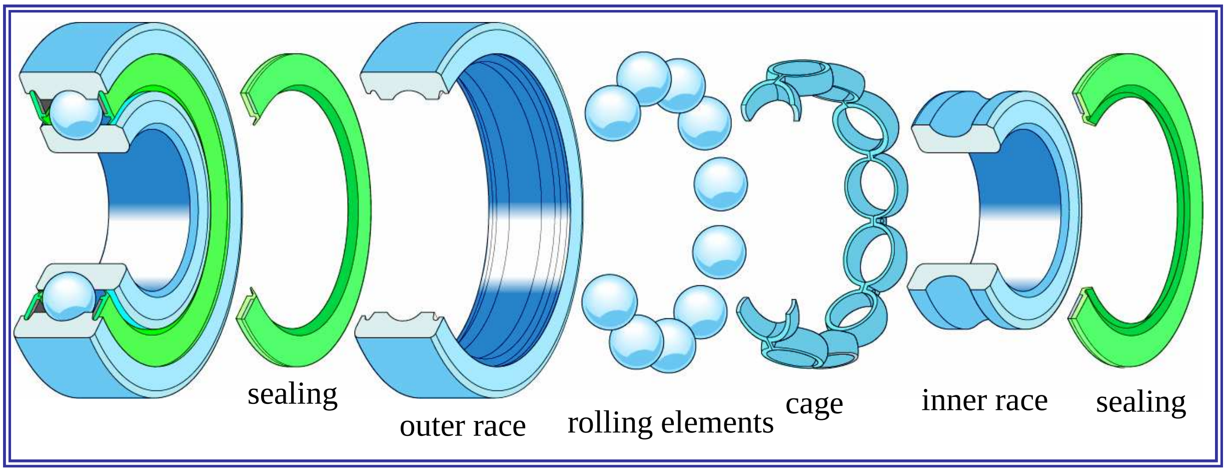


Fig. 4.5.1 - Rolling Bearing Components

4.5.1 Rolling Bearing Design

This text deals with rolling bearing design only to the extent that is necessary to understand and diagnose different types of bearings faults. Individual components of the rolling bearings are shown in Fig 4.5.1. Bearings are further divided by the type of rolling element and thus by the nature of the transmitted forces. Common types of rolling elements with schematic representation of transmitted forces are listed in Fig 4.5.2.

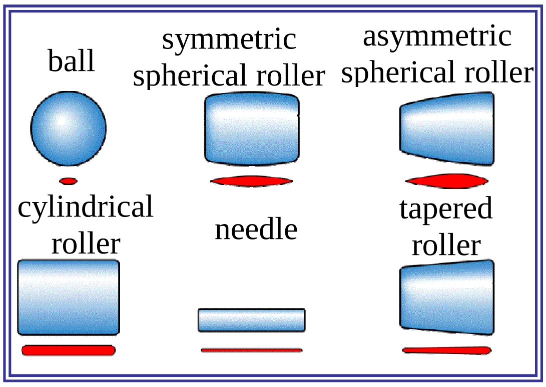


Fig. 4.5.2 - Common Types of Rolling Elements

4.5.2 Parameter for Assessing Rolling Bearing Condition

Condition monitoring of rolling element bearings and determining when they would need replacement are of great importance in terms of machine operation. When a rolling bearing becomes damaged, the vibration signal transmits to the stator part where it can be measured using accelerometer. However, for proper evaluation of rolling bearings condition one can rely neither on measuring the overall vibration only, nor on measuring the broadband value in the ultrasonic range (see chapter 4.5.3.1). When deciding which of the vibration parameters (displacement, velocity or acceleration) will be used to assess bearings condition, it is appropriate to realize the following:

Displacement - Since the displacement is significant at low frequencies, it tends to suppress or eliminate almost the entire spectral content that indicates bearing faults. Therefore, it is not used for assessment of rolling bearings.

Acceleration - Unlike the displacement, acceleration tends to over-emphasize most of the frequency content generated by rolling bearing faults. The result is that the acceleration spectra may cause a false alarm. Although acceleration is a better indicator of bearing problems in early stages, it is more suitable to use vibration velocity to evaluate the fault that is already developed. Vibration velocity more clearly indicates "the truth" about bearing condition. Acceleration spectra can detect bearing problems in earlier stages of the fault than velocity spectra, especially for high-speed machines. In addition, the envelope demodulated high-frequency spectra can provide warning of bearing wear or lubrication problems even earlier and, therefore, are widely used nowadays.

Velocity - Velocity spectra are one of the best parameters for evaluation of most rolling bearings problems. Generally speaking, the velocity remains "flat" (see Fig. 1.13) in the frequency range from 10 to 2000 Hz. This means that when the bearing defect frequency appears at 100 Hz or at 1000 Hz, the same weight can be used for evaluation.

4.5.3 Types of Vibration Generated by Defective Rolling Bearings

Defective rolling bearings generate three types of frequencies when a fault is developing, namely:

- random ultrasonic frequencies
- natural frequencies of bearing components
- bearing fault frequencies (depending on rotational speed)

4.5.3.1 Random Ultrasonic Frequencies

Measurement of ultrasonic frequencies from about 5000 Hz to 60 000 Hz involve measurements of *spike energy*, measurements of *spectral emitted energy* (SEE), measurement of high-frequency *acceleration spectral density* (HFD), measurement of *shock pulses* and others. Each of these techniques is considered as a parameter to detect an emerging fault. Generally speaking, a number that these methods give is only one of information that must be considered when assessing the condition of a rolling bearing.

4.5.3.2 Natural Frequencies of Installed Bearing Components

Natural frequencies of a rolling bearing are usually in the range from about 500 to 2000 Hz. If the bearing is defective, these natural frequencies are excited by periodic impacts of rolling elements to defects on the rolling traces and can be detected. When the wear is getting worse, sidebands appear around these resonant frequencies at intervals of rotational frequency or at distance equal to the bearing defect frequency.

4.5.3.3 Bearing Defect Frequencies

Over the years, many formulas have been derived that can help to detect specific defects in rolling element bearings. They are based on the geometry of the bearing, the number of rolling elements and the rotational frequency of the bearing.

Four types of faults are distinguished on the rolling bearing depending on where the fault occurs. The so called *bearing defect frequency* that can be calculated on the basis of bearing parameters and rotational frequency corresponds to each of these defects:

BPFI - Ball Pass Frequency Inner (defect on the inner race)

$$BPFI = \frac{N}{2}\left(1 + \frac{B_d}{P_d} \cdot \cos\varphi\right) \cdot n$$

BPFO - Ball Pass Frequency Outer (defect on the outer race)

$$BPFO = \frac{N}{2}\left(1 - \frac{B_d}{P_d} \cdot \cos\varphi\right) \cdot n = N \cdot FTF$$

BSF - Ball Spin Frequency (defect on a rolling element)

$$BSF = \frac{P_d}{2B_d}\left(1 - \left(\frac{B_d}{P_d} \cdot \cos\varphi\right)^2\right) \cdot n$$

FTF - Fundamental Train Frequency (defect on the cage)

$$FTF = \frac{1}{2}\left(1 - \frac{B_d}{P_d} \cdot \cos\varphi\right) \cdot n$$

where:

n ... rotational speed [Hz]

N ... number of rolling elements

B_d ... diameter of a rolling element [mm]

P_d ... pitch diameter

φ ... contact angle

Note: The above formulas apply to standing outer ring. In the case of a rotating outer ring, signs in calculations are inversed (except for the formula for rolling elements). Note that each of the bearing defect frequencies is given as a *multiple of the rotational frequency*. The severity of these equations is that they give the possibility to detect problems that occur on races, cage or rolling elements and to monitor these problems during their getting more severe.

To facilitate the diagnostician's work, catalogues and electronically available tables of parameters of individual faults for various types of bearings are at disposal, which, after multiplying by current rotational frequency, provide frequencies that are looked for in the spectra. See example at Table 4.4.

Table 4.4 - Table of Parameters of Bearing Defect Frequencies

Bearing	BPFI	BPFO	BSF	FTF		Pd	Bd	N	angle
N202	6,03	3,97	2,33	0,40		24,30 mm	5,00 mm	10	0 °
N203	5,98	4,02	2,45	0,40		28,00 mm	5,50 mm	10	0 °
N204	5,97	4,03	2,48	0,40		33,50 mm	6,50 mm	10	0 °
N205	7,01	4,99	2,88	0,42		38,50 mm	6,50 mm	12	0 °
6203	4,95	3,05	1,99	0,38		28,50 mm	6,75 mm	8	0 °
6204	4,95	3,05	1,99	0,38		33,50 mm	7,94 mm	8	0 °
6205	5,41	3,59	2,36	0,40		39,04 mm	7,94 mm	9	0 °
6206	5,43	3,57	2,31	0,40		46,00 mm	9,52 mm	9	0 °
6207	5,43	3,57	2,30	0,40		53,50 mm	11,11 mm	9	0 °
6208	5,42	3,58	2,34	0,40		60,00 mm	12,30 mm	9	0 °
6208E	4,93	3,07	2,04	0,38		60,40 mm	14,00 mm	8	0 °
7208C	7,75	5,25	2,43	0,40		60,00 mm	11,91 mm	13	15 °
7209CC	8,34	5,66	2,47	0,40		65,00 mm	12,70 mm	14	12 °

Some interesting facts can be said about the bearing defect frequencies:

1. How do the bearing defect frequencies differ from other defect frequencies?

One of the factors that distinguishes bearing defect frequency from other sources of vibration is that they are frequencies of existing defects. In other words, if a defect does not exist, bearing defect frequencies are not present. When they are present, it is the information that there is a problem arising. Other common frequencies, such as 1X rotational frequency, blade pass frequency in pumps, gear mesh frequency, etc., are always present and their presence does not necessarily mean that there is a fault or a problem. The presence of bearing defect frequencies sends a message "be careful" to the diagnostician. It is important to emphasize that the presence of such frequencies does not necessarily mean that there are defects directly in the bearing. They can also occur when insufficient lubrication allows the metal contact or when the bearing is loaded incorrectly.

2. Bearing defect frequencies are not integral multiples of the rotational frequency

Bearing defect frequencies are not harmonics of the rotational frequency. They represent one of the few vibration sources that generates non-integral multiples of the rotational frequency.

Fig. 4.5.3 illustrates how the bearing defect frequencies are generated in the bearing:

A *fault on the outer race* on the bottom of the bearing in the load zone (red dot) generates a pulse on the waveform always when the rolling element passes over the defect and hits it (ideally the impulses are of the same magnitude).

When a *defect is on the inner ring* (blue dot), the pulse occurs on the waveform each time when the inner ring passes over each rolling element (assuming that the inner ring is pressed onto the shaft). An important fact, which is shown in Fig. 4.5.3, is that the magnitude of the response from rolling elements that hit the fault on the inner ring depends on the position of the inner ring at the moment when the hit occurs. This means that if a defect on the inner ring is in the loaded zone, it will have significantly greater response than if the same defect is in the unloaded zone. This explains why the frequencies of inner ring defects are often surrounded by sidebands at intervals of about 1X - their amplitude is modulated by the rotational frequency.

Fault on a rolling element (green dot) generates a pulse at each contact both with inner and outer ring. Magnitude of the impulse depends again on whether the contact occurred in the loaded or unloaded zone.

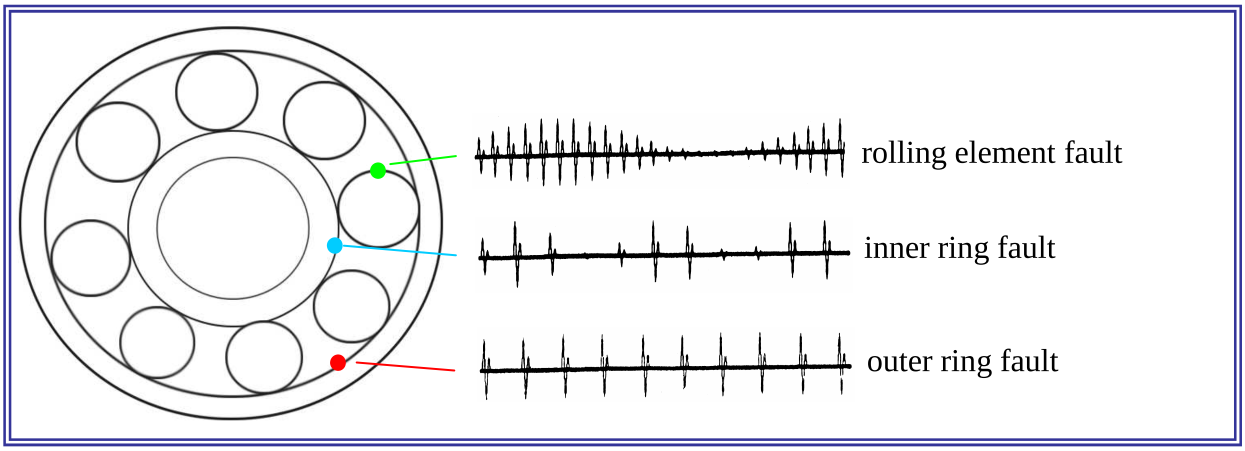


Fig. 4.5.3 - Bearing Defect Frequencies Generation

4.5.3.4 Permitted Vibration Values of Bearing Defect Frequencies

It is very difficult to determine the amount of vibration that is allowable for bearing defect frequencies in a similar way as it is for the 1X amplitude of unbalance. No absolute answer can be given. It depends both on the type of machine and bearing and on the way of the defect development. A key indication of significant damage of the bearing is often the presence of harmonic multiples of the defect frequencies, especially if they are surrounded by sidebands at intervals of 1X or at intervals comprising another bearing defect frequency.

4.5.4 Stages of Rolling Bearing Fault Development

It was found that most of the rolling bearings follow the well predictable course to failure from the early beginning to the eventual catastrophic failure. This course of fault development is illustrated in Fig. 4.5.5 where bearing damage evolution in time is drawn. Note the important fact that the bearing damage typically develops exponentially over the last 10 to 20% of its lifetime. The course of damage that consists of four stages is applicable to about 80% of rolling bearings failures.

Development of rolling bearing faults can be divided into four stages (see Fig. 4.5.4):

I. The earliest indications of bearing problems including poor lubrication occur in the ultrasonic frequency range from about 250 kHz to 350 kHz. Later, when the damage begins, the frequency drops to about 20 to 60 kHz. These are frequencies that require ultrasonic measuring equipment.

II. Small defects in the bearing begin to "ring" - they excite the natural frequencies of the bearing components, which are mainly in the range from 500 Hz to 2 kHz. These may also be resonances of supporting parts of the bearing. At the end of the 2nd stage sidebands around the resonance peaks appear. This stage can be detected using demodulated high frequency envelope spectra.

III. Bearing defect frequencies and their harmonic multiples occur. As the wear increases, more and more harmonic multiples of bearing defect frequencies occur and increase also the number of sidebands that are both around these harmonics and around multiples of bearing defect frequencies themselves. This stage can be determined from the vibration velocity spectra.

IV. Even the amplitude of the rotational component 1X is influenced at the end of the lifetime. It increases together with number of its harmonics. Discrete bearing defect frequencies and natural frequencies of the bearing components begin to disappear from the spectrum due to increased bearing clearances and are replaced by random broadband high-frequency "threshold noise".

Fig. 4.5.3 - Stages of Rolling Bearing Defect Development

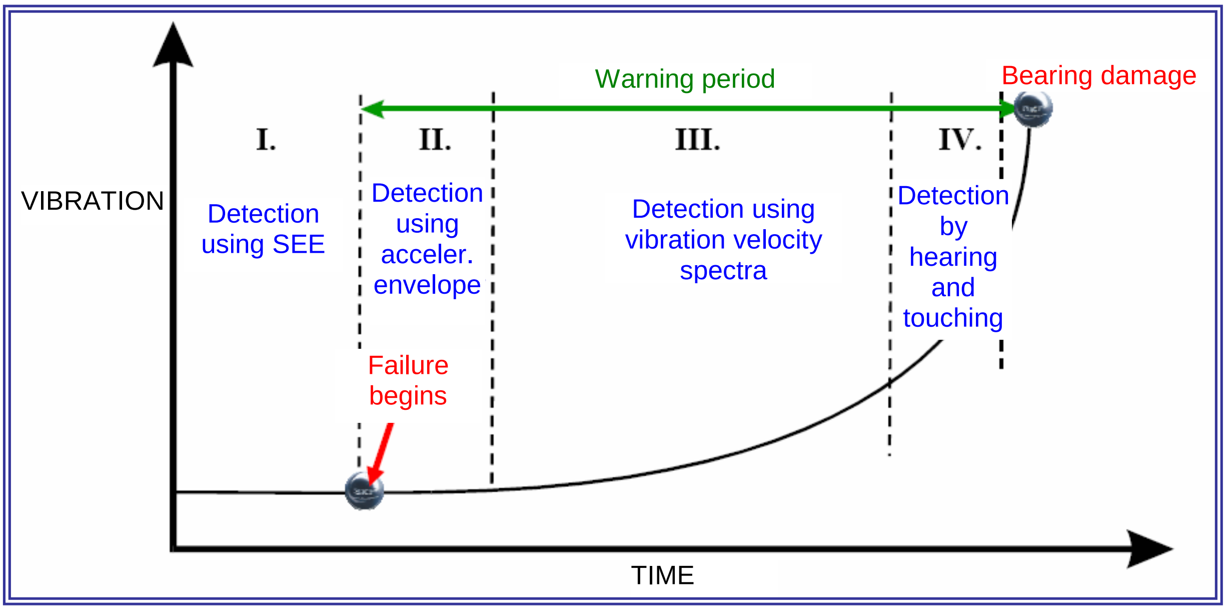


Fig. 4.5.5 - The Course of Rolling Bearing Defect Development

4.5.5 Acceleration Envelope

For a good data analysis, especially in the second stage of the rolling bearing defect development, a method called acceleration envelope is widely used nowadays. The principle of the method is explained in Fig 4.5.6.

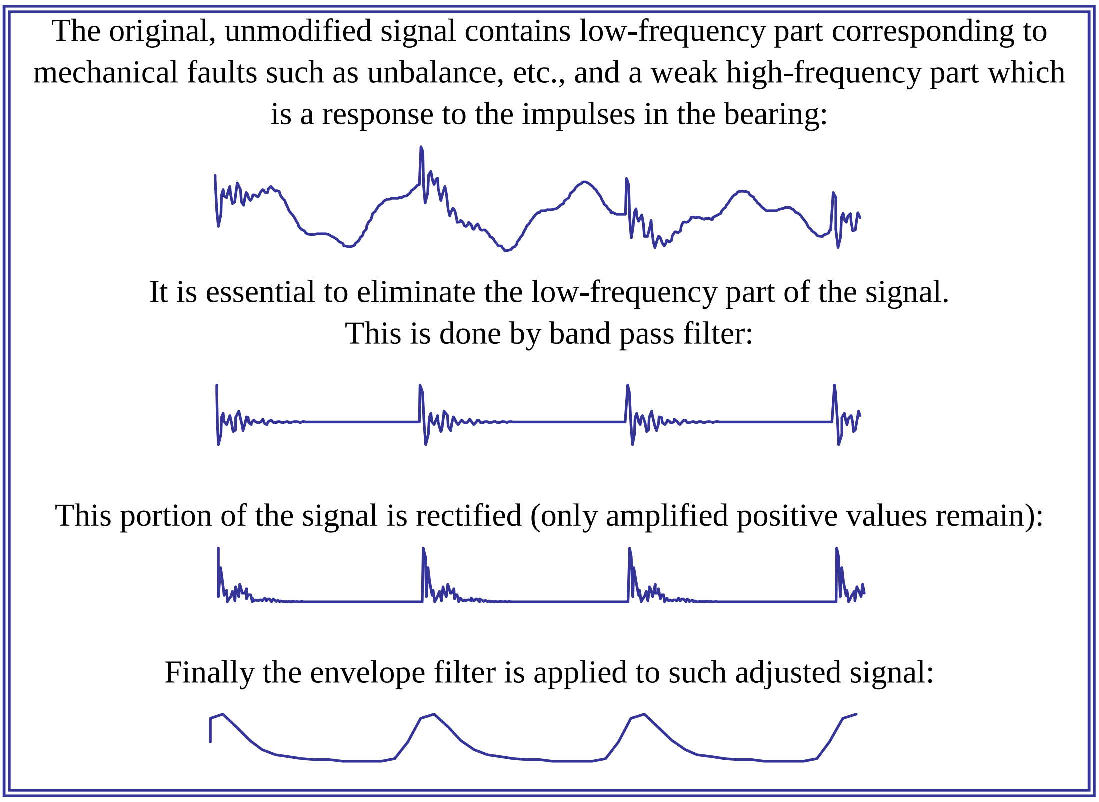


Fig. 4.5.6 - Creation of Acceleration Envelope

This signal is further processed, usually by two ways:

- The overall value is determined. Because of acceleration, dimension [g] is used, but because it is acceleration of an envelope signal, the letter E is added to the dimension, so the dimension of this quantity is **gE**.
- FFT of the signal is performed, yielding the acceleration envelope spectrum.

Practical notes:

To simplify the measurement task, analyzers are equipped with several band-pass filters to remove low-frequency portion of the signal. The principle for selecting the filter is based on the assumption that portion corresponding to other mechanical defects should be excluded from the signal. A rule of thumb is that the lower limit of the filter should be 10 times higher than the rotational speed (1X).

Example of filter setting:

5 Hz – 100 Hz	applicable for very slow-speed machines
50 Hz – 1000 Hz	applicable for slow-speed machines
500 Hz – 10 000 Hz	applicable for common machines
5 kHz – 40 kHz	applicable for gearboxes

4.5.5.1 Assessment of the gE Overall Value

As with all methods, it is recommended that a diagnostician would build a set of alarm values upon his experience. As an initial guide there is a recommendation of SKF company, compiled on the basis of extensive experiments. It is expressed both as formulas and graphically (see Fig 4.5.7):

Danger:

$$L = \left(\frac{f_{\max}}{1000}\right)^{0,43} \times 3,26 \times 10^{-4} \times n \times d^{0,55} \,[\mathrm{gE}]$$

Alarm:

$$L = \left(\frac{f_{\max}}{1000}\right)^{0,43} \times 1,09 \times 10^{-4} \times n \times d^{0,55} \,[\mathrm{gE}]$$

where:

L ... alarm limit for acceleration envelope measurements
f_{max} ... maximum frequency [Hz] for spectral band amplitude calculation
n ... rotational speed [rpm]
d ... diameter of bearing bore (indicator of load)
exponents ... empirical coefficients that should be set statistically exploring the existing databases

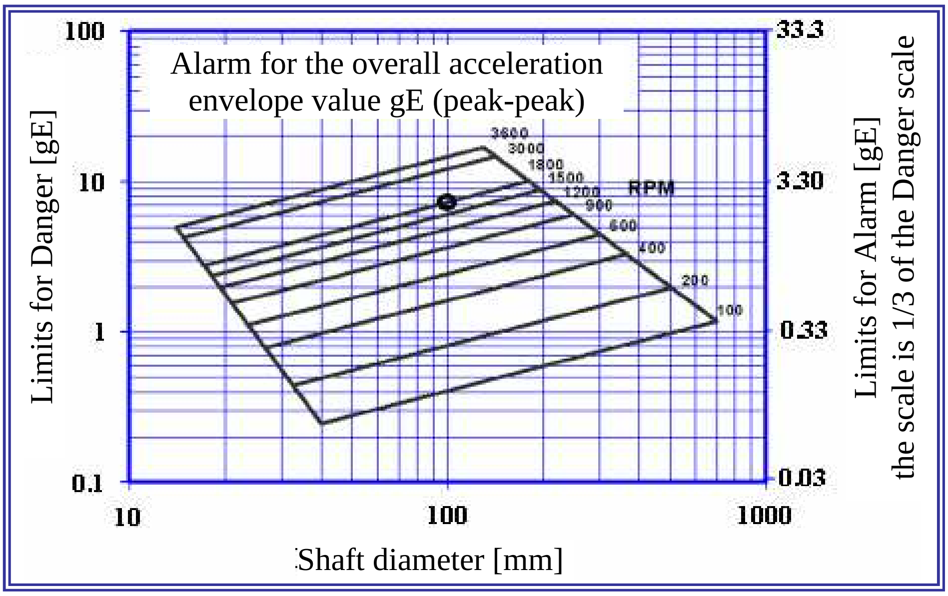


Fig. 4.5.7 - Setting of "Alarm" and "Danger" Limits on the Base of gE

4.5.6 Acceleration Envelope Spectra

As with ordinary spectra, when using the envelope spectra, it is also recommended to monitor *trends* of the individual bearing defect frequencies. Sometimes a bearing defect frequency occurs due to a larger load, but the fault does not develop. Fig 4.5.8 shows the envelope spectra for various extent of bearing damage; the damaged bearing is shown in Fig 4.5.9.

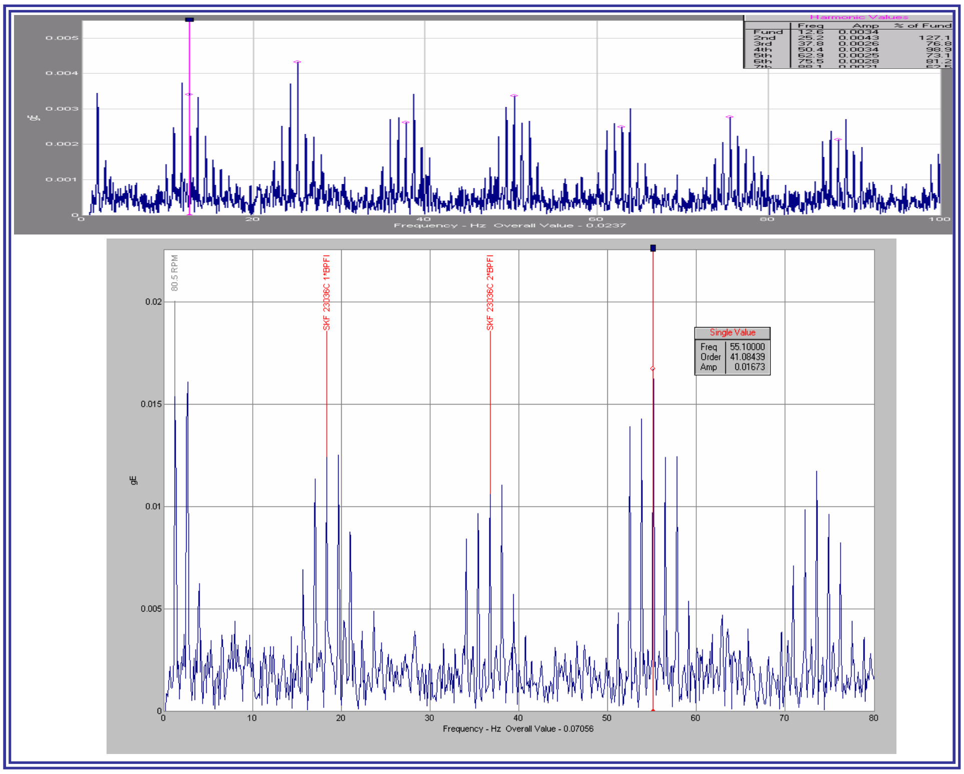


Fig. 4.5.8 - Envelope Spectrum of Medium (above) and Severe (below) Bearing Damage

Fig. 4.5.9 - Damaged Bearing

REFERENCES

[1] ISO 17359: *Condition monitoring and diagnostics of machines - General guidelines*

[2] ISO 13373-1: *Condition monitoring and diagnostics of machines - Vibration condition monitoring - Part 1: General procedures*

[3] ISO 13373-2: *Condition monitoring and diagnostics of machines - Vibration condition monitoring - Part 2: processing, presentation and analysis of vibration data*

[4] ISO 5348: *Mechanical vibration and shock - Mechanical mounting of accelerometers*

[5] ISO 10816-1: *Mechanical vibration - Evaluation of machine vibration by measurements on non-rotating parts - Part 1: General guidelines* and Appendix ISO 10816-1/Amd.1

[6] ISO 1925: *Mechanical vibration - Balancing - Vocabulary*

[7] ISO 1940-1 Review 2004: *Mechanical vibration - Balance quality requirements for rotors in a constant (rigid) state - Part 1: Specification and verification of balance tolerances*

[8] ISO 1940-2: *Mechanical vibration - Balance quality requirements for rotors in a constant (rigid) state - Part 2: Balance errors*

[9] ISO 11342: 1998 *Mechanical vibration - Methods and criteria for the mechanical balancing of flexible rotors*

[10] ISO 10814: 1996 *Mechanical vibration - Susceptibility and sensitivity of machines to unbalance*

[11] ISO 7919-1: *Mechanical vibration of non-reciprocating machines - Measurements on rotating shafts and evaluation criteria - Part 1: General guidelines*

[12] ISO 7919-3: *Mechanical vibration of non-reciprocating machines - Measurements on rotating shafts and evaluation criteria - Part 3: Coupled industrial machines*

[13] Randall, R.B. *Frequency Analysis*, Denmark: Brüel&Kjær, 1987

[14] Fryml, B., Borůvka, V. *Vyvažování rotačních strojů v technické praxi*. Praha: SNTL, 1962.

[15] Juliš, K., Borůvka,V., Fryml, B.. *Základy dynamického vyvažování*. Praha: SNTL, 1979.

[16] internal materials of SKF company

[17] Dow, S. *Understanding the Basic Theory Behind Vibration Analysis* [online], [quoted 2012-01-20]. Available from www<http://www.vibrationschool.com/mans/Indexa.htm>

[18] Lyons, J. *Dynamic Balancing* [online], [quoted 2012-01-20]. Available on www <http://www.irdbalancing.com/>

[19] Berry, J.E. *Illustrated Vibration Diagnostics Chart.* Technical Associates of Charlotte, P.C., 2007.

[20] Gasch, R., Pfützner, H. *Dynamika rotorů*. Praha: SNTL, 1980

[21] Bently Nevada Applications Note. *"Glitch" - Definitions, Sources and Methods of Correcting*. Minden, Nevada: Bently Nevada Corporation, 1993. Available also on www <http://www.ge-mcs.com/de/online-learning-center/glitch-definition-sources-and-methods-of-correcting.html>

[22] Giberson, M. *Dr. Mel's Technical Notes: Babbitt* [online], [quoted 2012-03-02]. Available on www <http://turboresearch.wordpress.com/page/2/>

[23] Trivanovic, D., Wier, W. *Using of Acceleration Envelope for Journal Bearings*, Aplication Note CM 3093 of SKF company

[24] Muszyńska, A. *Rotordynamics*. Boca Raton, FL: Taylor & Francis Group, 2005

[25] IRD Technical Paper No. 116. *A Practical Guide to In-place Balancing*.

[26] Wei, J. *Recommended Initial Criteria for Evaluation of Bearing Condition when Using Acceleration Envelope Measurements*. Aplication Note CM 3068 of SKF company.