Category

III

MOBIUS INSTITUTE

Vibration Training Course Book Category III

Produced by Sumico Technologies

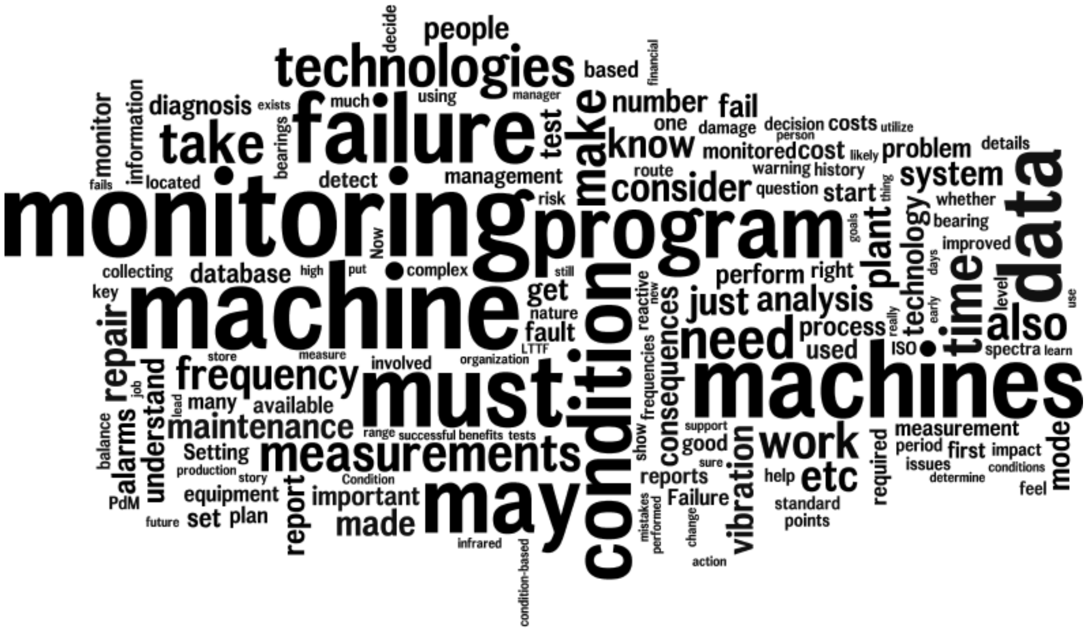
This manual is designed as a guide only.

In practical situations, there are many variables, so please use this information with care.

Rev. 12-05-12

Produced by Sumico Technologies

Table of Contents



Chapter 1

Condition Monitoring

Objectives:

- Getting a program started
    - Selecting machines to test
    - Selecting technologies to use
    - Setting up the database
- Reporting
- Running a successful program

Designing the program

The standard ***"17329:2003 Condition monitoring and diagnostics of machines -- General guidelines"*** is the standard that provides an overview of the condition monitoring process.

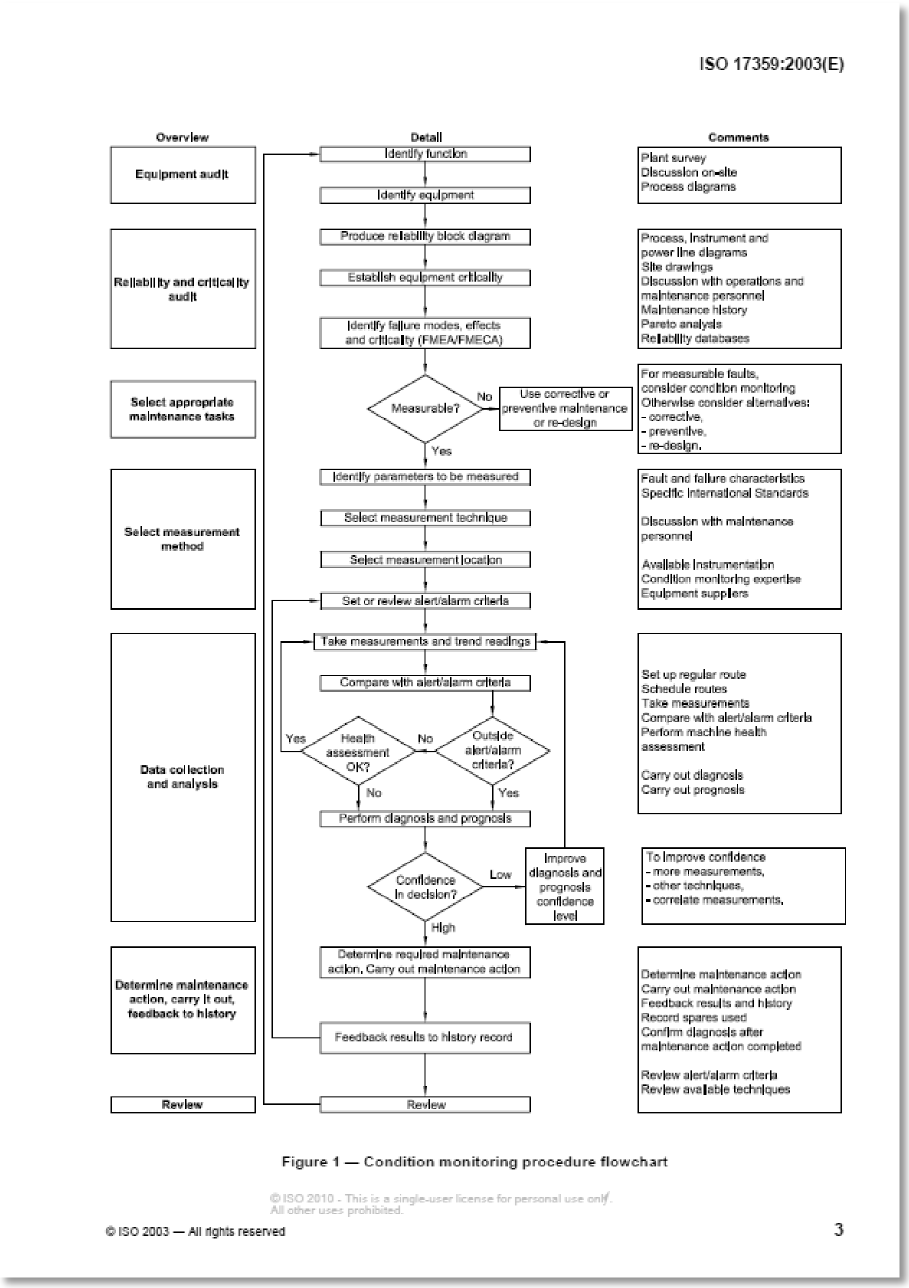


Figure 1-1

ISO 17359:2003 includes a flow chart of the "condition monitoring procedure". It is summarized in Figure 1-2. Per the ISO standard, in this chapter we are focusing on the four blocks highlighted in Figure 1-2.

Figure 1-2 - ISO 17359: 2003 condition monitoring procedure flowchart

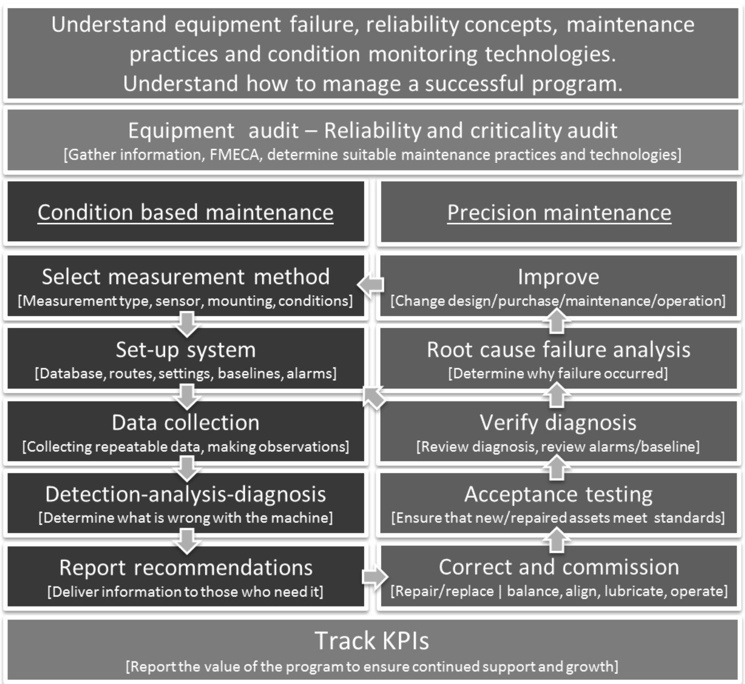


Figure 1-3

Review the plant

The first challenge is to determine which machines should be tested as part of your program, and if you are using multiple technologies, deciding which technologies should be used on each machine.

It comes back to the question; "*do the consequences of failure outweigh the cost of monitoring the machine?*". *Assuming that there are not regulatory or significant safety requirements*, you need to know the consequences of failure, and put a monetary value on the failure and balance that against the cost of monitoring the machine.

To make this assessment, you need to perform a Failure Modes, Effects, and Criticality Analysis (FMECA).

If you don't know how a machine fails, and the consequences of it failing, then it is difficult to decide which machines should be monitored; including which technologies should be utilized and what frequency of monitoring (permanently, hourly, monthly, etc.) is required.

We have developed a chart that will take you through the decision making process (Figure 1-4). For each failure mode, we can consider the consequences of failure, the likelihood of failure, the frequency of failure, the technologies available to detect the failure, and the lead time to failure (via the P-F interval).

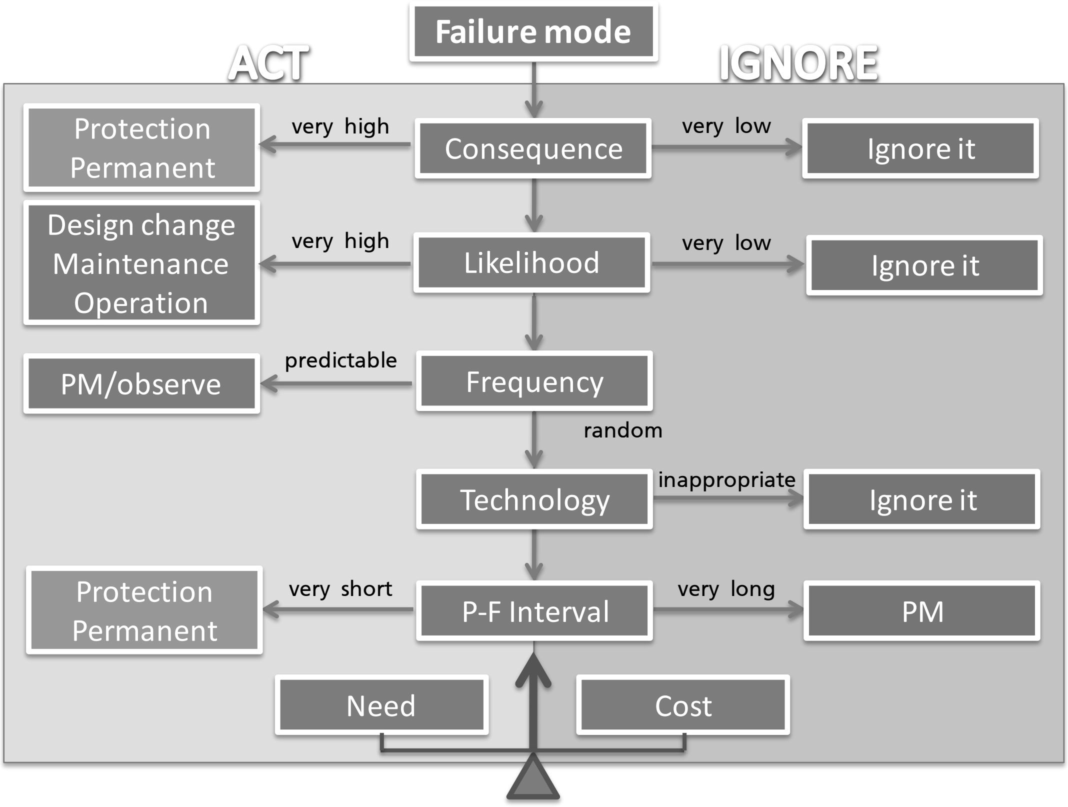


Figure 1-4

Ultimately we are trying to balance the need to monitor the failure mode (based on the consequences and risk) against the costs of implementing the technology. Remember, we considering whether, *on the one hand*, we should begin using a particular technology (e.g. infrared thermography), and *on the other hand*, whether it is worth the time and expense to monitor a specific machine with that technology. In the case of vibration monitoring, we don't just monitor a machine because it vibrates! Those costs have to be weighed against all the issues: consequences, probability of failure, nature of failure, and the method we have available to monitor it.

You have to consider the history of the machine. Has it been reliable? What has been (or would be) the impact of failure: downtime, secondary damage, and parts/labor requirements. What are the production downtime and capital costs involved? Is there a spare unit?

The following guide (Figure 1-5) may help you to assess the consequences of failure. If the analysis on a given machine results in *"Prevent it breaking"*, then you can continue with the process. Otherwise, we can ignore the machine/failure mode.

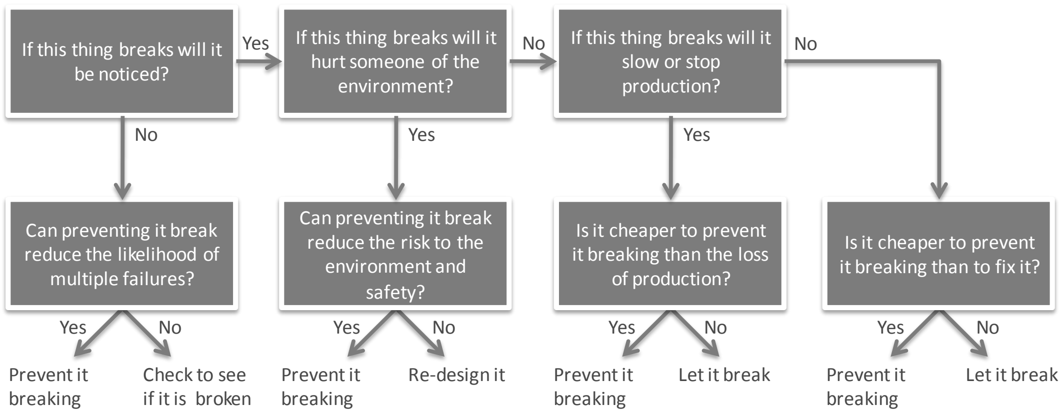


Figure 1-5

Risk analysis

You really need to perform a risk analysis. That analysis will help you to decide whether it will be monitored, and also how often and with which technologies. It will also help you later on when a fault is detected: at what point should the repair be made?

The risk may change during the year. At times of high demand, or adverse weather conditions (summer for cooling plant, winter for power generation, for example), the monitoring frequency and repair plan may need to change.

You must then consider the technologies required to detect future problems. Can you be successful with vibration? Do you need to include other technologies: infrared, wear particle analysis, etc?

Technology

Application	Vib	Lube	Wear	MCA	IR	US	Vis
Generator	✓	✓	✓	✗	✓	✓	✓
Turbine	✓	✓	✓	✗	✓	✓	✓
Pump	✓	✓	✓	✓	✓	✓	✓
Electric motor	✓	✓	✓	✓	✓	✓	✓
Diesel engine	✓	✓	✓	✗	✓	✓	✓
Fan	✓	✓	✓	✓	✓	✓	✓
Gearbox	✓	✓	✓	✗	✓	✓	✓
Cranes	✓	✓	✓	✓	✓	✓	✓
Electric circuit	✗	✗	✗	✓	✓	✓	✓
Transformer	✗	✓	✗	✓	✓	✓	✓

Figure 1-6 - From Keith Young, paper in Maintenance Technology, June 1995

Consider the nature of the symptom: wear occurs, heat is generated, impacts occur, corrosion exists, fatigue occurs, etc.

Then consider whether each technology is capable of detecting that symptom.

	Vib	Lube	Wear	MCA	IR	US	Vis
Wear	✓	✗	✓	✗	✗	✓	~
Heating	✓	✓	✓	✗	✓	✗	~
Impact	✓	✗	✓	✗	✗	✓	~
Corrosion	✗	✓	✓	✗	✗	✗	~
Fatigue	✓	✓	✓	✗	✗	✗	~

Figure 1-7

Physical issues

With any monitoring technology, you must consider safety and practical issues. If you have remote plant, or equipment located in a hazardous environment, or bearings located in inaccessible locations, you may need permanent sensors or on-line monitoring.

Permanent sensors can be installed within a machine with their connection points located in a convenient location.

Criticality and failure mode

The key is to take measurements frequently enough so that you don't miss the start of the warning period. You may be able to do all the measurements yourself or utilize operators and simpler recording devices.

Test frequency

The key is to take measurements frequently enough so that you don't miss the start of the warning period. You may be able to do all the measurements yourself or utilize operators and simpler recording devices.

- Too many measurements may cost too much.
- Too few measurements may result in a lead time that is too short (or the machine could even fail).
- It is important to consider the failure modes: use the appropriate technology and measurement frequency.

How many machines to test?

If you are starting a program it can be daunting to take-on the most complex machines in your plant. And you must be careful not to have too many machines in your program initially; you can easily become overwhelmed.

This topic can be a little controversial because your natural instinct may be to do the exact opposite.

If a machine is very complex there is a higher probability that you will make an error when setting it up in your database and analyzer, and when performing the analysis and diagnosis.

And if you test too many machines you are more likely to make mistakes and inundate yourself with data.

We feel that it is a good idea to start small, have some success, learn from your mistakes (small ones), and then grow the program. Just take a few months to get things right, build your confidence, then move on and do a better job with a larger number of machines, and with machines that are more complex.

Setting goals

And you must be sure that management knows which machines you are monitoring. Once you have spent all that money for a system, they may think that every machine is instantly immune from failure. Everyone will be in for a nasty surprise when a machine fails.

You must set realistic goals for yourself, and make sure that management agrees with those goals.

If you do not have management buy-in, you will not be successful. Happy manager = happy employee.

Know your machine

Now that you have made your selection, you need to compile the list of key details on each of the machines. At a minimum you should know the:

- The shaft speeds
- The motor size (hp or kW) and number of rotor bars
- The number of teeth on each gear
- The diameter of each sheave and the distance between their centers
- The number of vanes in pumps
- The number of blades on fans

It is also very helpful to know the details of the bearing - either the physical details (number of balls, etc.) or the forcing frequencies. Remember, however, that it is often "obvious" when a bearing begins to fail (as described in the diagnostic section), so knowing the details is not required to be successful.

Some of this information can actually be derived from your first set of spectral measurements. If you are monitoring a pump and there is a high 6X peak, then there is an excellent chance that the pump has six vanes.

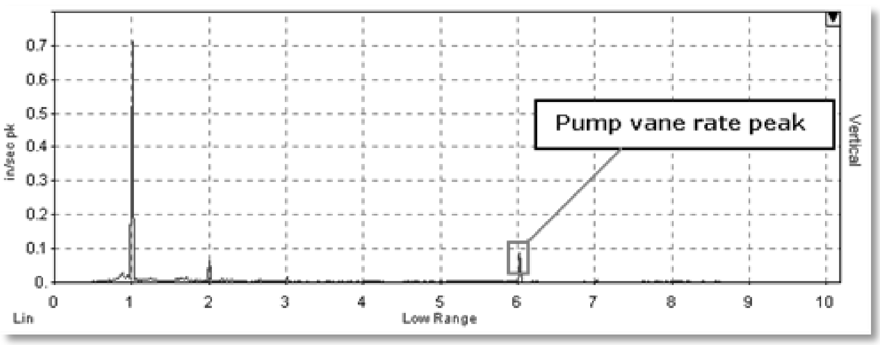


Figure 1-8

Standard test conditions

You must also establish standard, repeatable test conditions. It is very important that the machine is running at the same speed (within just a few percent) and the same load because when comparing vibration measurements the only thing we want to change between tests is the condition of the machine.

What do you measure?

Now you must decide what data to collect. There are really two parts to this question. First you have to decide the types of measurements (overall level measurement, spectrum, etc.) and secondly, where required, the specifications of the measurement (frequency range, number of averages, etc.).

Every piece of data collected will take time to measure, store, analyze and document. However in the early days of your program it may be wise to err on the side of too much data, rather than too little. If it proves that some data is not useful, stop collecting it.

The obvious options to consider are overall level measurements (10-1000 Hz typically), time waveforms, spectra, bearing measurements (spike energy, shock pulse, HFD, etc.) and demodulated spectra.

Overall level measurements can be useful for trending purposes, and for comparison against standards charts. They may help you to get started by acting as a basic warning system. But in the longer term they may prove to provide little information.

Vibration spectra should be collected, however the frequency range and resolution must be chosen carefully. It is recommended that you first compute the forcing frequencies (as described earlier), and then take a test measurement with a frequency range at least 30% higher than the highest frequency of interest. An examination of this measurement may reveal important frequencies that should be monitored routinely.

Setting up the database

Each condition monitoring system is different, however the better systems now employ database setup "wizards" that walk you through the steps required to not only set up the database, but to establish the forcing frequencies and bands so that exception reports and expert systems can function correctly.

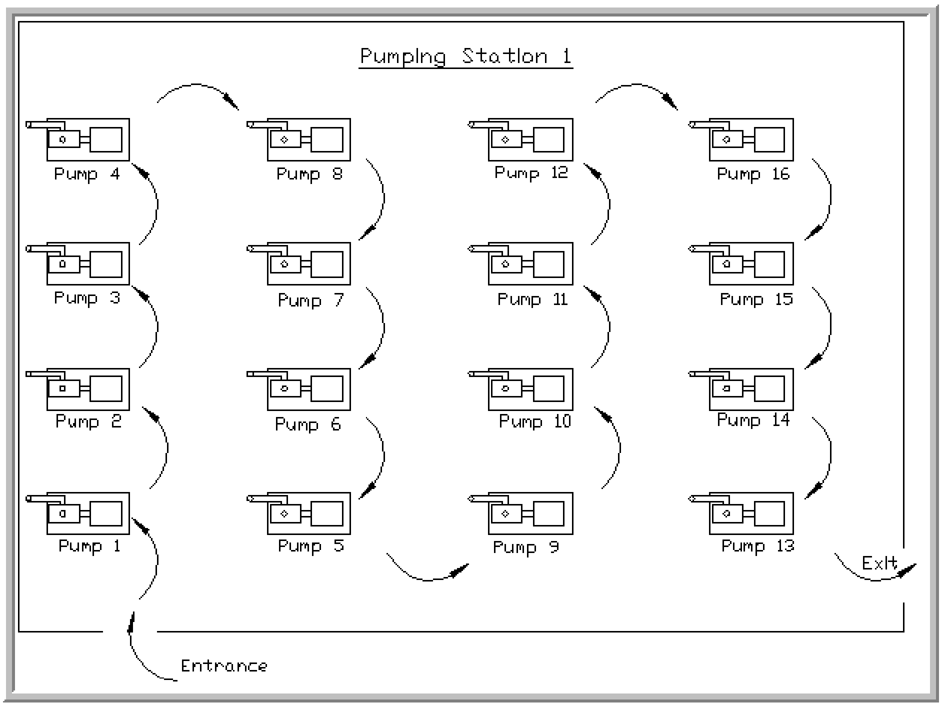
Setting up a database varies from system to system, and is beyond this course to discuss. But there are four VERY important issues:

- Naming machines and points
- Organizing the machines and points
- Creating routes
- Generating reports

Setting up the route

The machines must also be placed on to an ordered list of machines, typically called a "route" or a "survey". The route is used to guide the person collecting data around the plant.

The route will guide you through the points you must test, and they can be ordered as well. Time can be saved and collection errors will be reduced if you consider the path you will take and the most convenient order for the tests.



Pumping Station 1 is an 'area' within a Plant 'site'

Figure 1-9

Plant database structure

The vibration monitoring software is used to store the data and provide a means to perform analysis. It will enable you to name every *machine* that you test and organize the machines according to the *area* where it is located, and the *plant* where it is found.

Take time to organize the plants, areas and machines carefully.

- Reporting and data access can be greatly improved
- You can get more out of the system – don't just treat it as a place to store data

Reporting

All vibration monitoring systems support reporting. These reports may be your link to management.

Make sure you know how to get the most out of the reporting capability

Figure 1-10

Setting and maturing alarms

In the Analysis Module we discussed the methods available for setting alarms levels. The task is to set sensible alarms to get the program started, and then refine them as your program matures.

The basic principle is to start with alarms using the best available information (vendor recommendations, ISO standards, etc.), but then, if your system supports it, employ statistics to compute new alarms levels when you have five or more sets of data.

The statistics can even "learn" from other identical machines, so if you have a series of four identical pumps, compare all the data, weed out any data that indicates a problem with the machine (if it exists), and then build a set of alarms from the remaining data.

This approach ensures that the alarms are based on whatever is "normal" for that machine - based on its size, make, location, function, mounting, etc. When the abnormal occurs, the exception report (or expert system report, as the case may be) will indicate that a problem exists.

Making the recommendation

Now, with a complete picture, you should be ready to make your recommendation. You must understand the way in which your maintenance organization works so that you know how to report machine faults.

The actual action taken will depend on the nature and severity of the fault, but also the demands on the equipment, the availability of spares, and more. It is an economic decision.

It can be frustrating that all of your work to detect and diagnose a problem may appear to be ignored. However you must realize that someone else is likely to be responsible for the continued operation of the plant and for the maintenance budget. Many factors have to be considered before an unscheduled repair is performed.

Ideally you will give plenty of warning about a problem so that the repair can be performed during the next planned shutdown period. That is the ultimate goal.

Also understand that not all of your diagnoses will be correct. Even the most experienced practitioners get it wrong. It is one thing to detect the nature of the fault, but getting the severity right can be a challenge. If a repair is performed and the bearings show only slight damage, or no visible damage for example, then you might cop some flack!

But just take it on the chin, explain your logic, and learn from the experience. When you do make important calls that prove to be right on the money, you will gain their respect. But again, if you have involved them in the decision making process they will be less defensive and less critical if the diagnosis is incorrect.

You may not have support at first

You must also understand that, particularly in the early days, other maintenance, production and engineering staff may be skeptical about your diagnosis. You must be patient, but you must also be proactive.

You should try to educate these people so that they understand how you arrived at your conclusion, and you should try to involve them in the diagnosis process.

If you have consulted them with questions about past failures, repair history, the history of similar machines, the current performance of the machine, any strange noises heard, and so on, then they will feel part of the process. Make them part of the team.

The last thing you should do is give the impression that only you could possibly understand your technology. People who have been working with the equipment for a long time will feel that they know a lot more than you do with your new data collector. Respect those feelings and ask their advice - you will get a lot further.

In addition, you should ensure that you explain your diagnosis, not with spectra and other complex data, but with a description that supports your conclusions. If they understand your reasoning, and understand how confident you are in your conclusions, then they will be more accepting.

But do consider educating them in condition monitoring techniques. They do not need to know as much as you do, but if they simply understood that the vibration pattern reflects the forces within the machine then they will feel more comfortable with your diagnoses.

Don't just look for faults

Don't stop at the diagnosis. You must go into the next phase: root cause analysis. Do what you can to determine why the fault condition existed in the first place. Was the equipment under rated, was the last repair faulty, was the balance/alignment precise, is lubrication adequate, etc.

If you can determine why the fault occurred, and then work towards correcting the situation, the machine will be more reliable in the future.

Do not forget that when the repair is made to the machine, regardless of the accuracy of your call, you must perform additional tests to ensure that the problem has been corrected, and that the machine is now running in good condition. Now we are in the final phase: verification.

Check the balance and alignment. Check that the bearings are OK. Check that the hold-down bolts are tight, and so on. Your data collector and newly gained skills are vital at this stage.

Report your successes

You have the opportunity to save your organization a great deal of money. Your efforts can result in significant cost avoidance. Your work can improve the profitability of the organization through improved up-time and improved product quality.

For your future viability, do everything you can to report your impact to the managers that make the key financial decisions. The work involved may be quite different to what you are accustomed to, but if you want to keep your job and even grow your group, this is a necessary step to take.

First you must record and report on the "saves" you make. But don't just document that you "found that the motor bearing would fail and it was replaced before it did fail". You must put it in economic terms.

You will have to research the likely secondary damage, labor costs, and impact on production and safety. Don't exaggerate the numbers, but don't be too conservative either.

Over a period of time, as the equipment reliability improves, you will not have as many "saves" to document. So you have to take your reporting to the next level.

You must research the annual maintenance, up-time, spares inventory, overtime labor costs, and other metrics, and then show how these have changed since the time that your program has been put into place.

It is difficult to report on these figures and take all of the credit, but you must show the cause and effect. You must resell the benefits of condition monitoring and show the impact that has been made. Be aggressive, make sure that the right people know.

Why programs fail

Failure can be measured in two broad ways:

1. Failure to deliver the benefits that are really possible.
2. Failure to continue to operate the program: loss of staff or cancellation of the program and loss (or reassignment) of everyone involved.

This graphic (Figure 1-11) illustrates some of the reasons why a program fails.

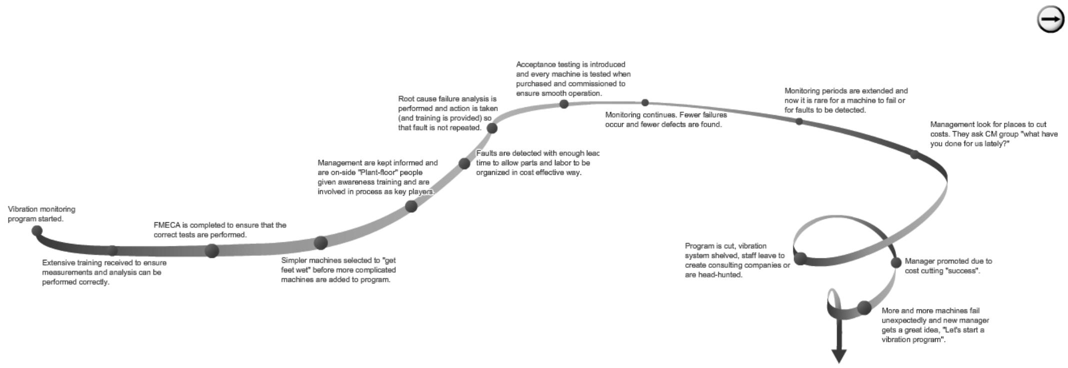


Figure 1-11 - The growth and failure of the program

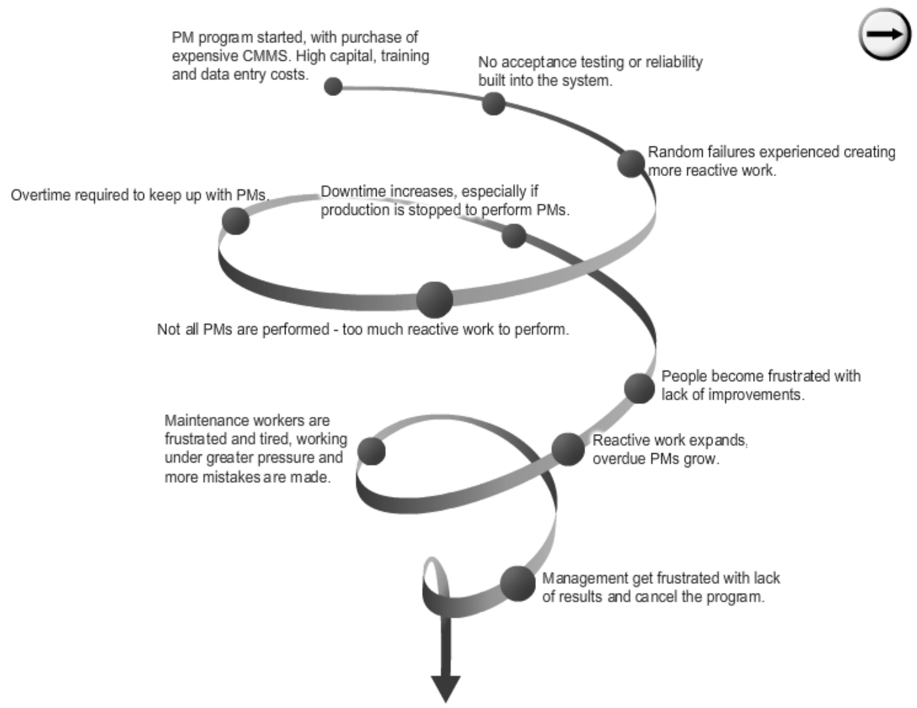


Figure 1-12 - The condition monitoring program death spiral

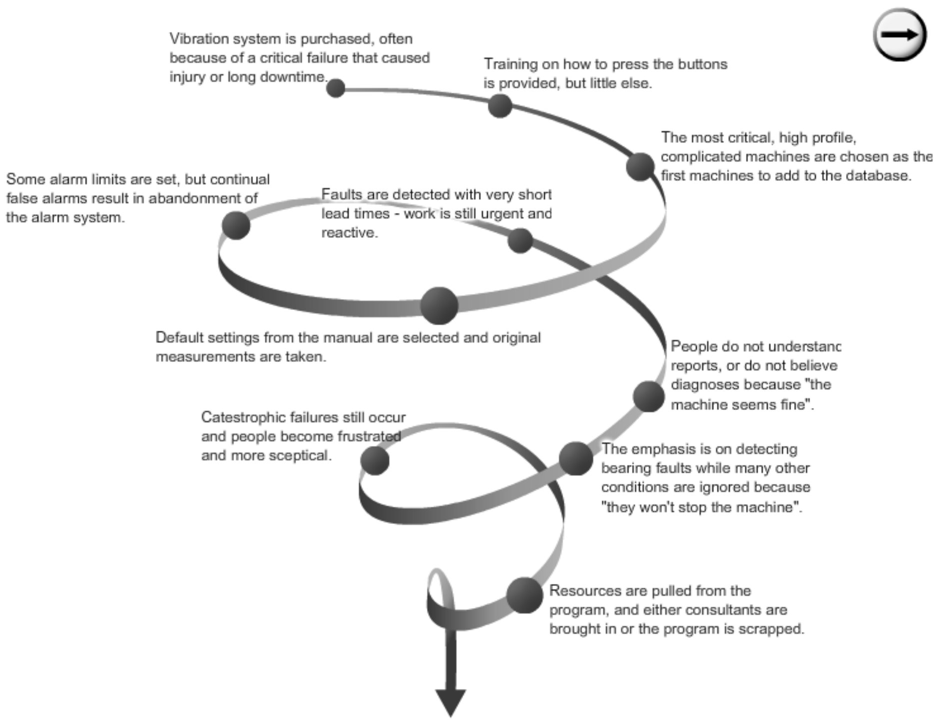


Figure 1-13 - The vibration program death spiral

Financial benefit is not understood or reported

A common reason for program failure is the failure to track and report the benefits.

All it takes is a new manager, or tough financial times, for a decision to be made that savings can be made by cancelling the PdM program.

This happens with alarming frequency!

No belief in upper management or shop floor workers

It does not matter which technologies you utilize or which products you own, if the rest of the plant does not believe in the idea of condition-based maintenance then the program will never succeed. If people do not conduct work based on condition and still perform reactive work, or they wait until the bearings are screaming and are so hot that they would burn your hand, you are not performing CBM, you are still in reactive mode.

No strategy

It is common for people to purchase one or more condition monitoring products and simply "start using them". Not all machines *that should be monitored* are monitored. Multiple technologies are not used *where they should*.

No commitment

It is also common for people to purchase just one or two condition monitoring technologies, and then assign a person to start collecting data on some machines. That person may have other responsibilities which always seem more urgent than collecting data.

Wrong people

If the PdM technicians do not have the right attitude (or intelligence), or the PdM manager does not have the drive and persistence to make the program work (or ability to lead), then again, the program will fail.

It takes a certain set of skills, but it also takes a special kind of personality to be able to make this plan work.

Poor execution

Many of the condition monitoring technologies are easy to use incorrectly. Anyone can measure vibration, point an infrared camera at things that "look hot", listen to sounds with ultrasound that are misinterpreted, incorrectly take oil samples and then ignore the reports that come from the laboratory.

Without proper training, certification, and careful design, the condition monitoring technologies will not deliver the desired results.

Condition monitoring versus troubleshooting

As a follow-on issue, programs also fail because the condition monitoring tools are used in reactive mode; to troubleshoot problems rather than to continuously monitor the condition so that faults can be detected from an early stage.

Poor reporting

There is not much use having PdM technicians who know the current health of an asset if that information is not communicated to the people who need it. Two mistakes are commonly made:

1. Reports are not made available to those who need them (in a timely fashion)
2. The reports do not clearly state the status, the severity, and the recommended action. Providing data in support is OK, but managers/planners need actionable information.

Condition monitoring, not improvement

Just because you monitor machines does not mean their reliability is improved. It is important to stop machines from failing catastrophically, but the greatest financial rewards are gained when the life of the machine is extended through defect elimination.

Short Lead Time to Failure: LTTF

When you detect a fault you will have a certain amount of time before the risk of failure becomes so high that the maintenance department will be forced to take action – this is the LTTF.

Maintenance planners would like to have more time to plan the repair work – to find the most cost effective time to perform the work.

If the technologies are used correctly, the planners should have weeks to months to plan, not just days to plan.

Calendar-based not condition-based maintenance prevails

As a follow-up comment, the real problem with the short LTTF and a lack of belief, is that the plant stays in reactive mode, or at best, still performs most maintenance based on time instead of condition.

The goal must be to reduce failures, reduce time-based PMs, increase condition-based work, and improve the reliability of all assets (where it is shown to be viable to do so).

Conclusion

There is story after story of condition monitoring groups that have been shut down because they did too good a job. Managers question "what have you done for me lately". If you don't have a good answer, you're history!

So, get a good answer, and give it to them before they ask the question. You need to become a success story, something that managers want to be associated with, not just a cost on the balance sheet that no one quite understands.

Chapter 2

Signal Processing

Objectives:

- Describe the purpose of filters and what they do
- Describe how sampling affects spectral data
- Determine Lines of Resolution needed to resolve frequencies
- Explain the purpose for Windowing
- Describe three Averaging methods and how they work

This chapter discusses what really happens to the signal from the transducer; the processes related to making the measurement from the perspective of capturing and manipulating the signal to produce the data we need. It involves digitizing the signal and producing the spectrum. Issues related to that are covered such as filters, sample rates, resolution and windowing.

A Quick Overview

Inside the data collector/ analyzer the "analog" signal from the sensor is converted to provide usable data: waveforms, spectra, rms readings and more

There are many settings and we need to know what is happening inside the analyzer in order to better understand the information it is giving.

Step One: Power the sensor - The ICP accelerometer requires power sent to it. The signal from the sensor is an electrical signal – an analog signal.

Step Two: Digitize the signal – The analyzer is a digital instrument so the data must be converted into a digital signal so the analyzer will have numbers to work with.

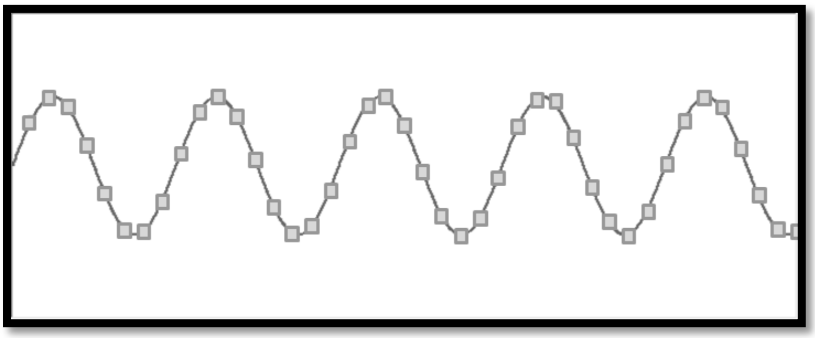


Figure 2-1 The analog signal must be digitized so the analyzer has numbers to work with.

There are a number of considerations:

- Does the voltage from the sensor match the input range of the analyzer?
- How quickly should the signal be sampled?
- How many samples are kept?

Step Three: Process the signal - generate other data that can be worked with: spectra, rms overall readings, envelop spectra

Considerations in the processing:

- Filtering
- Integration
- Windowing

- Averaging

Filters

When a transducer is mounted on a machine, the electrical output is a continuous analog signal representing the vibration at that location. It is important to understand what filters do and how they are used in the field of vibration.

There are four types of filters:

- Low pass filters: allow low frequencies to pass through
- Band pass filters: allow frequencies within a band to pass through
- Band stop filters: blocks frequencies within a band from passing through
- High pass filters: allows high frequencies to pass through.

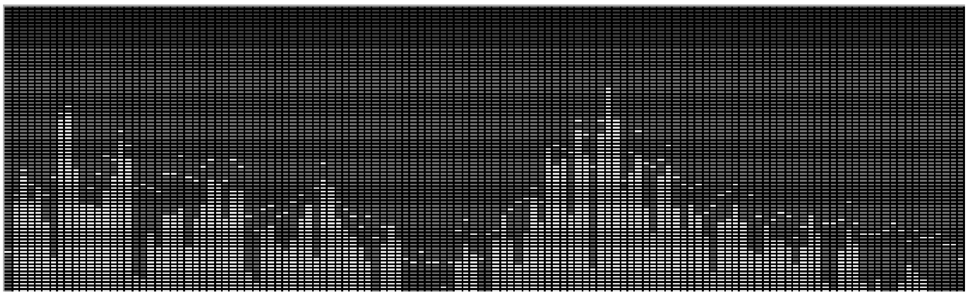


Figure 2-2 Data from a machine, no filter applied.

Figure 2-3 is the same data but with a Low Pass filter applied. It is allowing the low frequencies through while blocking the high frequencies.

Low-pass filters are by far the most common filter type, earning wide popularity in removing alias signals and for other aspects of data acquisition and signal conversion. For a low-pass filter, the pass-band extends from DC (0 Hz) to the frequency specified, and the stop-band lies above the specified frequency.

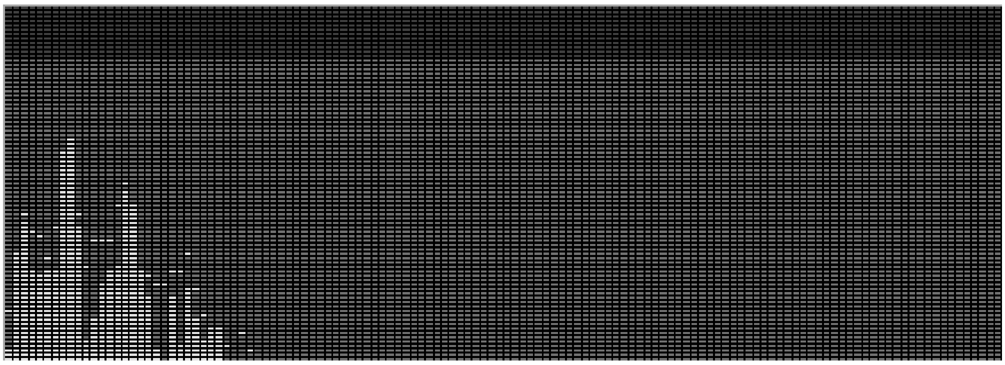


Figure 2-3 Same data but with a Low Pass filter applied.

Band-pass filters transmit only those signal components within a band around a center frequency. An ideal band-pass filter would feature brick-wall transitions at fL and fH , rejecting all signal frequencies outside that range. Band-pass filter applications include situations that require extracting a specific tone, such as a test tone, from adjacent tones or broadband noise. See Figure 2-4

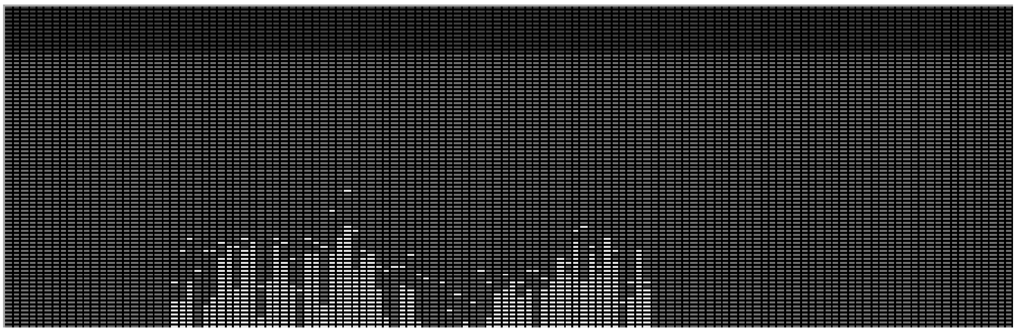


Figure 2-4 Band Pass filter applied. It effectively blocks the high and low frequencies.

Band-Stop (sometimes called band-reject or notch) filters transmit all signals except those between specified frequency ranges. These filters can remove a specific tone - such as a 50 or 60 Hz line frequency pickup - from other signals.

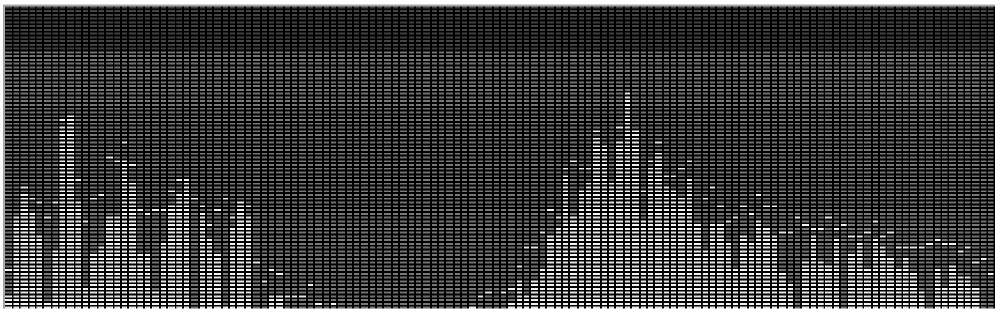


Figure 2-5 Band Stop filter blocks the frequencies in a particular frequency band.

In a **High-Pass filter**, the pass-band lies above a specified frequency, while the stop-band resides below that point.

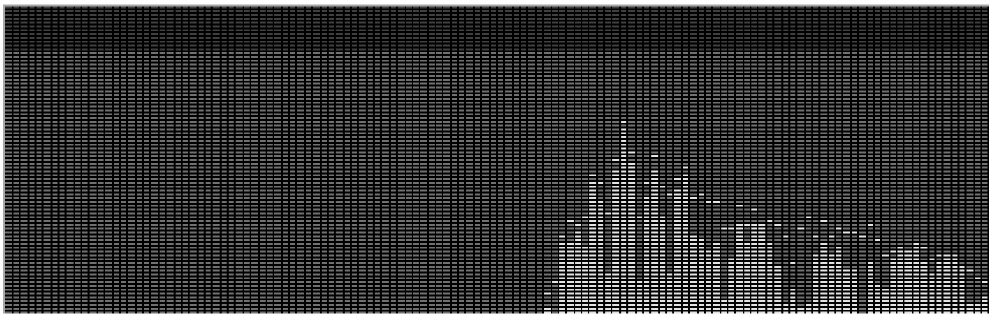


Figure 2-6 High Pass filter blocks all frequencies below the specified frequency.

High pass filters are normally used in early bearing wear detection. It blocks the high amplitude, lower frequencies to enable the scaling to adjust to the low amplitude, high frequency signals generated by bearing wear in the early stages of wear.

Tracking Filters

A tracking filter is a low pass filter or a band pass filter which automatically tracks the input signal versus the rpm.

They are used in order tracking to follow a machine speed precisely during changes. It is useful in run-ups and coast-downs to capture phase data along with the peak of vibration.

Some tracking filters can track multiples of 1x.

It is used on some strobes to accommodate an accelerometer to generate a once per rev signal.

Filter Characteristics

Two types of filters are common used in the field of vibration, Digital and Analog.

Digital filters are achieved with special "digital signal processing" (DSP) chips or software.

Analog filters are created with electronic components such as capacitors and resistors. Many analyzers still have some analog filters.

There are two issues to be concerned about:

- The filter characteristics (the cut-off)
- Settling times

Ideally, the filters would block the unwanted frequencies like putting up a brick wall. See Figure 2-7. This would provide a clean break and keep out all unwanted signals. However, this is not the case.

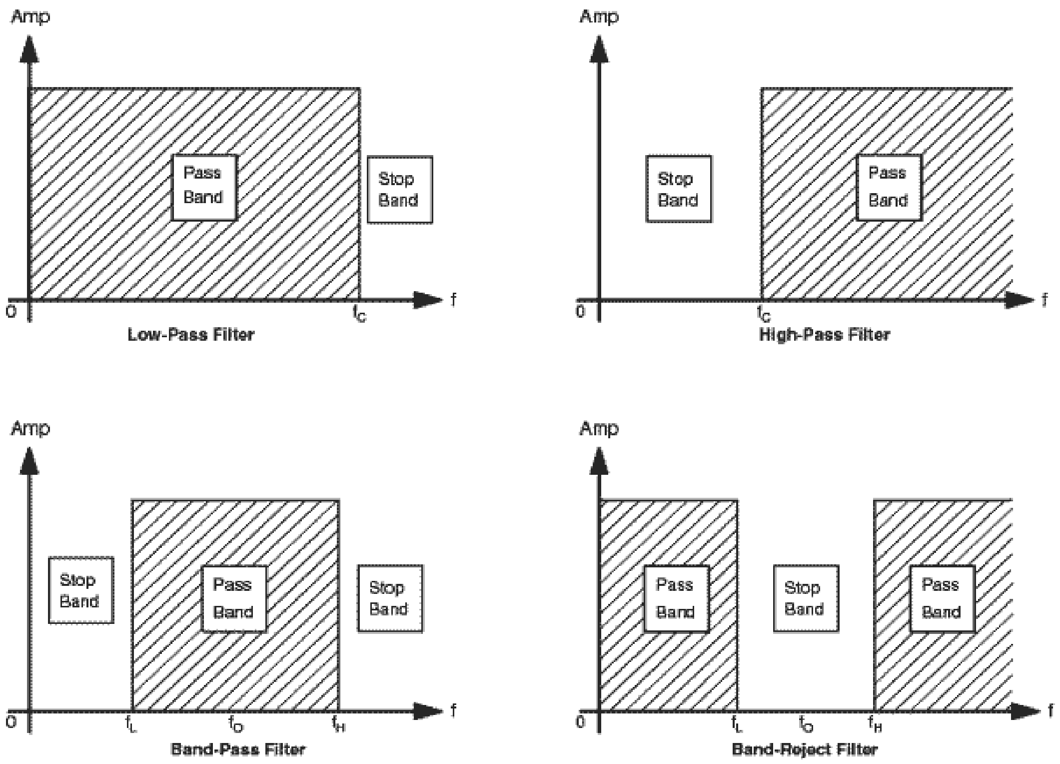


Figure 2-7 Ideal Filter Design would block out all unwanted frequencies.

In reality, especially with analog filters, there is a region where some frequencies will be attenuated, but not blocked.

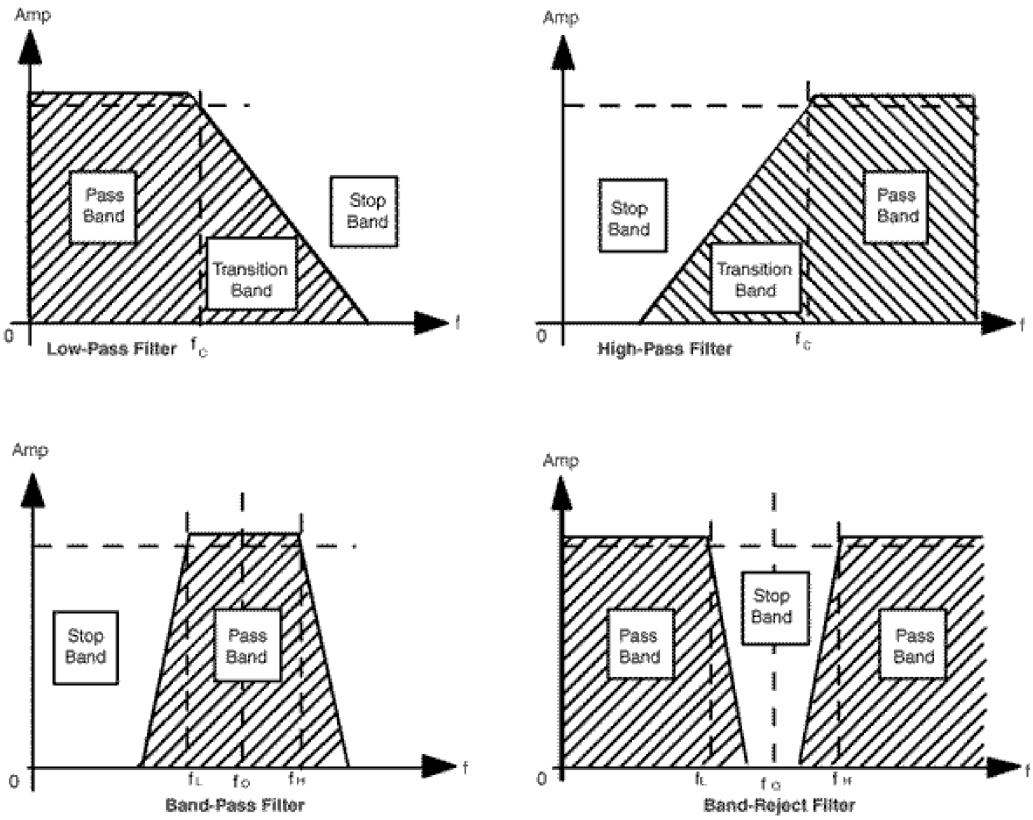


Figure 2-8 Actual Filter design has a transition band that lets some of the frequencies through.

There is one concept that is important to understand regarding filter electronics. When data is applied to the filter circuitry, it causes the output of the circuit to "ring". This requires the settling time to be set for a duration long enough to allow the circuit to settle to normal levels before capturing the data.

Integration, differentiation and signal to noise ratio

The analog signal from the accelerometer is proportional to acceleration. We typically view the data in acceleration or velocity.

- Converting the data from acceleration to velocity is called "integration". The reverse of the process is called "differentiation".

- To convert from acceleration to displacement the process is called "double-integration". To convert from displacement to acceleration is "double-differentiation"

This process can be achieved with analog circuitry (a filter) or it can be performed digitally. Some data collectors offer both methods, while others provide one option or the other.

Integrating reduces high frequency vibration and amplifies low frequency vibration (and introduces a 90° phase change), as shown in this example.

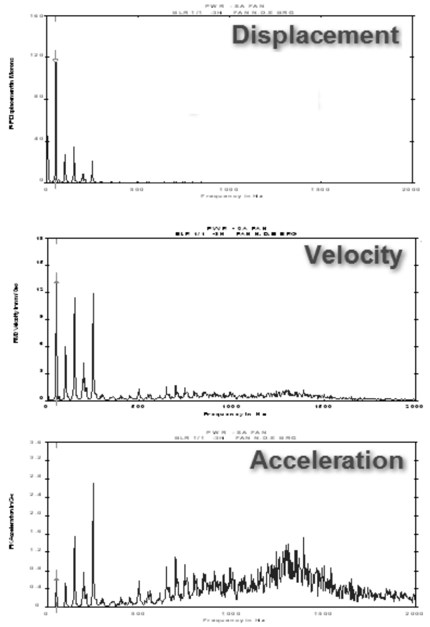


Figure 2-9

The relationship between acceleration and velocity is 1/f (where 'f' is frequency). [From acceleration to displacement the conversion is $1/f^2$]

- If 'f' is very small, velocity is very large.
- If 'f' is zero, velocity is infinite!

Due to this fact, and the characteristics of filters:

- Low frequencies can be amplified by a factor of 10 to 100 times
- High frequencies can be attenuated

Therefore a high-pass filter is used to remove/attenuate the low frequencies.

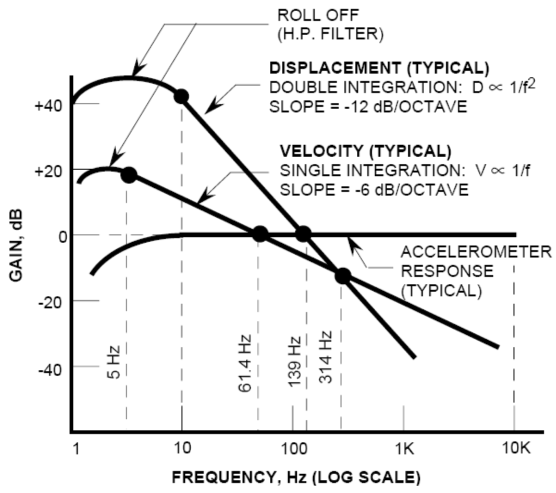


Figure 2-10 - [From SKF CM1001]

Signal-to-noise (S/N) ratio

It is very important to have a high signal-to-noise ratio. We do not want noise to mask the important vibration signal.

Noise can come from a number of sources:

- Noise can come from the process itself or external machines.
- Noise can come from electronics and transducers.
- The S/N ratio may be low if the signal is low:
    - Transducer output is low
        - Displacement probe >1 KHz
        - Accelerometer at low frequency
        - Quiet machine

One way to avoid this problem is to avoid integration.

- View data from displacement probes in mils or μm
- View data from accelerometers in g's or mm/sec^2

Some systems offer digital integration as well as analog integration. The benefits of digital integration are:

- Signal is digitized while still in acceleration
- Converted to velocity (or displacement) in software
- Spectrum and optionally time waveform can be converted to any units upon demand. (*Some systems cannot integrate the waveform.)*

Low speed machines

The issue of unit conversion and integration are most important when testing low speed machines (below 3 Hz or 180 RPM). The vibration of interest may be low amplitude, and will be low frequency. To improve the signal-to-noise ratio:

- A special accelerometer should be used:
  - Very low noise
  - High sensitivity (e.g. 500 mV/g instead of 100 mV/g)
- Beware of thermal and physical transients.
- Mount the sensor in the load zone.
- Be careful with the mounting method.

Note that the frequency response of the sensor will change if you use high sensitivity accelerometers.

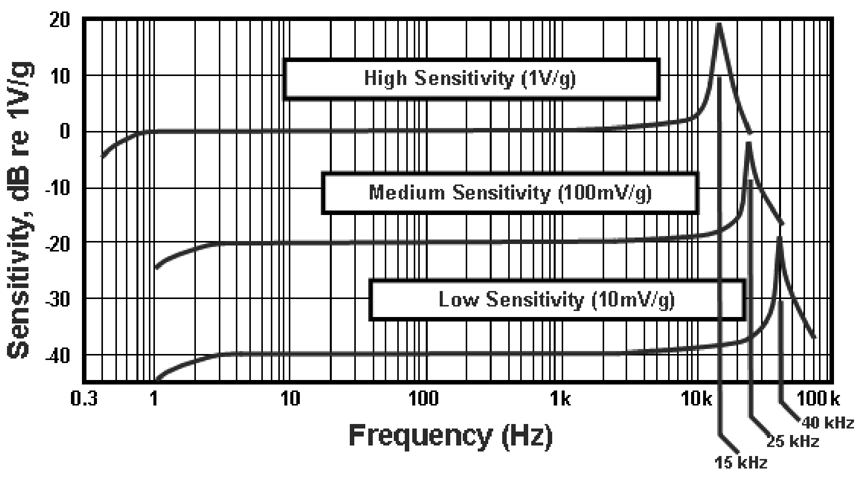


Figure 2-11

Sampling and Aliasing

Now we will look at what really happens to the signal from the transducer. Sophisticated software and hardware capture and manipulate the signal to yield data that can be used for accurately determining machinery condition.

Sampling the signal and the sample rate

When the transducer is placed on the machine, its electrical output is a continuous analog signal representing the vibration at that location.

It contains all of the information available from the machine proportional to either velocity, displacement, or acceleration - depending upon the type of transducer.

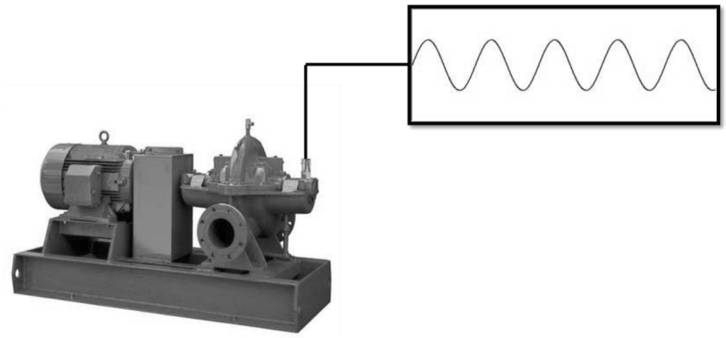


Figure 2-12 Continuous Analog signal from the sensor

The signal must be digitized to enable the analyzer to process, display and store the vibration. To do this the continuous analog signal is converted into a series of discrete numbers called a **time record**.

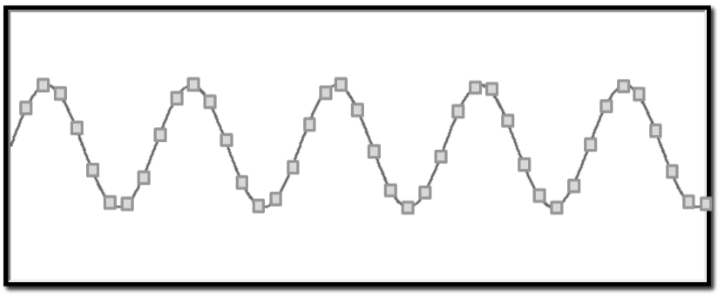


Figure 2-13 The time waveform is sampled at discreet intervals. Each point is an individual sample. The rate of the sampling is called the **Sampling Rate.**

Figure 2-13 shows the process of digitizing a standard analog waveform. It has been marked where individual samples have been made at discreet time intervals. The samples are converted to numbers and stored so the waveform can be rebuilt later and build the FFT.

The rate at which the waveform is sampled is called the **Sampling Rate or sample rate.** It is the number of samples made in one second.

To represent the time waveform, the data collector only has the individual samples to work with. Anything that happened between the samples is lost. When the data collector or software draws the time waveform it simply "connects the dots." This is called the **Time Domain.**

What is actually stored in the data collector is therefore not as detailed as the original continuous analog signal.

Triggering

During routine data collection, the sampling process begins when you press the button to start the measurement. However there are a number of tests where you may wish for the system to delay the data acquisition until an event occurs. Most commonly that event would be a tachometer pulse (during time synchronous averaging or phase measurements), or the impact of a hammer during modal testing.

Phase, order tracking and TSA

Your data collector will have an external trigger input. The data collector can be instructed to wait until the pulse is detected from the tachometer signal before acquisition begins.

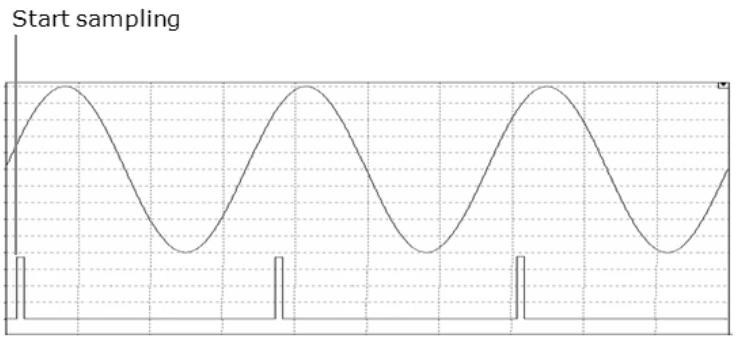


Figure 2-14

- Single channel phase:
    - Requires the external signal.
    - The data collector measures the lag between the tach and the high spot
- Time synchronous averaging:
    - The waveform capture is synchronized to the tach pulse.

- Order tracking:
  - The sample rate is controlled by the tach pulse rate.

These applications are discussed in greater depth later in this chapter.

Data driven trigger

Instead of using a tachometer, the data itself can trigger the data acquisition. In this case we have the impact from a hammer strike during a modal analysis test.

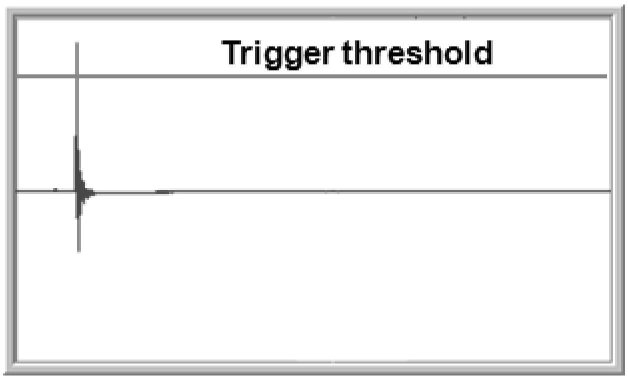


Figure 2-15

Pre- and post-trigger delays

It is possible to instruct the data collector to buffer the data so that it may begin sampling before the trigger arrives. In this example we have acquired data that came in before the impact. This is pre-trigger. It is typically defined as a percentage of the time record.

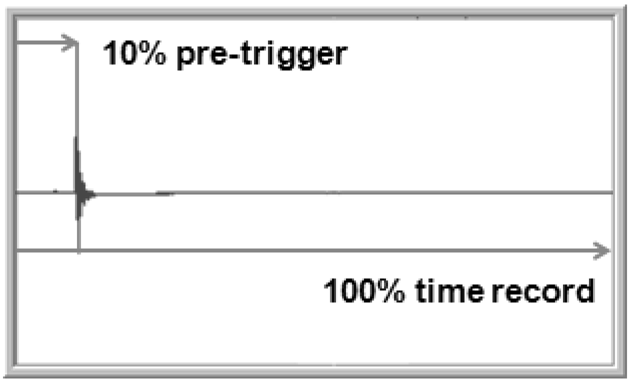


Figure 2-16

If triggering is used you must be sensitive to input gain. Normally a data collector will set its gain when it first begins recording the vibration signal – this will introduce a delay. When triggering we cannot afford that delay. It is best to set the gain ahead of time.

The Frequency Domain

The FFT process takes the time waveform and creates the spectrum which is referred to as the **Frequency Domain**.

Fast Fourier Transform or FFT

The vibration spectrum is the result of a Fourier Transform, named after the mathematician who developed the equation which turns a complex wave into its harmonic components.

Two other mathematicians, Cooley and Tukey, in 1965 developed an algorithm to effectively speed up the processing of the Fourier Transform and thus the name Fast Fourier Transform or FFT.

$$f(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin(2\pi f_i + \phi_i)t$$

Fourier, Jean Baptiste Joseph - French baron, physicist, mathematician 1768 - 1830

Discrete Fourier Transform (DFT)

The DFT is the broader description of creating the spectrum from the time series, with a finite time series. The DFT process can be used to transform a time series of any length to a spectrum – but it is very time consuming. The FFT was derived as a fast algorithm to compute the DFT. To achieve the speed, the number of samples in the time series must be a power of 2: 256, 512, 1024, 2048, etc.

Limitations of the FFT

The FFT works best on waveforms that include only periodic signals. It will build sine waves out of those signals. Each peak in the spectrum represents a sine wave.

- Amplitude: $\mathbf{a_i}$
- Frequency: $\mathbf{f_i}$
- Phase: $\boldsymbol{\theta_i}$

$$f(t) = a_0 + \sum_{i=1}^{\infty} a_i \sin(2\pi f_i + \phi_i)t$$

When you compute the FFT, each line in the spectrum contains a phase component. Unless we perform cross channel measurements, or specifically take a phase measurement, the phase is discarded. We will explore a number of techniques where the phase can be put to good use.

Time Records and FFT Records

A time record is defined to be N consecutive, equally spaced samples of the input. For computational reasons, N is a power of two; typically 1024, 2048, or 4096.

The FFT is a calculation that creates a spectrum from the time record. The resultant spectrum actually contains N/2 lines (i.e. 512, 1024, 2048) – this will be explained further in the aliasing section.

The spectrum only gets half as many numbers to work with because each frequency line actually contains two pieces of information, amplitude and phase. The phase is discarded.

Note that the terms: “bins” or “lines,” not “samples” are used when discussing the spectrum.

- $N = 2^9 = 512$ ––> 256 numbers
- $N = 2^{10} = 1024$ ––> 512 numbers
- $N = 2^{11} = 2048$ ––> 1024 numbers
- $N = 2^{12} = 4096$ ––> 2048 numbers
- $N = 2^{13} = 8192$ ––> 4096 numbers

Nyquist Theorem / Criterion

The sample rate affects the range of waveform that can be reconstructed.

In Figure 2-17 a 50Hz signal is sampled at a sampling rate of 50 times per second (50 Hz). This yields one sample during each cycle. Connecting the sample dots yields a straight line. The waveform cannot be reconstructed. There are not enough samples.

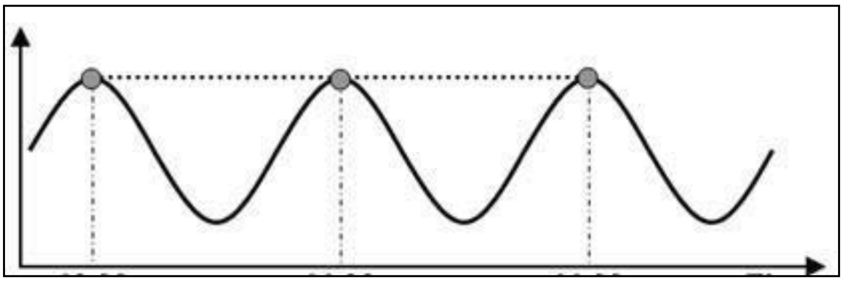


Figure 2-17 A 50 Hz signal sampled at a rate of 50 times per second yields one sample per cycle. Connecting the dots yields a straight line.

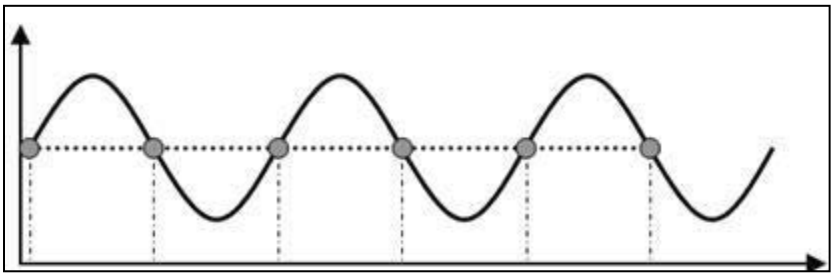


Figure 2-18 The 50Hz signal sampled at a rate of 100 Hz. Connecting the dots still yields a straight line.

In Figure 2-18 the sampling rate has been doubled to 100 Hz. But connecting the dots again yields a flat line. There are still not enough sample points to reconstruct the waveform. The waveform must be sampled at a rate greater than twice the frequency of the waveform.

The **Nyquist Theorem** or **Nyquist Criterion** states that the waveform must be sampled at a rate greater than twice the input signal. This frequency is often referred to as the Nyquist Frequency.

Most digital analyzers sample at 2.56 times the maximum frequency of interest, which will be explained shortly.

A digital vibration measurement consists of a large number of samples, enough to be representative for the time record. The speed at which these samples are taken depends on the maximum frequency to be displayed when the spectrum is produced.

The sampling frequency must be higher than twice the maximum frequency in the spectrum. To get a spectrum line at 400 Hz it must sample at faster than f_s = 800 Hz, that is, collect more than 800 readings per second. This means that there is less than 0.00125 seconds for each reading.

As will be explained shortly, we use 2.56 as the multiplier. If the maximum frequency of interest (often called the Fmax) is 400Hz, the **sample rate** would be 400x2.56 = 1024 times each second.

We can control 2 parameters: The Sample Rate and the Total Number of Samples collected.

The Sample Rate is controlled by selecting the Fmax.

The Number of Samples is defined by the Lines of Resolution.

Sampling time: ts

Sampling Frequency: fs (fs = 1/ts)

Nyquist theorem: fs > 2 x Fmax

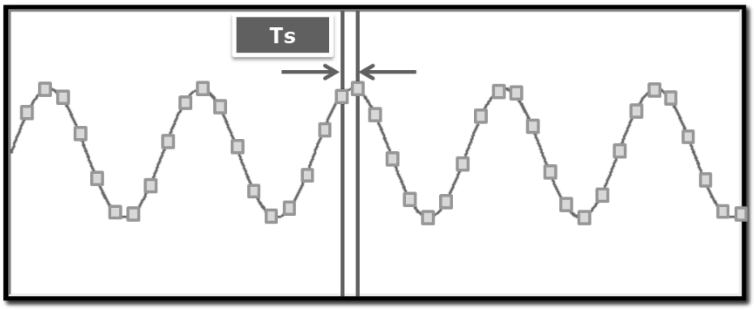


Figure 2-19 - the sampling frequency depends on the maximum frequency to be displayed in the spectrum

Aliasing

Sampling means that we measure a continuous real time event in a number of points and then try to reconstruct the event by connecting these points. If we do not sample fast enough, the reconstruction will not match the event. In Fig. 5-14, the reconstructed curve has the right amplitude but the wrong period. This effect is called "aliasing" and the effects must be corrected.

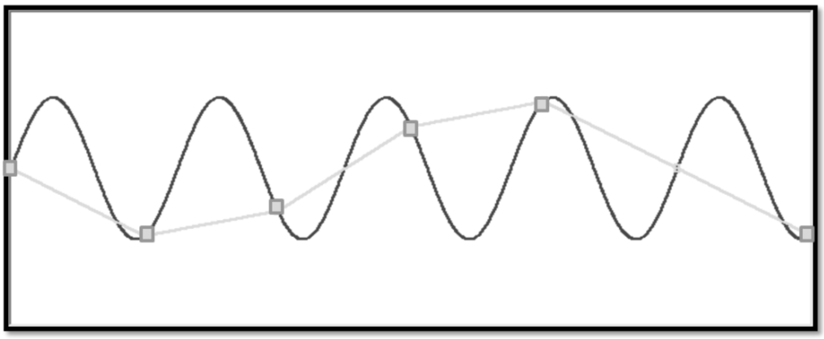


Figure 2-20 The effects of aliasing is that a false signal is constructed. It has the correct amplitude but the wrong period.

Two signals are said to alias if the difference of their frequencies falls in the frequency range of interest. This difference frequency is always generated in the process of sampling.

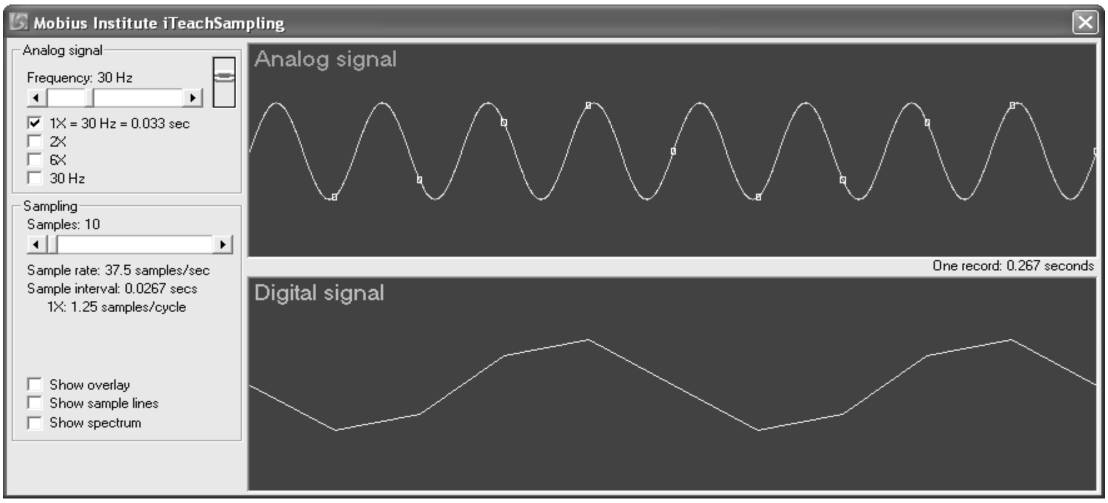


Figure 2-21 The analog signal showing points of sampling and the digital reconstruction.

In Figure 2-21 the analog signal is not sampled at a rate fast enough to reconstruct it. Instead, it generates a waveform at a much lower frequency. Note that the analog signal is a real signal, but there are not enough samples to reconstruct it accurately. The reconstructed waveform is at a much lower frequency, it is an aliased signal; a signal that appears real in the spectrum but does not really exist.

Figure 2-22 shows an analog signal with a 1x signal of 48 Hz and a 6x signal at 288 Hz. The Sampling Rate is 960 samples per second. So the 1x peak is sampled 20 times per cycle and the 6x pk is sampled 3.333 times per cycle.

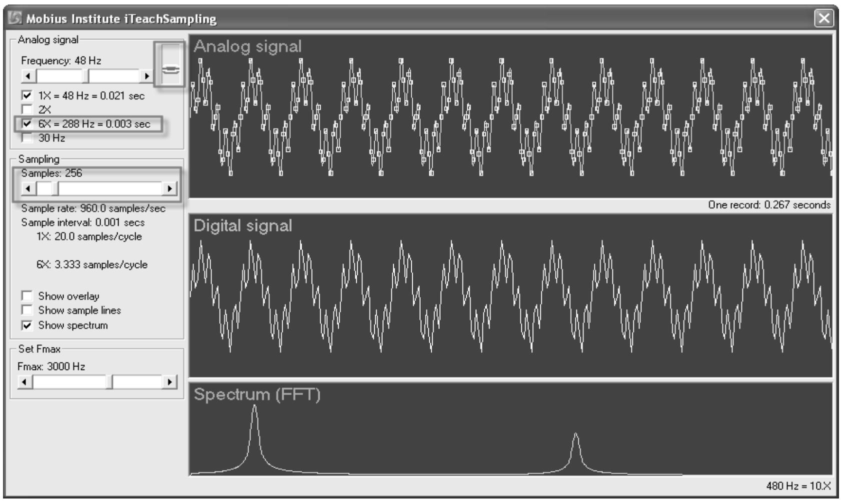


Figure 2-22 The 1x and 6x signals are sampled more than twice per cycle so they are reconstructed accurately in the digital waveform.

Since both signals are sampled more than 2x their frequencies, they are both reconstructed accurately in the digital waveform. Both signals are also accurately reflected in the spectral data.

Another example, Figure 2-23, has a 65 Hz signal. The spectrum has an Fmax of 100 Hz so the sampling rate is 256 samples a second. The peak in the spectrum is at 65 Hz.

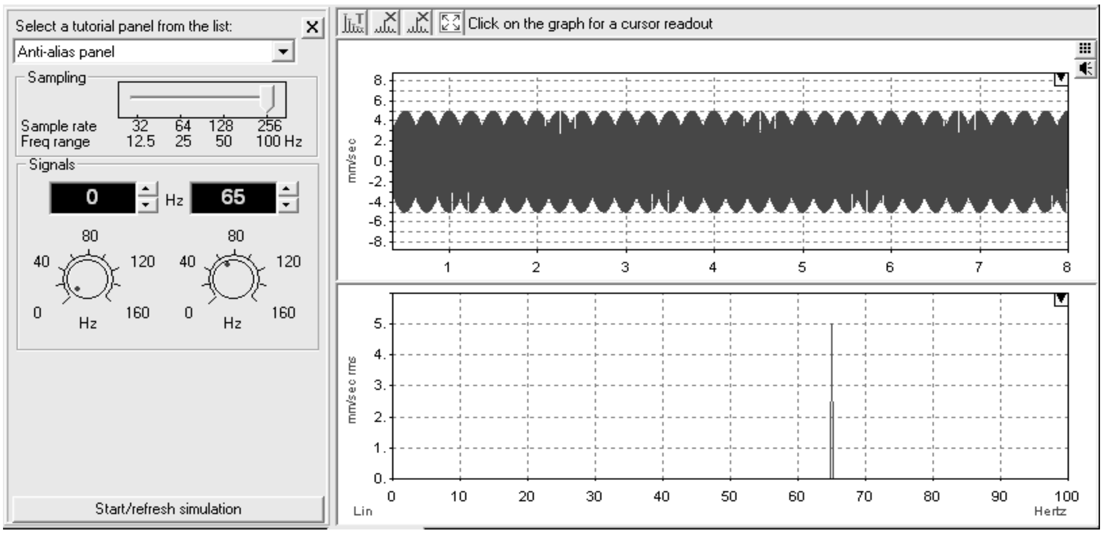


Figure 2-23 A signal is sampled at a rate of 256 times per second includes a 65 Hz signal. It is at 65 Hz in the spectrum.

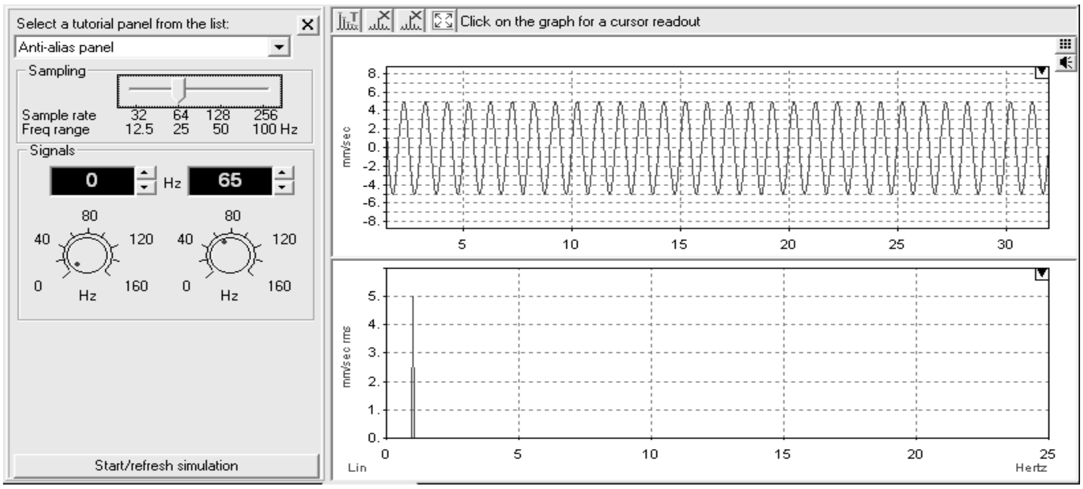


Figure 2-24 A 65 Hz signal is sampled at a rate of 64 times per second. The aliased peak shows up at 1 Hz.

Figure 2-24 is the very same signal but sampled at a rate of 64 samples per second. The frequency range is 25 Hz. The 65 Hz signal is in the analog signal but cannot accurately be reconstructed because the sampling rate is too low. It placed the aliased signal at 1 Hz.

Coping with Aliasing

To demonstrate the aliasing process, we will begin by assuming that we have a signal of interest of 20 Hz. We therefore select an Fmax of 25 Hz.

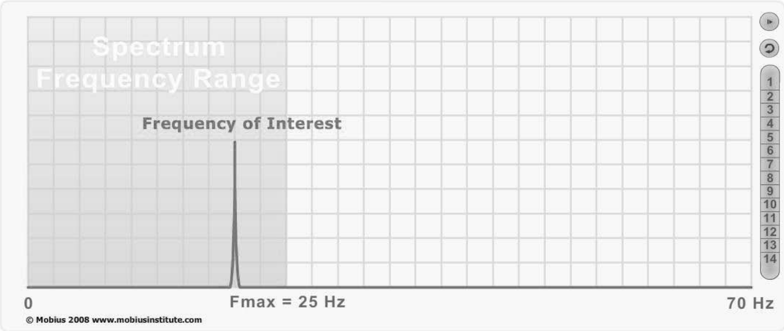


Figure 2-25

To begin with we are using the standard Nyquist criterion which states that we must sample the signal at over twice the greatest frequency of interest. Fmax is our greatest frequency of interest, so our sampling rate is simply greater than 2xFmax Hz.

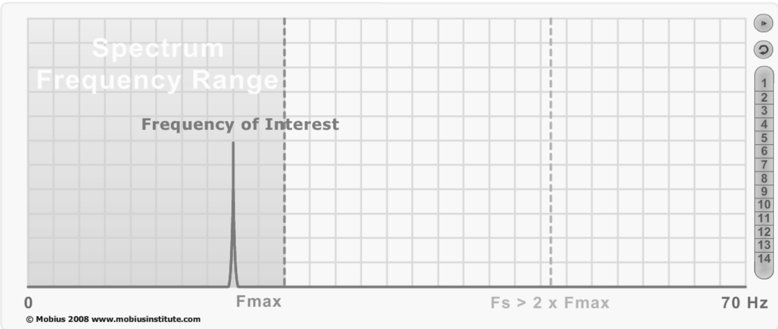


Figure 2-26

However, what we find is that if there also happened to be a frequency (Fin) in the input signal that was higher in frequency than half the sampling frequency (Fs), it would alias. A peak would appear at the sampling frequency less the actual frequency (Fs-Fin). Therefore our spectrum would contain a peak even though there was not a signal at that frequency.

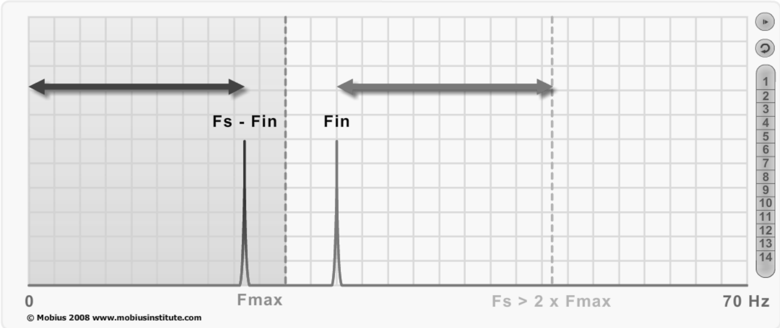


Figure 2-27

No matter what that Fin frequency happened to be, any frequency above Fs/2 would alias.

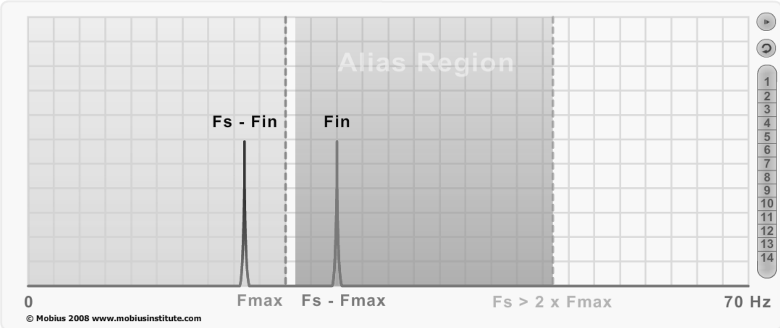


Figure 2-28

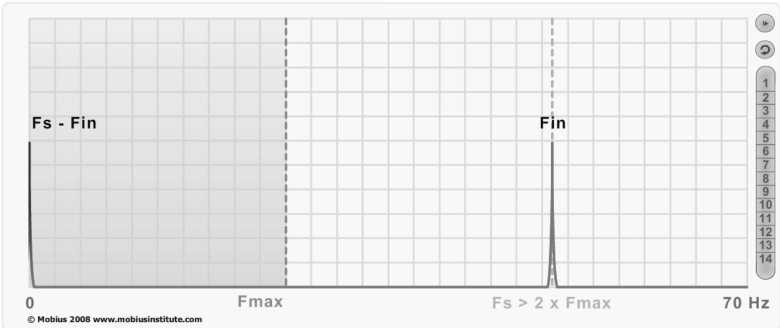


Figure 2-29

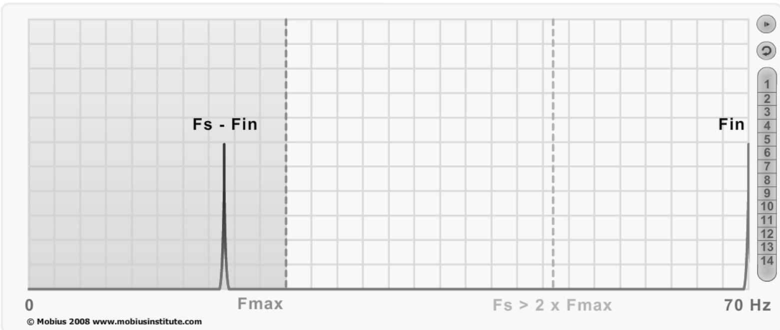


Figure 2-30

The solution is to add a filter – an anti aliasing low pass filter. We could filter out all of these frequencies above Fmax so that nothing would alias.

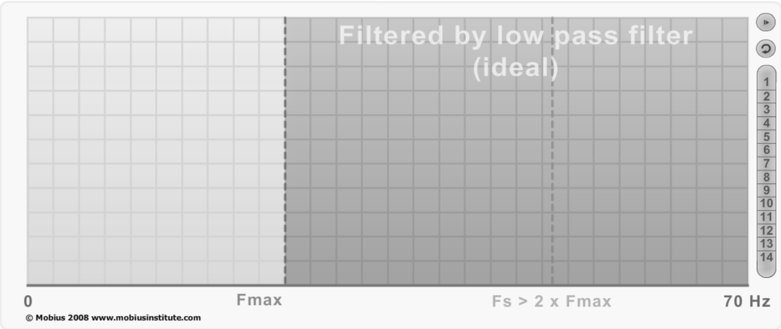


Figure 2-31

The only problem is that there is no such filter that is so perfect that one frequency is completely blocked, and a slightly lower frequency is allowed to pass unscathed. Instead, we can draw a curve (representative of the older analog filters) that shows just how much of the signal is allowed to pass.

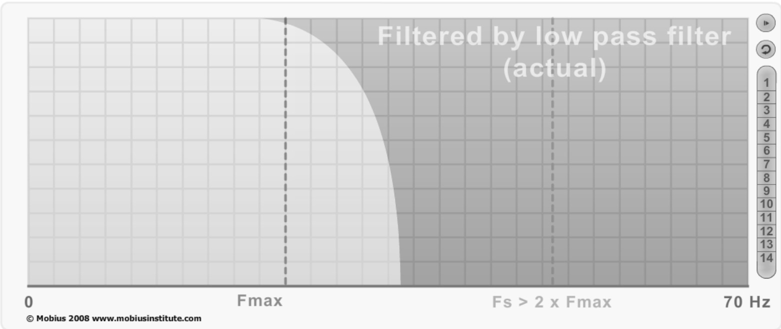


Figure 2-32

What that means is that any signal that is above half of the sampling rate, and below the point at which the filter is adequately blocking the frequencies, will still alias (albeit at a lower amplitude).

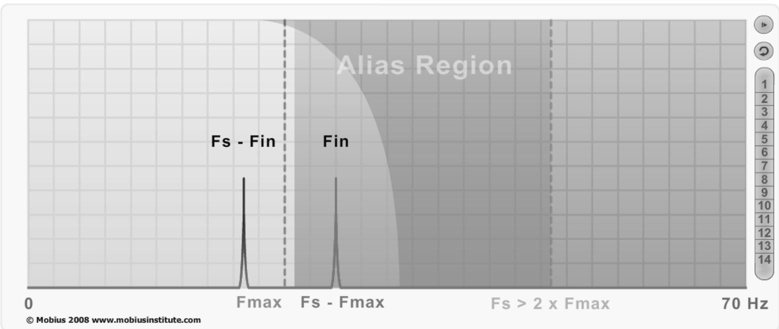


Figure 2-33

The solution is to set the sampling frequency much higher, but leave the filter where it was (i.e. low pass at Fmax). For convention, the industry selected 2.56 as the multiplier of choice. If we desire a given Fmax, we will sample at 2.56xFmax.

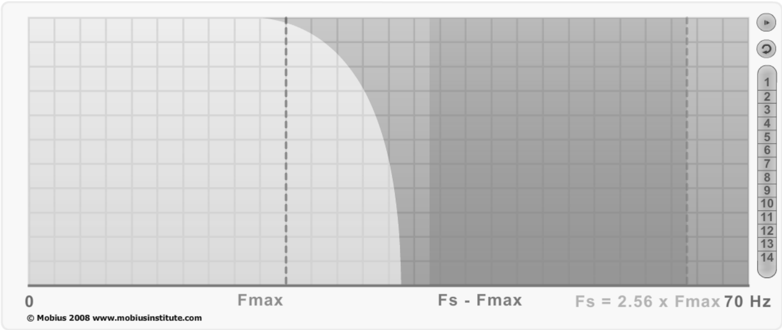


Figure 2-34

Any signal that has a frequency above half the Fmax will be rejected completely – the filter will make sure of that.

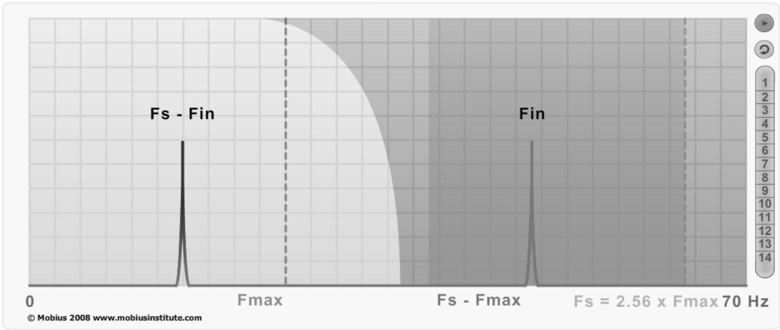


Figure 2-35

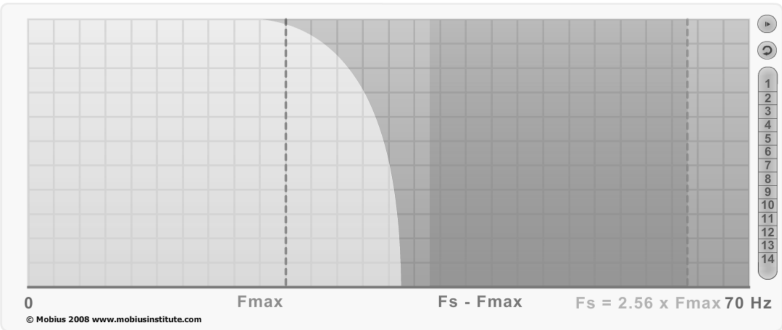


Figure 2-36

That does still leave a "grey area". Some signals will still alias into the area between Fmax and half Fs. So we simply discard that data. All of the data representing the frequencies above Fmax are "thrown away" – even though the FFT calculation provides data up to Fs/2.

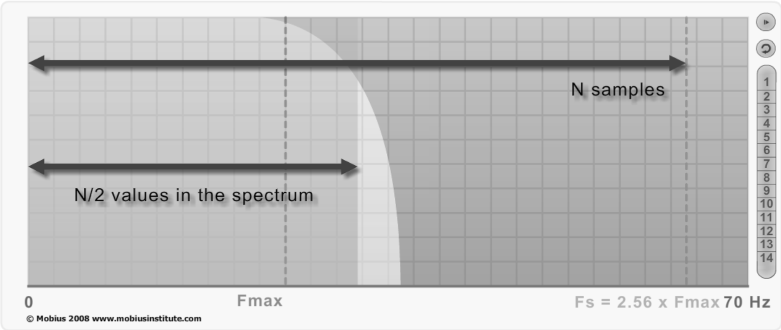


Figure 2-37

The result is that if we have 2048 samples, the raw FFT will contain 1024 amplitude values, but we discard the data after the 800th sample (2048/2.56 = 800).

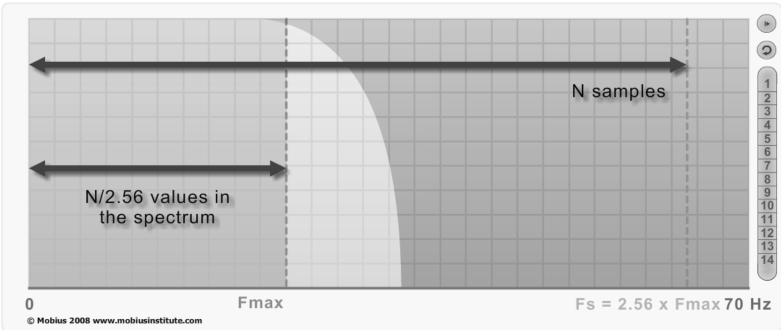


Figure 2-38

So, not only should that discussion explain how aliasing works (and how anti-aliasing filters work), it should also explain where the number 2.56 comes from.

The bottom line on aliasing

Two steps are taken to eliminate the effects of aliasing.

First, a low pass filter is applied before sampling so that the signal components that may have been aliased will be removed.

Secondly, recall that for N samples in the time record the FFT produces a spectrum with N/2 lines. So a 2048 sample time record produces a 1024 line spectrum. However, to correct for anti-aliasing, only 800 lines are kept. The accepted rule is that the number of lines in the spectrum is equal to the number of samples in the time record divided by 2.56.

- $N = 2^8$ = 256 ––> 100 lines
- $N = 2^9$ = 512 ––> 200 lines
- $N = 2^{10}$ = 1024 ––> 400 lines

- N = 2^{11} = 2048 ––> 800 lines
- N = 2^{12} = 4096 ––> 1600 lines
- N = 2^{13} = 8192 ––> 3200 lines

Sigma-Delta

Since the advent of digital signal processors, there are now two ways to solve the aliasing dilemma. One is to use an analog anti-aliasing filter and the other is to "over-sample" the incoming signal (i.e. sample it more quickly than needed) and use "digital" filtering and "re-sampling" to reconstruct the signal of interest. This process is called the "sigma-delta" method.

Most modern data collectors now employ the sigma-delta method, as it has better performance, and they are less expensive to manufacture. In reality it uses a much more effective filter which would allow many more lines to be kept in the spectrum. For convention, the industry has kept the same number of lines as we used in the past.

Sampling and Resolution

Taking a closer look at the sampling process leads to some conclusions. The sample rate must be 2.56 times the highest frequency of interest. To recreate a signal of 1000Hz requires a sampling rate of 2560Hz. This produces a spectrum with an Fmax (maximum frequency) of 1000Hz. This yields Rule #1.

Rule #1

Sample Rate is proportional to frequency range

(The Sample Rate is 2.56 x Fmax)

Another relationship to remember involves the waveform and spectrum. The number of samples in the time waveform determines the bandwidth or resolution in the spectrum. In most analyzers the **bandwidth or resolution** is set by the **"number of lines"** chosen by the operator. This number is referred to as **Lines of Resolution** or **LOR**.

One "line" is one data point in the spectrum. It is sometimes referred to as a cell or bin. Two frequencies cannot be resolved if they are closer together than the lines. The bandwidth or resolution of each line is calculated by dividing the Fmax by the number of lines.

For example, a spectrum with a maximum frequency of 1000Hz and 400 lines results in a maximum resolution of 1,000/400=2.5Hz. This means that two discreet frequencies will appear as one if they are closer than 2.5Hz apart.

If the Lines of Resolution is increased to 800 lines, the resolution is 1000/800=1.25Hz.

The Number of Lines is proportional to resolution

But this is not the whole story. There is a link between the number of lines in the spectrum and the number of samples in the time record. The number of samples required is **2.56 times the number of lines (LOR).** For 400 lines in the spectrum, the number of samples is 400 x 2.56 = 1024. Rule #2 in its broadest scope then is:

Rule #2

The Number of Samples is proportional to resolution

(Number of Samples = 2.56 x LOR)

The net result is that good resolution can be achieved. It requires a greater number of samples and increases the time to collect the data.

To resolve frequencies that are close together, more samples are needed.

Number of samples 'N' is proportional to resolution

- N = 256 ––> 100 lines
- N = 512 ––> 200 lines
- N = 1024 ––> 400 lines
- N = 2048 ––> 800 lines
- N = 4096 ––> 1600 lines
- N = 8192 ––> 3200 lines

Sample Time

There is a trade off in results vs. time. The faster the data is sampled the more data that is collected – which has to be stored, transferred, graphed, and used in calculations. The more data collected to get greater resolution, the longer it takes to collect it, which means more time at the machine.

The time required to collect the samples is equal to the number of samples required, divided by the sampling rate.

For example: an Fmax of 400 Hz sets the sampling rate at 1,024Hz. A resolution of 400 lines sets the number of samples at 1,024. So the sample time is 1,024/1,024 = 1 second.

Another way to compute the Time for one average is to use the numbers that we can control, i.e. the Fmax and LOR. So **T = LOR / Fmax.** Using this simplified formula for the previous example yields… T = 400LOR/400Hz Fmax = 1 second.

Another example using the simplified formula for an Fmax of 100 Hz and 12,800 lines of resolution…. T = 12,800LOR/100Hz Fmax = 128 seconds to collect one average.

Figure 2-39

Example: Sampling Rate

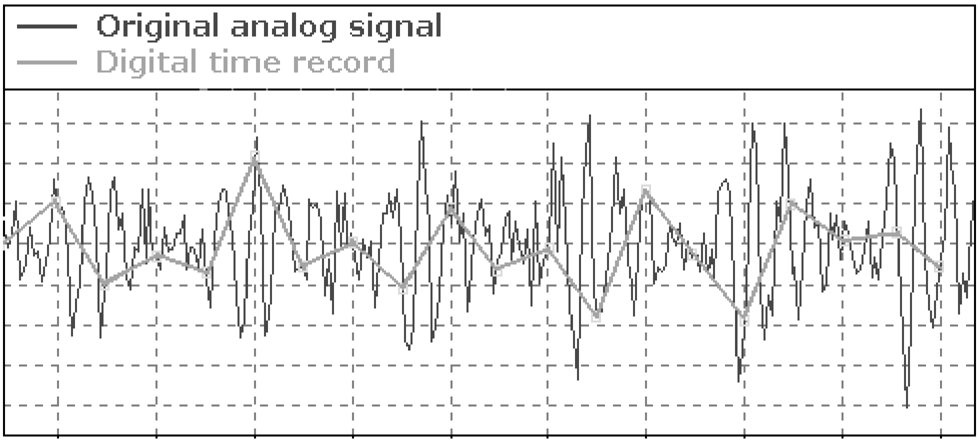


Figure 2-40 The sampling rate is set too low to capture the high frequency occurrences in the waveform.

In Figure 2-40the sampling rate is too low to capture any of the high frequency signal. Information contained in the waveform signal will be lost.

Example: Resolution

Figure 2-41 is a waveform which contains 3 signals, a 6 Hz, an 8 Hz, and 380 Hz signal. However, there are only two peaks in the spectrum.

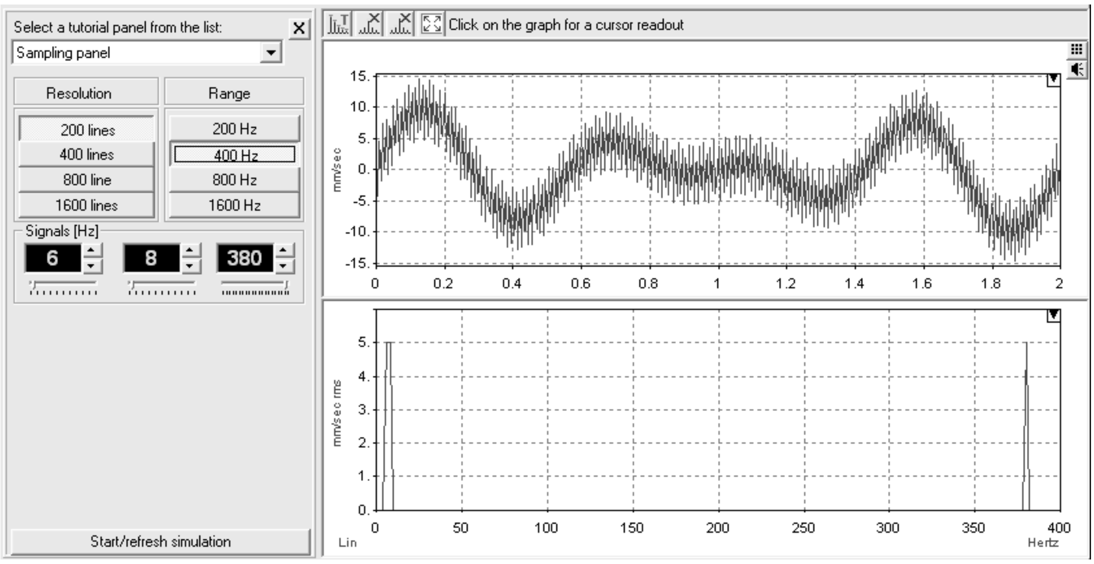


Figure 2-41 There are three signals but only two peaks

The frequency range is set to 400 Hz which is a sampling rate of 1024 samples per second. This captures the 380 Hz signal.

However the number of samples is 512 which produces a 200 line spectrum. The 6 Hz and 8 Hz signals combine and produce a peak around seven Hz. The resolution is not good enough to separate the peaks. The resolution is 2 Hz (400 Hz Fmax /200 LOR). This only separates signals greater than 2 Hz apart.

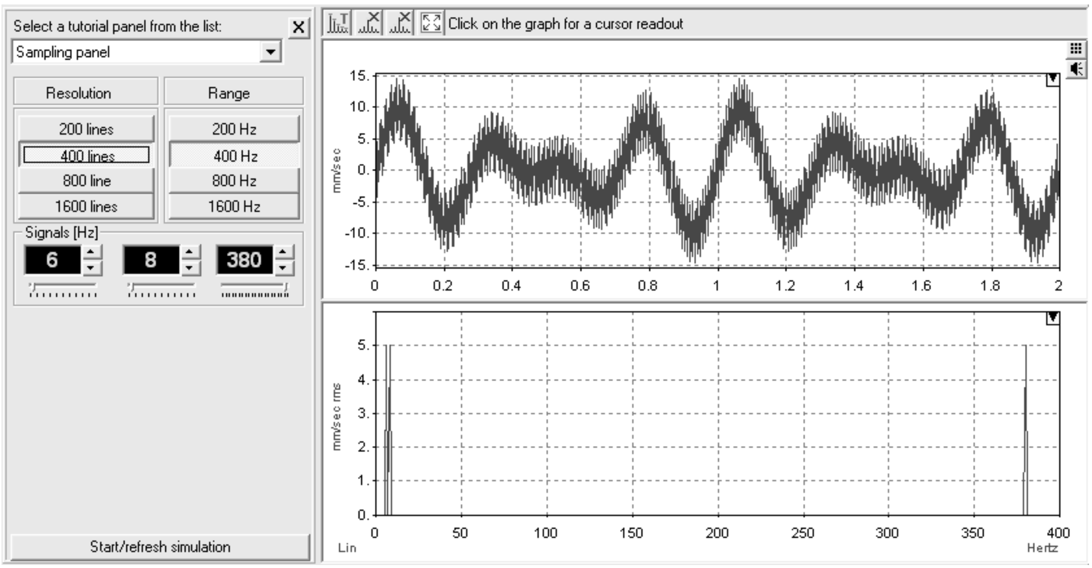


Figure 2-42 The Resolution is increased to 400 lines and the two signals are now separate peaks

Figure 2-42 is the same signal but with 400 lines of resolution, 1024 samples in the waveform. The resolution is now 1 Hz and the two signals have their own discreet peaks.

It takes a combination of Sampling Rate and No. of Samples (Fmax and LOR) to capture and see the data that may be needed.

Figure 2-43 has two signals, a 15 Hz and a17 Hz signal. The Fmax is set to 800 Hz and the number of lines is 800 lines. The resolution is one Hz. The spectrum is zoomed in to show only the first 25 Hz. The dotted lines indicate the lines of resolution.

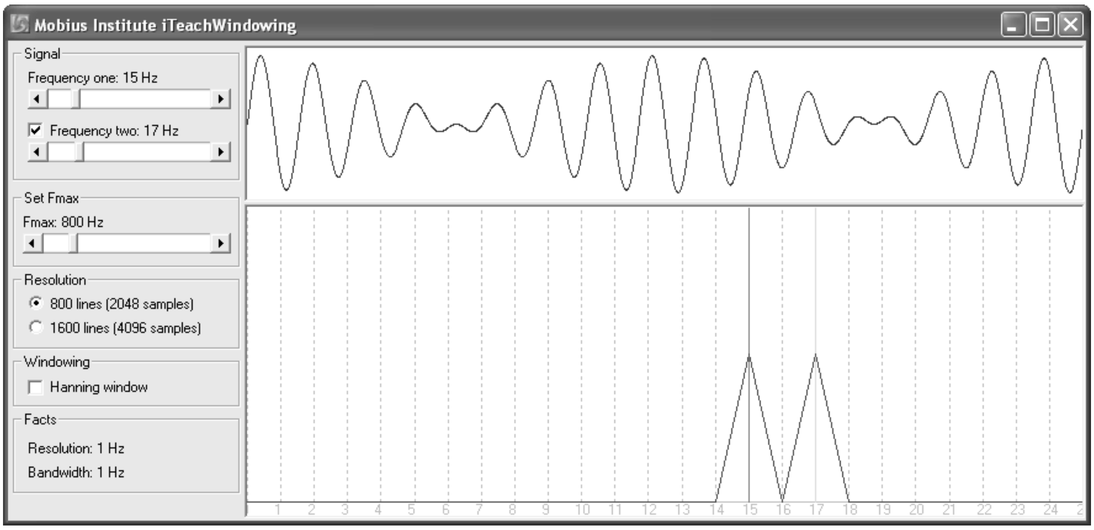


Figure 2-43 Two peaks are separated but smeared over two lines of resolution.

The peaks are "smeared over two lines of resolution because there is not enough resolution to make them more precise. So although the resolution is 1 Hz, these peaks are 2 Hz wide.

NOTE: The Resolution is further limited by the Window factor which will be discussed in a later section. See ***Window Type Affects Frequency Resolution*** *in the Windowing section.*

The illustration below shows part of a spectrum with a range to 800Hz. It has 800 lines of resolution. The frequency resolution is the Fmax / number of lines or **Fmax / LOR**.
800/800 = 1Hz

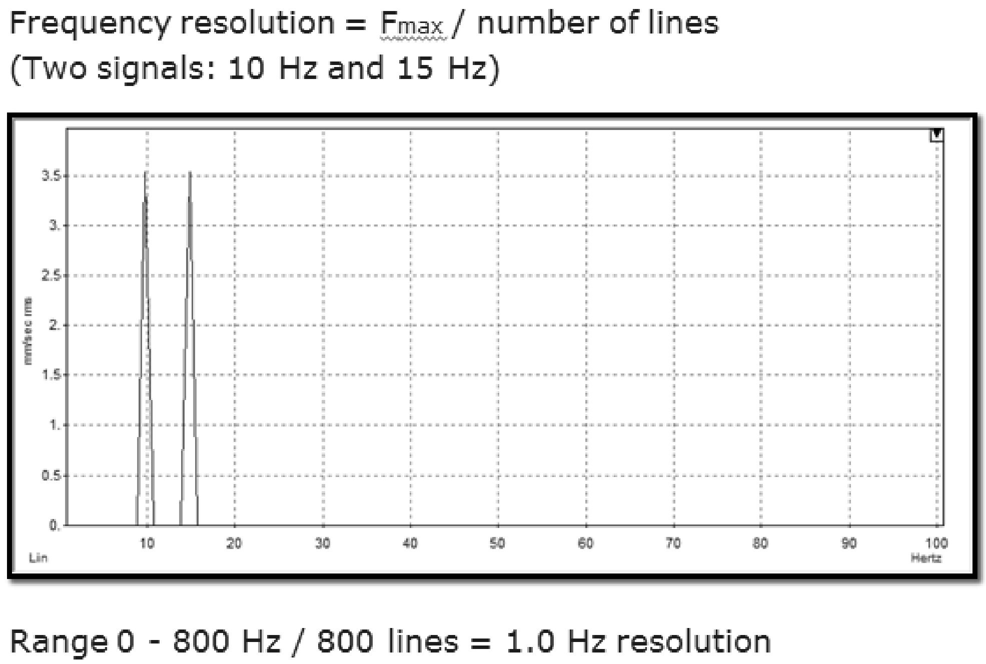


Figure 2-44

Table 2-1 lists the increasing number of lines for a constant Fmax showing the resolution or bandwidth, the sample rate, the number of samples and the sample time.

Number of lines	Fmax (Hz)	Resolution (Hz)	Sample rate	Number of samples	Sample time (sec)
400	1000	2.50	2560	1024	0.4
800	1000	1.25	2560	2048	0.8
1600	1000	0.63	2560	4096	1.6
3200	1000	0.31	2560	8192	3.2
6400	1000	0.16	2560	16384	6.4

Table 2-1 Effect of resolution on the Sample Time. Fmax and therefore the sample rate are constant.

Zoom measurements

When we set the Fmax of the data collector we get a spectrum from 0 Hz to the Fmax frequency. Deciding on the Fmax and resolution are very important decisions:

We need to see >3 x Gearmesh frequency
We would like to see pole-pass sidebands

We have three options:

1. Analysis zoom
2. Pseudo-zoom
3. Measurement zoom (or "true-zoom")

Analysis zoom

If you collect a 1600 line spectrum with an Fmax of 1600 Hz, we will have 1 Hz between each line in the spectrum. On a data collector or computer screen, we can't see each line in the spectrum – we may have more data than pixels. Even on a high resolution screen it is difficult to see the detail.

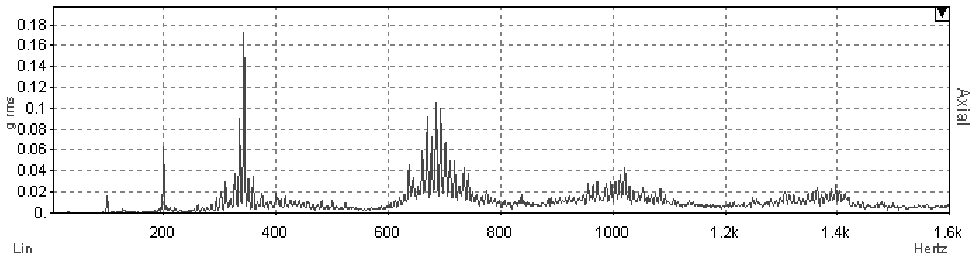


Figure 2-45

We can graphically zoom-in on the data of interest. We are not improving data resolution.

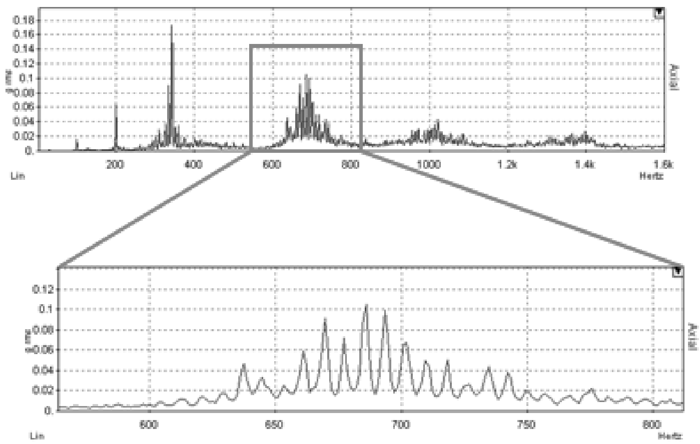


Figure 2-46

Pseudo-zoom

Many data collectors offer a pseudo-zoom option. Options will be presented with specific Fmin and Fmax combinations.

For example: Fmin: 400 Hz, Fmax: 800 Hz, 800 lines

The data collector is actually collecting a normal spectrum with Fmax = 800 Hz and 1600 lines. It only keeps the 800 lines between 400 Hz and 800 Hz. Options are limited due to the requirement to have the desired resolution in the desired frequency range.

True-zoom

True-zoom, also known as "band selectable analysis", performs a process called heterodyning on the incoming signal. In effect it shifts the desired frequencies into a baseband where the collector can use the normal FFT process.

If the desired Fmin is 400 Hz, and Fmax is 800 Hz, the signal is shifted down by 400 Hz, and then the FFT is performed with the desired resolution from 0-400 Hz.

The user is presented with data as 400 – 800 Hz.

Most analyzers/data collectors will offer limited resolution options, perhaps only 800 lines.

Dynamic Range

Up to this point the focus has been on the accuracy of the frequency. But there are concerns regarding the resolution of the amplitude. Vibration waveforms from machinery contain a great deal of information, and some signals are very low in amplitude compared with the more dominant frequencies.

Remember, when analyzing the data we do not just look at the high peaks, but also look at the lower peaks. Harmonics of bearing tones, for example, can be quite low in amplitude, but they are still critical to the diagnostic process.

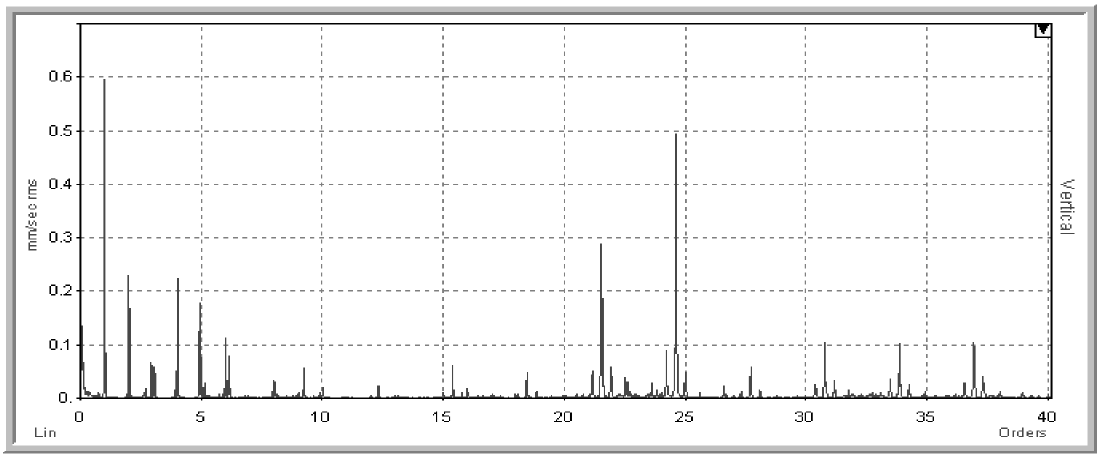


Figure 2-47 low level bearing peaks in the presence of other moderate peaks.

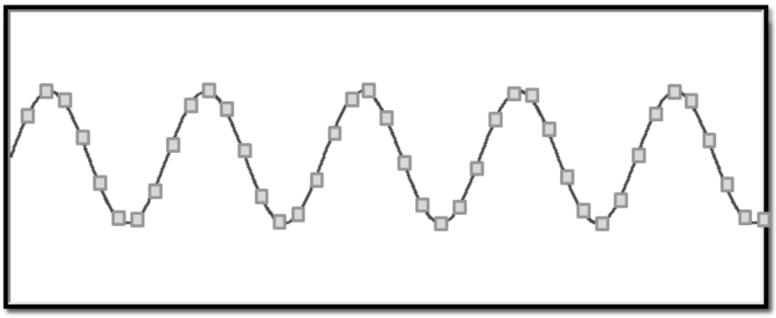


Figure 2-48

When the data collector digitizes the incoming signal it can only assign a limited number of amplitude values to each sample it takes. Let's take the extreme. If a digitizer (the A/D) was only 5 bit, that means it can assign one of 32 numbers (2^5 = 32) to the sample.

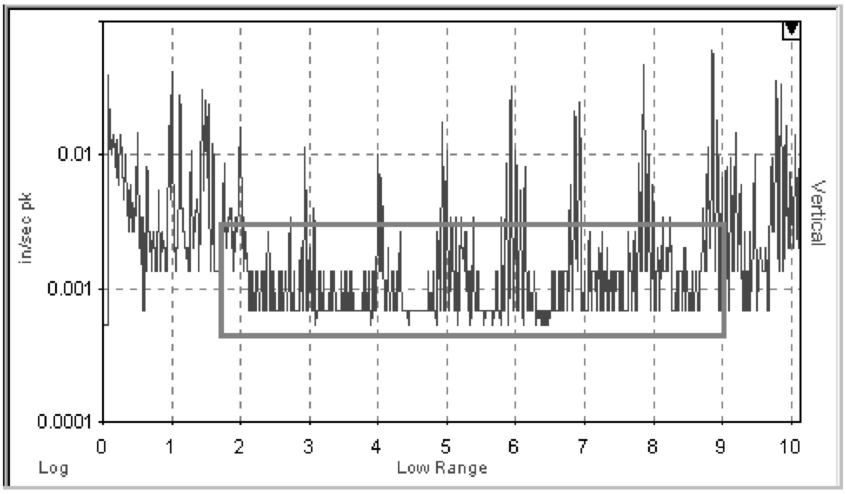


Figure 2-49

This is not nearly enough dynamic range. The waveform is chunky and the spectrum will be similar.

Dynamic range is a measure of the ability to 'see' small signals in the presence of large ones. It is a ratio of the smallest signal to the biggest, and is typically represented in decibel (dB).

For many years data collectors had 16 bit A/D converters which provides a dynamic range of 96 dB. This really means that if the input signal was 1 volt, we could still detect changes as small as 0.015 mV - which is excellent. The most recently released analyzers use 24 bit A/D converters!

Data collectors do not have only one input range. There is usually an amplifier (a gain stage) before the A/D to increase the amplitude to best suit the input range of the A/D converter.

If we go back to our 5 bit A/D example (Figure 2-48) the ideal situation exists when the signal is actually 1 volt in amplitude. All of the available input is being used.

The measurement quality of a data collector is actually a measure of the resolution of the A/D (for example 16 bit), and also its ability to amplify the signal so that the collector utilizes the majority of its input range. We would like to be able to change the range settings in small steps to cater for the wide range of possible signals.

While most data collectors will allow you to manually set the input range settings, in most cases you will allow the collector to auto-range - it will determine the best range automatically. It actually does this by trying each range setting to find which is best.

Auto-range vs. Auto-scale

Many people confuse auto-ranging with auto-scaling. They do not mean the same thing at all.

As just described, auto-ranging is the process of selecting the optimum gain setting. Auto-scaling is a simple graphical routine which takes the spectrum and displays it with the optimum display range.

Windowing

There is another property of the FFT which affects its use in vibration analysis. You may remember that the FFT is performed on a block of samples called the time record. One assumption made in the FFT calculation is that the time record is continuous. That is, the signal just before the captured time record, and the block immediately after our time record are identical.

In this example, although we are performing the FFT on the block of data with the black background, the FFT calculation "assumes" that the data continues endlessly before and after this block of data - as shown with the data with a gray background.

In this example it is true that the single frequency sine-wave begins and ends at zero amplitude. Four complete cycles live within the time record

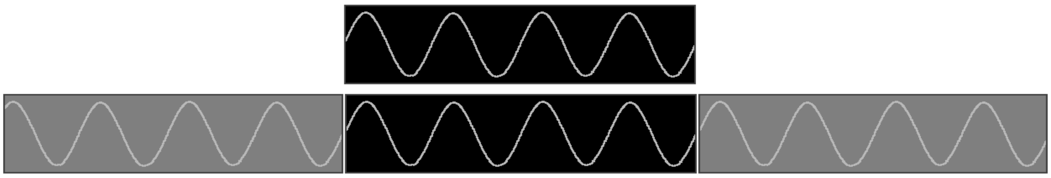


Figure 2-50 The FFT process assumes the block of data goes to zero at the ends. In this case it does.

If we are analyzing a pure sine wave, i.e. just one frequency, and there is an integer number of cycles in the time record, then this assumption is correct.

However it is seldom true that the time record starts and ends at zero. More commonly they are similar to Figure 2-51.

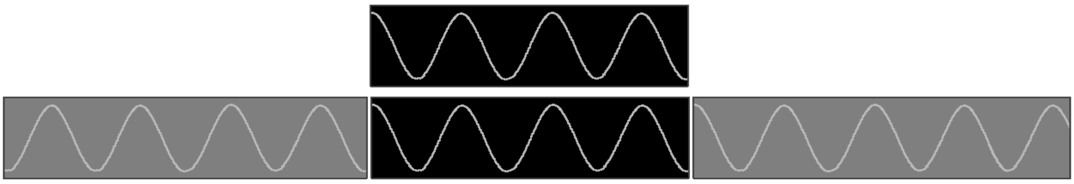


Figure 2-51 The time window does not start nor end at zero.

When the FFT calculation is performed the signal is discontinuous. It seems to have a step increase in level and looks similar to an impact to the FFT calculation. It generates a peak that is spread over a wide frequency band similar to an impact.

Remember that an impact generates energy in a wide frequency range. That is not what we want to see as a result of this.

The real data in Figure 2-52 shows that the ends of each sample do not have the starting and ending amplitude at zero.

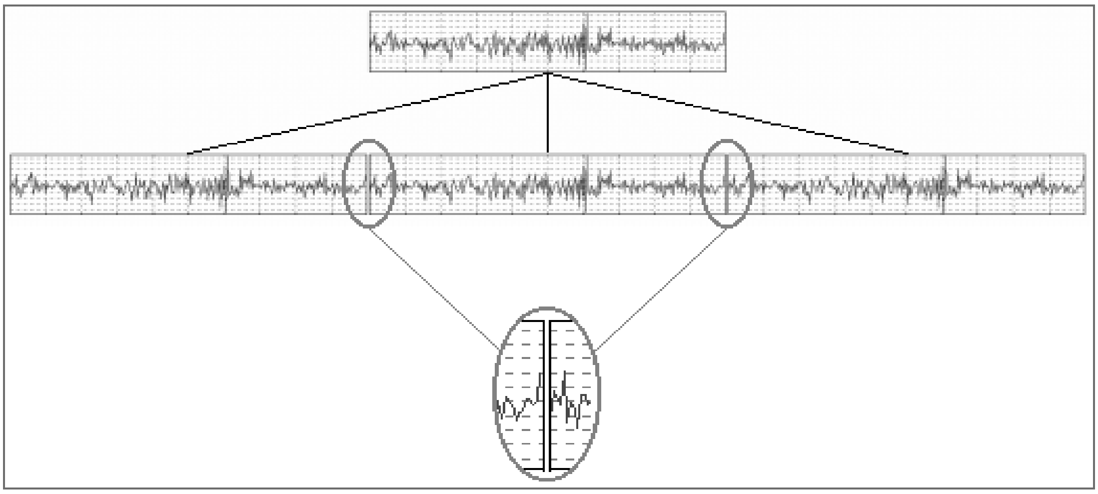


Figure 2-52 Real data example where the ends of the sample blocks do not end at zero amplitude.

This phenomenon is called **Leakage.** The result is a broadening of peaks within the spectrum.

To get around this problem the shape of the time record is actually changed so that there is no data at the beginning or end of the record. The ends are pushed to zero amplitude. This is known as **windowing** the data. The window has a minimal effect on the frequency content, but it does affect the shape of the spectral peaks and the amplitude levels.

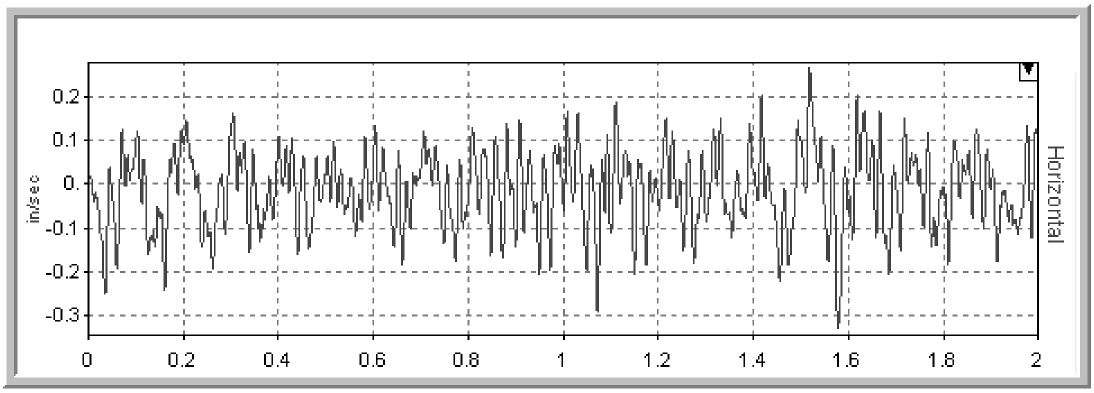


Figure 2-53 This is the raw data before any window is applied.

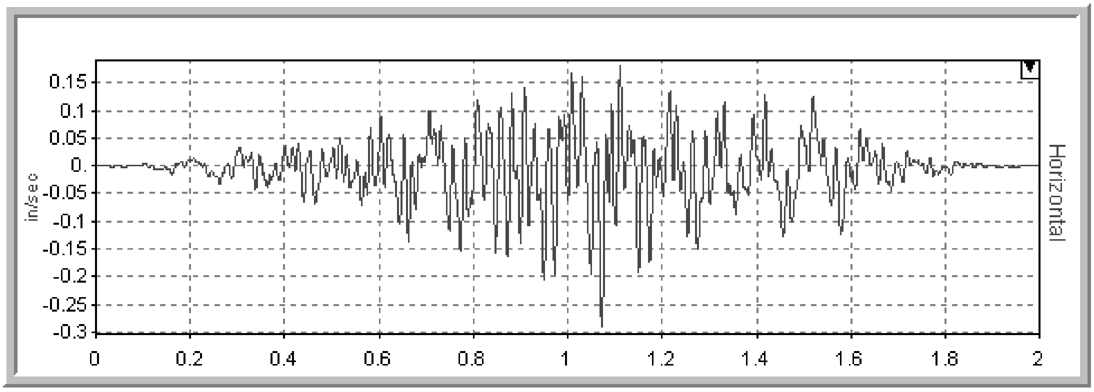


Figure 2-54 This is the same data with the Hanning Window applied. Bothe ends have zero amplitude

In this illustration, (Figure 2-54) it is clear that when the windowed time records are placed end to end there is no longer any sudden change in amplitude at the start or end of the record, and thus there is no leakage.

There are a number of window functions to choose from, each with a different shape, and each having a different effect on the resultant spectrum. The **"Hanning"** window is most commonly used in vibration analysis of rotating machinery. It yields the best frequency accuracy although it does affect the amplitude.

Window Type Affects Frequency Resolution

The Hanning window will result in a correct amplitude a when the frequency corresponds exactly with a spectrum line. However, there will also be "leakage": a spectrum line with a reduced amplitude to either side of the "true" line.

For a near miss, the Hanning window will produce two lines, both with an amplitude of 85%, plus small leakage.

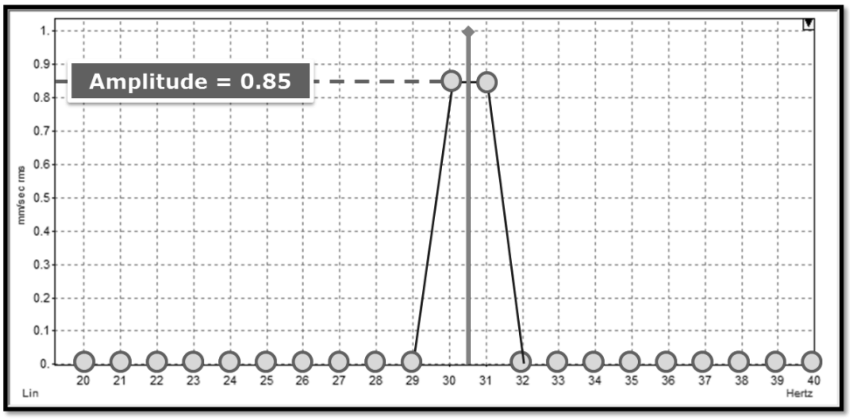


Figure 2-55 - the vibration occurs at 48.25 Hz but can only be shown at 48 Hz or 48.5 Hz, the resolution being 0.5 Hz

Although the Hanning Window has the best frequency accuracy, it still affects it. To correct for this phenomenon, a Window Factor is used. The Window Factor is a multiplier to apply when determining resolution.

The Window Factor for the Hanning Window is 1.5. This means that after determining resolution in the normal manner, the result is multiplied by 1.5. This effectively decreases the resolution. For example:

If the frequency range is1600 Hz and the Lines of Resolution is 800 lines, then the formula for determining resolution is 1600/800=2. Now multiply that by 1.5 and the true resolution using the Hanning Window is 3Hz. This means that two discreet frequencies that are closer than 3 Hz will appear as one peak.

Another way to express the resolution formula using a Hanning Window is:

$$Resolution = 1.5 \times \frac{F_{max}}{Lines}$$

Example: For a spectrum with an Fmax of 400 Hz and 400 lines of resolution **RES= Fmax / LOR = 400 / 400 x Window Factor = 400 / 400 x 1.5 = 1.5 Hz**

Other Window types include:

- Flat Top Window– has better amplitude accuracy but the frequency accuracy is poor.
- Hamming Window– similar to Hanning window except the ends do not go to zero amplitude.
- Rectangular, Uniform, or No Window– the Rectangular and Uniform are the same as having no window at all. These can be used during special testing situations and are discussed in Category II.
- Exponential Window - A special windowing function for minimizing leakage in lightly damped structures that is used in impact testing. In a lightly damped structure, oscillations may not die out within the sampled time data block, T, which results in leakage error. An exponential window adds damping to the time signal to force it to die out within the time T, thus minimizing leakage. The added damping is then removed mathematically after the signal is processed.

Most analysts use only the Hanning Window for routine measurements.

Importance of resolution

When deciding upon the data collection parameters for each machine, you must look at the machine, determine the rotation rates of the important components, and decide what the lowest and highest frequencies of interest will be. You should remember that you also need to analyze harmonics of key frequencies, which will dictate the highest frequency of interest - a gearmesh frequency or rotor bar rate for example.

A great many less-experienced vibration analysts just stick with one frequency range, for example 1000 Hz, almost regardless of the machine being analyzed. This can lead to poor resolution when testing low speed machines, and missed key frequencies on higher speed machines.

Many practitioners would recommend a frequency range of 50 orders of running speed for most "normal" machines. For example, if the machine were running at 1750 CPM (30 Hz), the frequency range would be 1500 Hz. A 50 order frequency range would give you a number of harmonics of the bearing tones and blade/vane pass frequencies, and may also give you a second harmonic of a gearmesh frequency.

However, if you also used the standard 800 lines of resolution (and thus 2048 samples in the time waveform), there would be 1.5 * 1500/800 = 2.81 Hz (or 168.75 CPM) between lines in the spectrum (assuming a Hanning window), i.e. the bandwidth would be 2.81 Hz. This makes the exact determination of the frequency a little difficult. In a belt driven machine, or a machine with suspected electrical faults, and in other situations where there are multiple peaks at close frequencies, you may require greater resolution.

Sometimes two measurements are better than one. You need to have higher resolution in the first ten orders of rotation where many of the key frequencies will be found: 1X, 2X, looseness (1X-10X), oil whirl/whip (sub 1X), fundamental bearing tones (sub 1X and 2.5X-8X), and vane/blade pass (4X-10X). And then another measurement of say 50 to 100 orders of rotation rate. This second measurement would be very quick due to the high frequency, and would add valuable information.

Another approach is to collect a single higher frequency, higher resolution measurement from the machine. It would be performed on gearboxes or other components that contain high speed rotating elements. Rather than the normal 800 line measurement, you may use 1600 lines or greater.

Averaging

In an ideal world, the data collector would collect a single time record free of noise from a never changing vibration signal, then produce the FFT and store it. But the vibration is constantly changing slightly and there is noise in the signal. Changes occur as rotating elements go through cycles and there is random noise from inside and outside the machine.

There is a way to minimize the effects of the noise and keep more of the changes due to cycles inside the machine. The process used to correct this is called **Linear Averaging**. It simply computes the average value for each line in the spectrum. This process helps in maintaining repeatability so that two measurements taken 5 minutes apart should be the same.

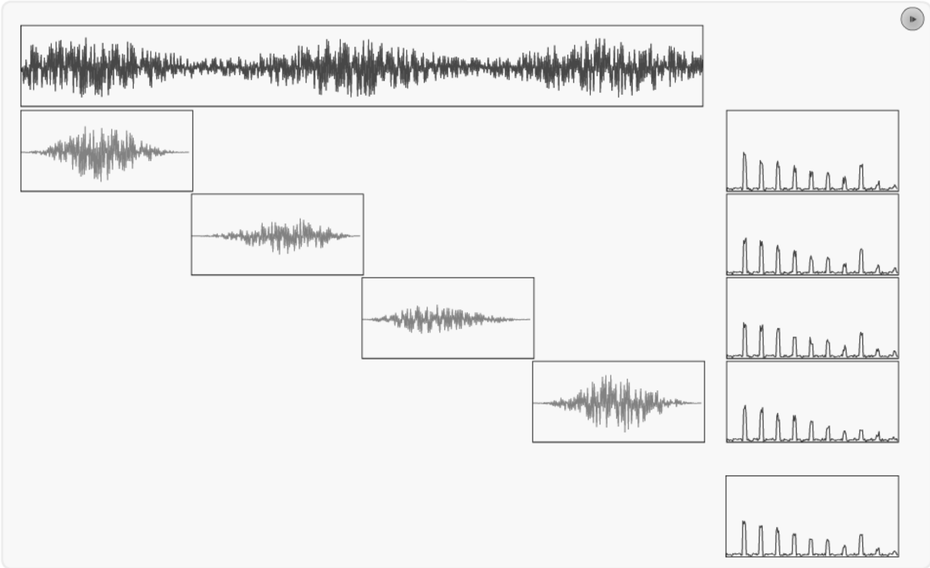


Figure 2-56 Linear Averaging with 4 samples

Linear Averaging collects a time record, produces the FFT and holds the FFT or spectrum. The second time block is collected immediately after the first and produces the FFT and holds it. This is repeated for each average (usually 4). The spectra are then averaged to yield an Averaged Spectrum. It is stored and the rest are dropped from memory.

$$A_3 = \frac{S_1 + S_2 + S_3}{3}$$

Overlap averaging

In order to increase the speed of averaging, and to ensure that we do not lose important data, we use overlap averaging. In this example we have 50% overlap.

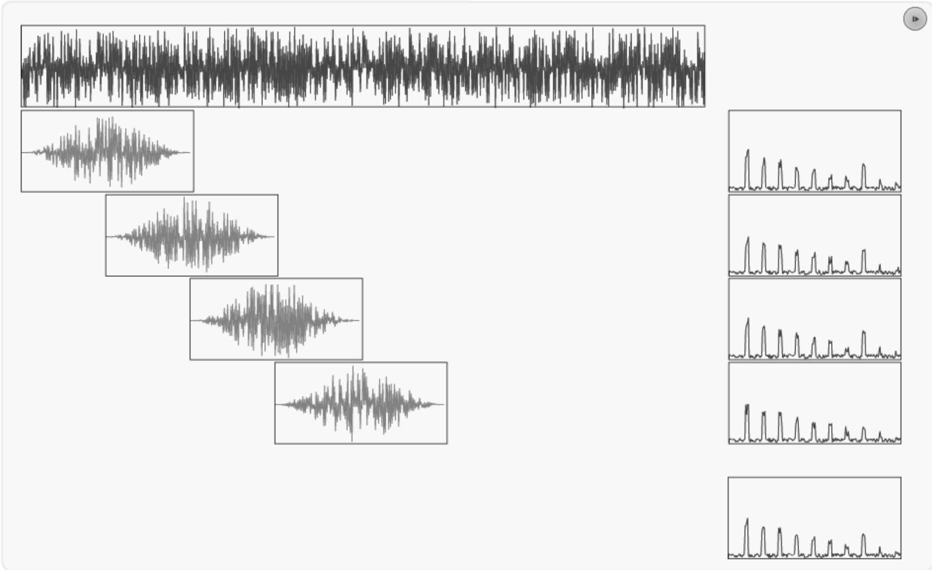


Figure 2-57

When the measurement calls for storing the waveform, most analyzers store either the first or last time record without any window applied. This way the analyst can see the actual time block that produced the spectrum. Some manufacturers collect and store a separate waveform after the spectral data has been averaged and stored.

The time saved is depicted in this graphic:

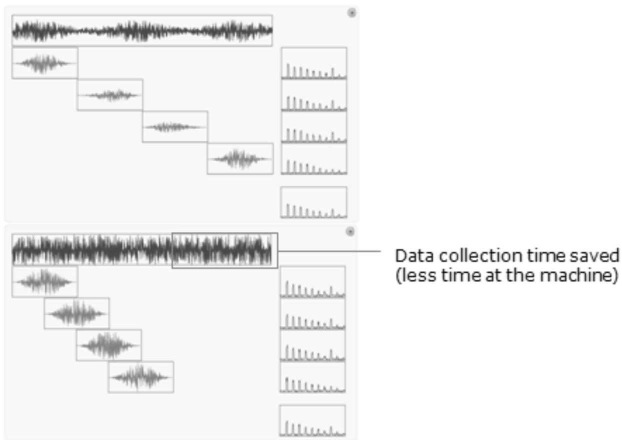


Figure 2-58

Reducing Noise

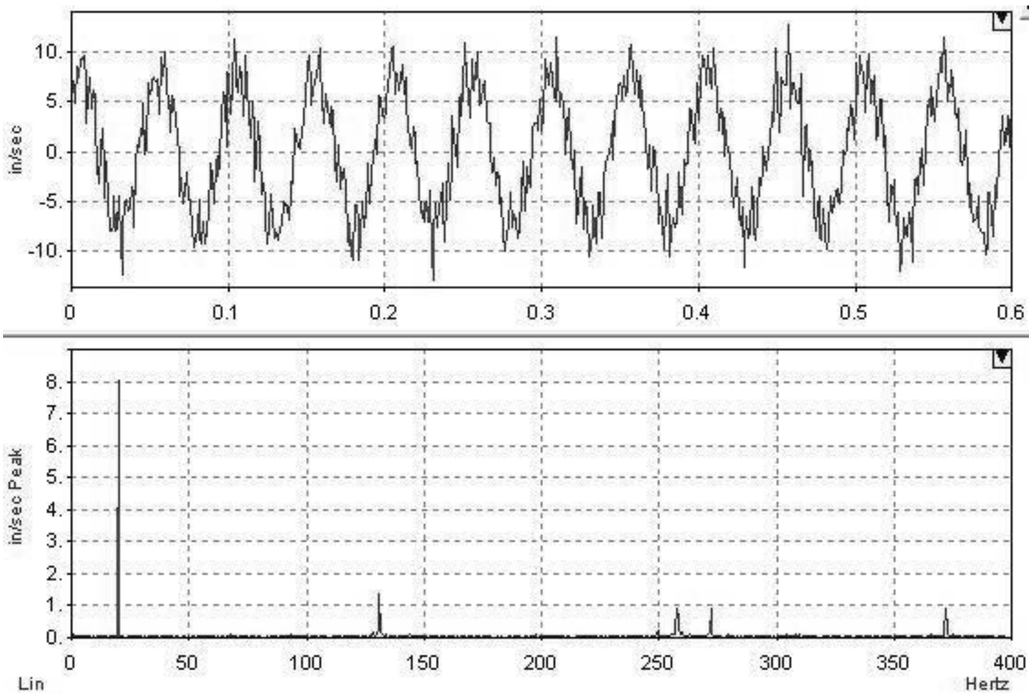


Figure 2-59 One average. Are all the peaks real?

Figure 2-59 shows a waveform and the resulting spectrum. It is one average, or essentially no averaging. Notice the dominant peak. It is likely that it is the frequency of the background sinusoidal waveform.

But what about the other peaks? Are they real? Are they really a result of conditions in the machine such as bearing frequencies?

Additional averages help in resolving that. If it is due to a mechanical or electrical vibration it should be in every block of data. If it is random noise it will be averaged down.

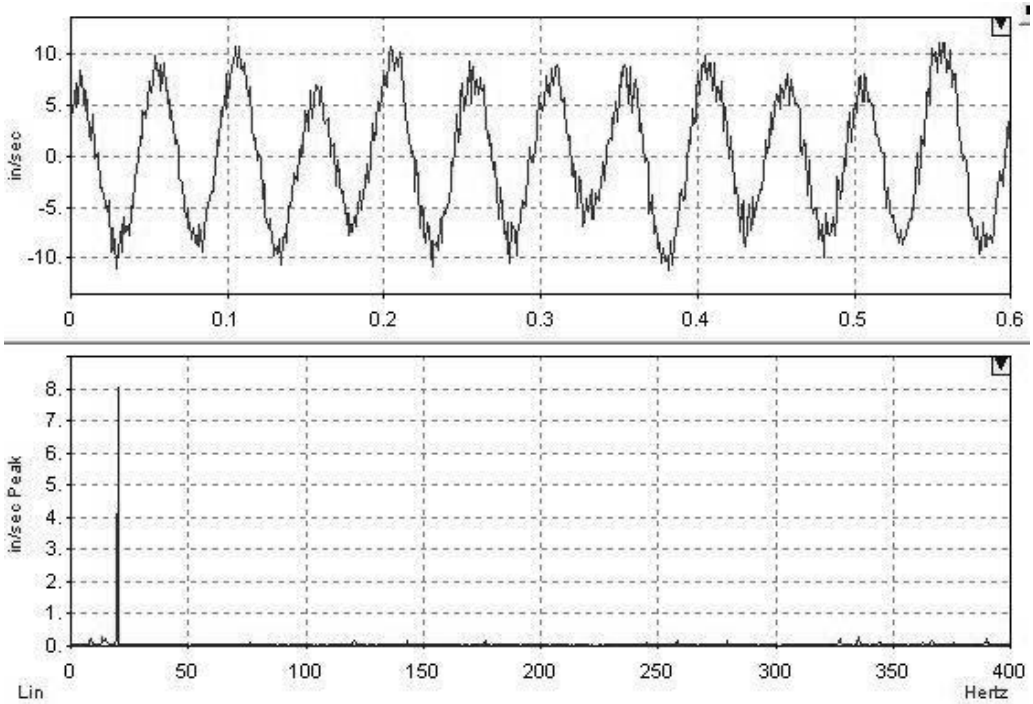


Figure 2-60 Four averages. Most of the peaks are drastically reduced. They are random noise.

The spectrum in Figure 2-60 is the result of 4 averages. Notice how the spectrum has changed. The primary peak is still present but all the other peaks have been reduced significantly. They were random noise and were averaged down to $1/4^{th}$ the amplitude because of the averaging (4 averages). Notice that the noise is NOT removed, it is only reduced.

Averaging methods

There are a number of averaging techniques that can be used. By default data collectors use linear averaging. There are three other techniques used in special occasions.

Exponential averaging

Whereas linear averaging uses all of the history of data equally, exponential weights recent data more heavily than older data.

$$A_1 = Spectrum_1$$

$$A_2 = \frac{3}{4}A_1 + \frac{1}{4}A_2$$

$$A_3 = \frac{3}{4}A_2 + \frac{1}{4}A_3$$

$$A_N = \frac{3}{4} A_{N-1} + \frac{1}{4} A_N$$

Subtraction (negative) averaging

Subtraction averaging is a function which subtracts the power spectrum from the power spectrum after summation (linear) averaging. In this case we allow the analyzer to record a number of averages while the machine is in one mode, and then we collect a second set of averages while the machine is in a different mode. During the second stage the analyzer is actually subtracting the data from the first set. The result is that we see the difference between the two modes.

Negative averaging is very useful when performing bump tests:

- Vibration is recorded and averaged while bumping the machine while it is running.
- Vibration is then recorded while the machine is not being bumped.
- This data is subtracted from the first set, thus leaving only the response to the bump – the resonance.

Peak hold averaging

Peak hold averaging is used when we want to keep the highest amplitudes at each line in the spectrum. Peak hold averaging is used when performing bump tests and when performing run-up or run-down tests (when other more advanced techniques are not available).

Time Synchronous averaging

Thus far we have discussed averaging the spectrum. The linear average does not actually remove the noise, it just improves the statistical accuracy of a noisy spectrum.

Averaging the time data can actually reduce the level of noise, and thus uncover low level signals that may have been obscured by the noise.

Time synchronous averaging is performed with the time waveform rather than with the spectra. It uses a tach signal and averages to zero all frequencies that are not integer multiples of the tach signal. The remaining energy is turning speed and its harmonics.

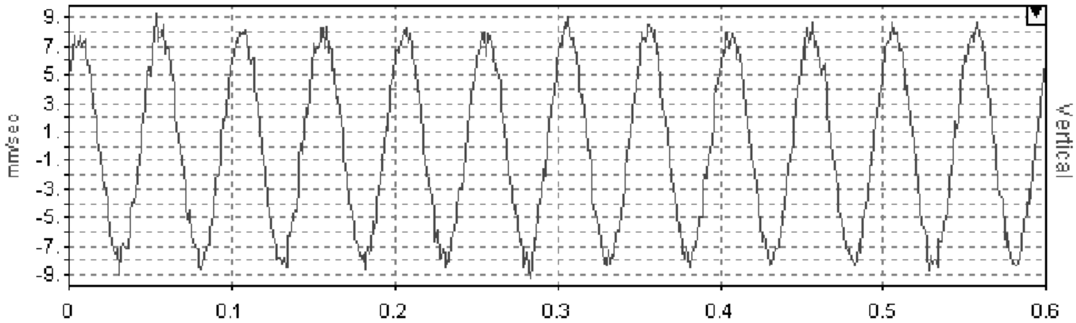


Figure 2-61

If we were to average these time records (Figure 2-62), the data would average away – nothing is synchronized.

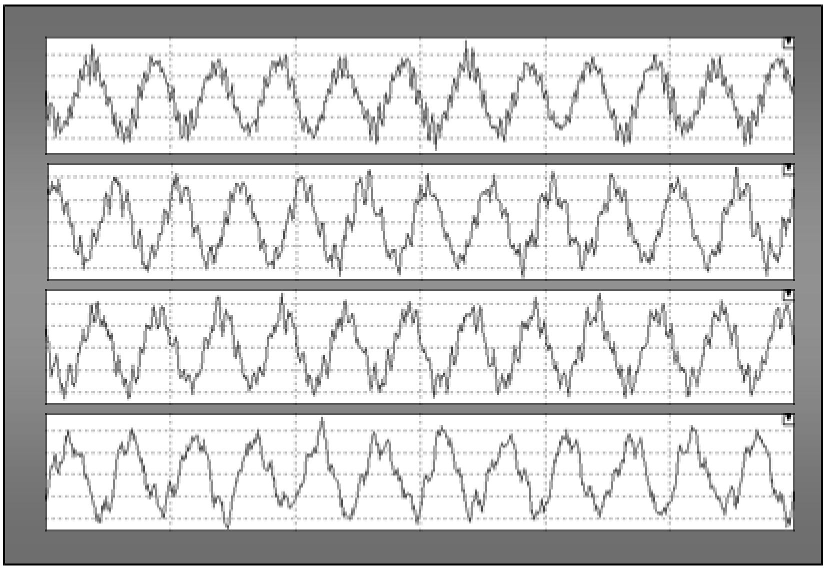


Figure 2-62

If we use a trigger signal to synchronize the beginning of the time records, the data would like as follows:

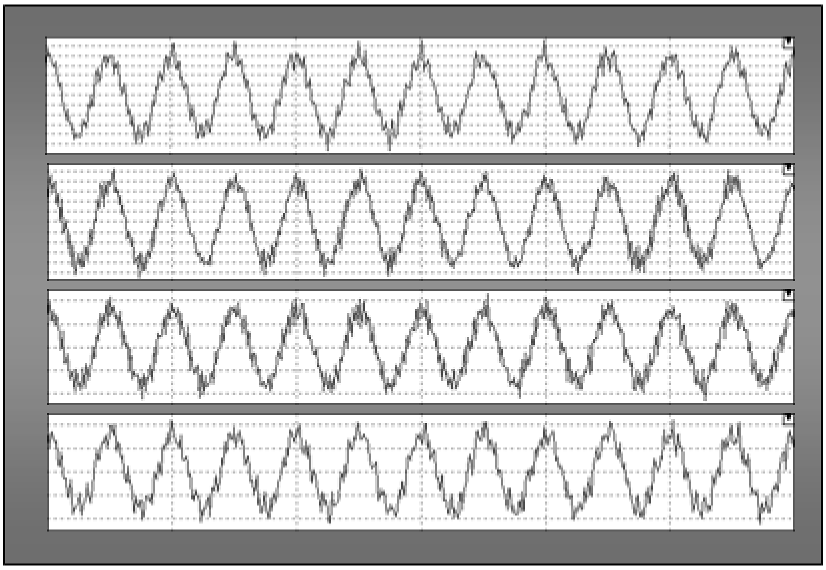


Figure 2-63

The need for the trigger/synchronizing signal (the tachometer signal) means that this method is not often used. In addition, a large number of averages are typically required, so the test can take a long time. However, in certain situations (such as gearbox analysis), the additional effort will pay dividends.

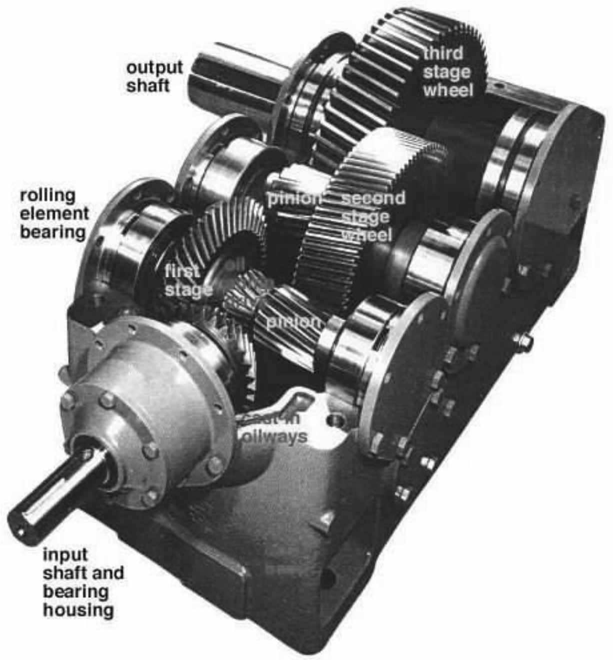


Figure 2-64

Signal patterns: How the FFT behaves

We already know that the FFT will take a time waveform and generate a spectrum. It attempts to find the series of sine waves that, added together, would build the original waveform. Each sine wave would have a different amplitude, frequency and phase. When a waveform is more complicated, the FFT can generate some "interesting results". Although we can try to learn from experience what the various characteristics of a spectrum (harmonics, sidebands and raised noise floor) may mean, this section will provide some "rules" on what the spectrum will look like when various waveforms are processed.

Rule 1: The sine wave

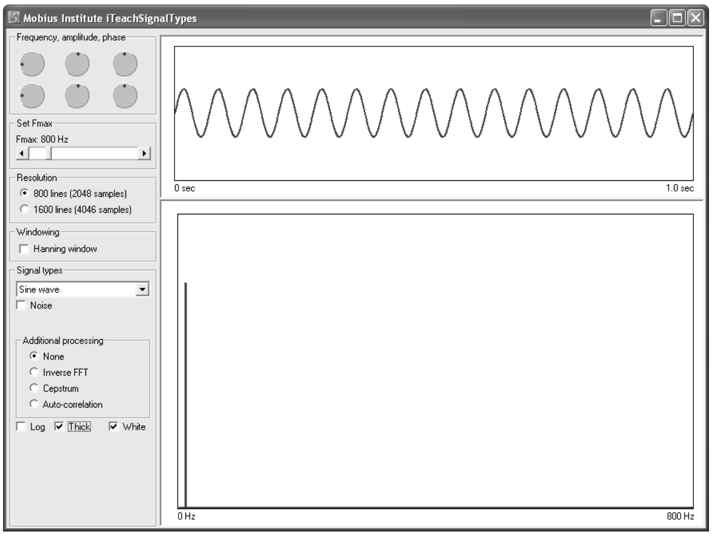


Figure 2-65

Even with a few sine waves added to the original, the spectrum comes out just as you would expect – if they are added together.

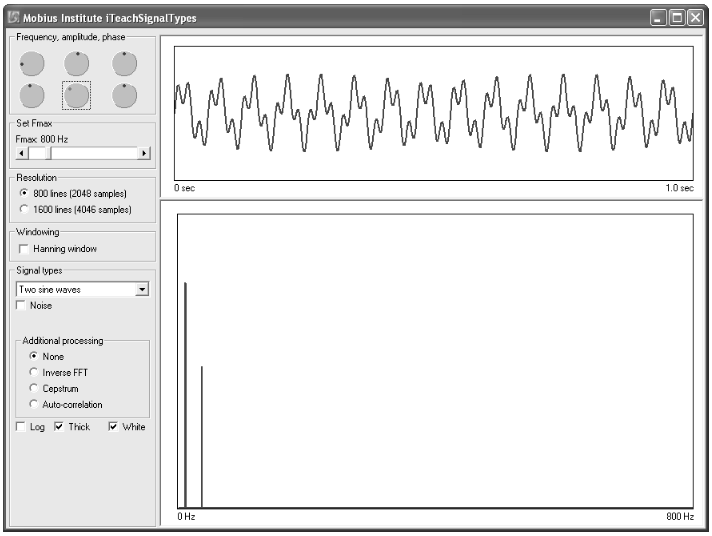


Figure 2-66

Rule 2: Harmonics

If we distort the sine wave we get harmonics. With more distortion comes more harmonics. It is very common to see a small number of harmonics. You should not read too much into small harmonics. If the harmonics are larger, think of reasons for the distortion.

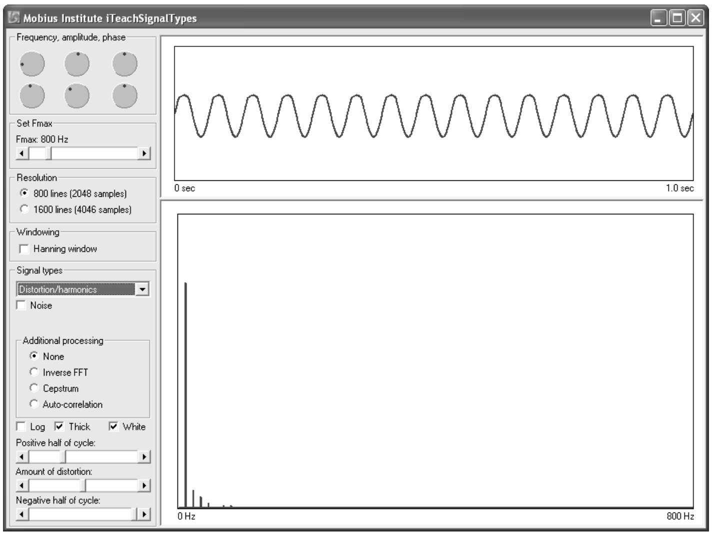


Figure 2-67

Rule 2b: Transients

Transients (with rings) generate very strong harmonics. Impacts in machines (bearings, gears, looseness) all generate strong harmonics.

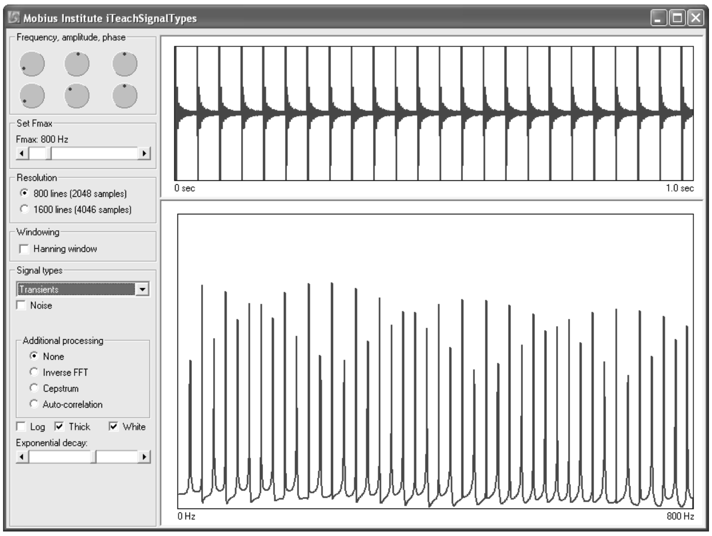


Figure 2-68

Rule 3: Sidebands

Amplitude modulation is the periodic rise and fall of the amplitude of a signal with a different frequency. The higher frequency signal is the carrier, and the lower frequency is the modulator. Amplitude is very common: bearings, gearboxes, motor vibration, etc.

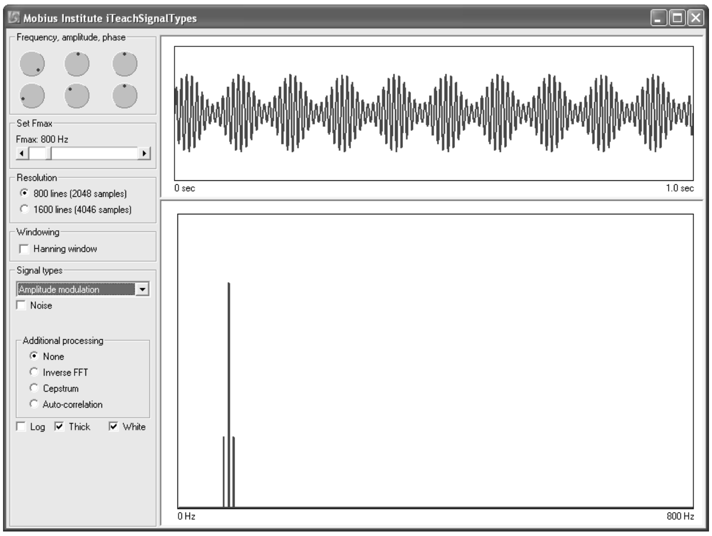


Figure 2-69

Frequency modulation is the periodic change of frequency of a signal. This can occur in bearings and gears when the speed is changed as the torque changes. The result is a stronger series of harmonics.

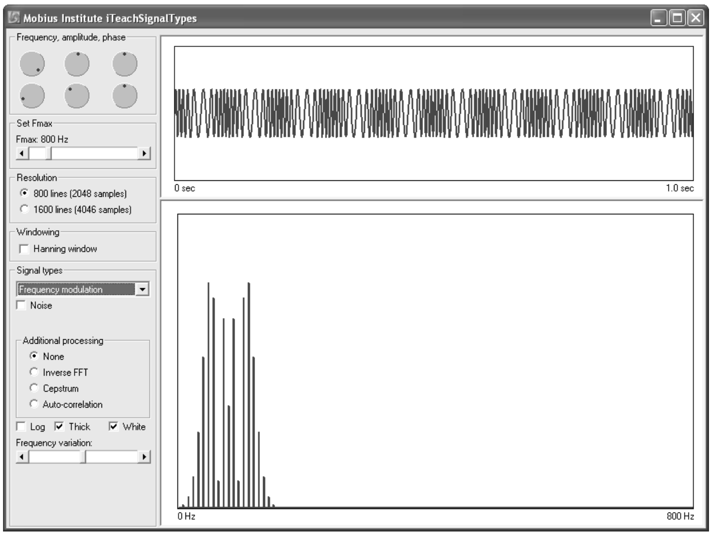


Figure 2-70

Beating

Beating is different to amplitude modulation. Beating occurs when two signals of very similar frequencies interact with each other. As they go into phase their amplitudes add. When they go out of phase they subtract. Therefore we hear a throbbing sound with a frequency which is the difference in frequency between the two signals.

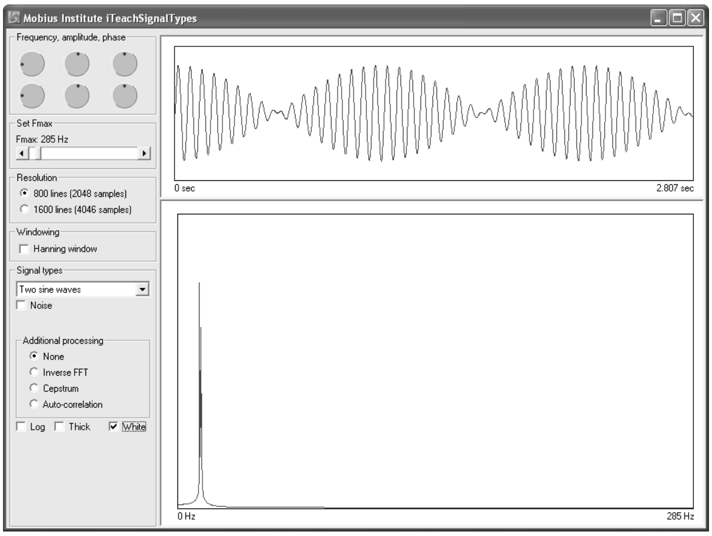


Figure 2-71

Rule 2+3: Harmonics and sidebands

If we distort an amplitude modulated signal, there will be harmonics with sidebands.

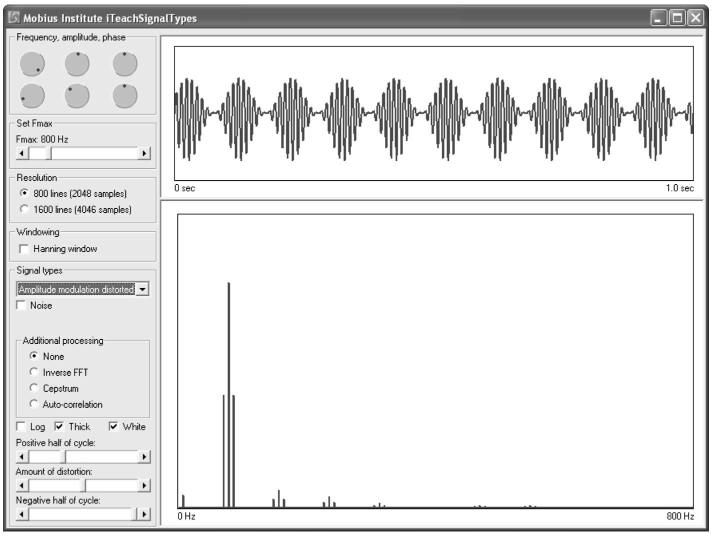


Figure 2-72

Rule 4: Single impulse

A single impulse generates a "spectrum" with energy at all frequencies.

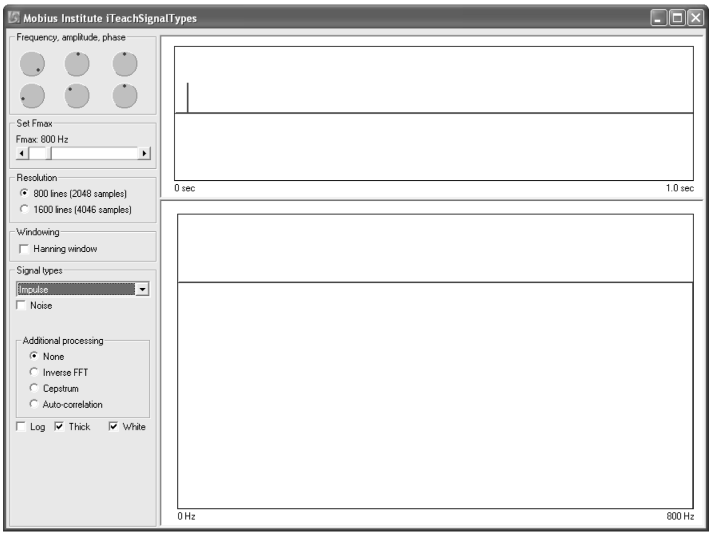


Figure 2-73

A true impulse is difficult to generate in reality, therefore the spectrum will have energy over a range of frequencies. When we impact a structure to excite resonances, we can think of this as generating an impulse.

Rule 5: Pulse

The pulse generates a pattern which is dependent on the width of the pulse.

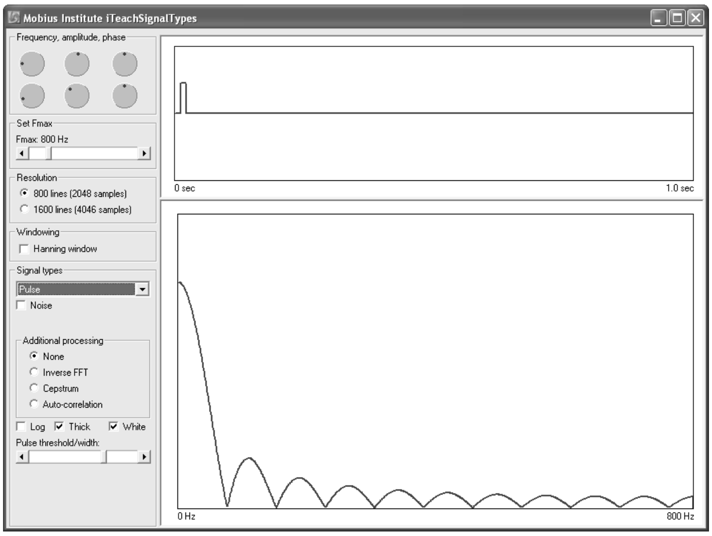


Figure 2-74

Rule 6: Pulse train

The frequency of the harmonics is the inverse of the period between pulses.

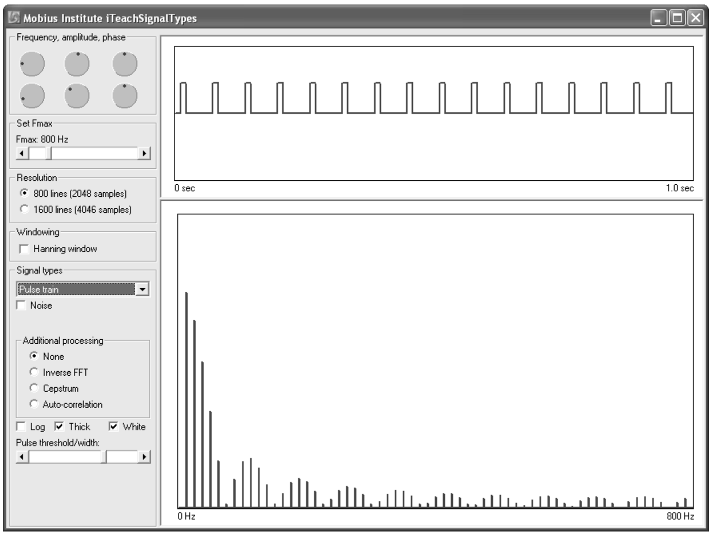


Figure 2-75

Rule 6b: Square wave

A square wave is a modified version of a pulse train. It also generates harmonics, however only odd harmonics exist in the spectrum.

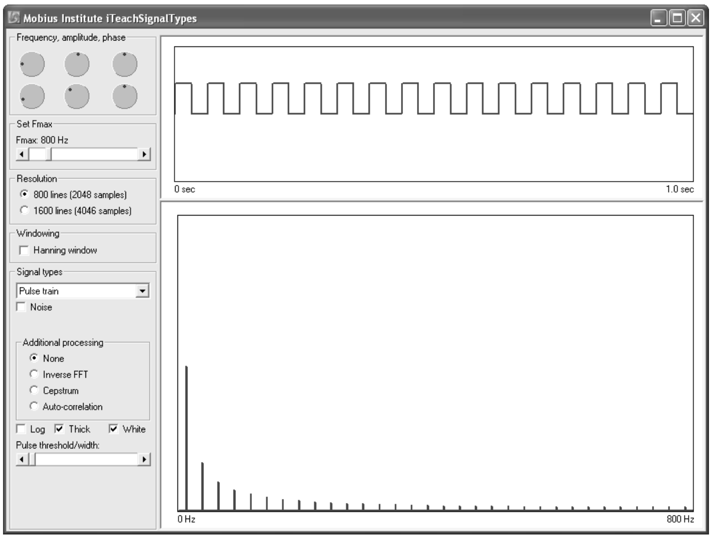


Figure 2-76

Dealing with complex vibration

We have seen how well the FFT deals with sine waves. We have learned about the relationship between the sampling rate, Fmax, and resolution. But now we must deal with a variety of situations that the real world throws at us:

- Noise reduction and extracting the desired signal
    - Averaging, TSA and autocorrelation
- Dealing with variable speed machines
- Dealing with low speed machines
- Using triggers for one-off events
- Recognizing transients, non-linear vibration and modulation

For many years we simply performed the FFT on the time waveform, and then studied the spectrum. In some cases the spectrum was clean and we were able to extract the desired information.

However, there were many other cases where it was necessary to "read between the lines" in order to really understand what was going on within the machine. We might observe a series of peaks and use their existence as evidence of a particular type of vibration: amplitude modulation, frequency modulation, impacting, and more. There were also cases where the information we needed was buried somewhere in the data. Perhaps the amplitude of the vibration of interest was too small in comparison to the rest of the vibration. Perhaps the vibration looked similar to other patterns. Perhaps the machine varied in speed during data collection.

We will now explore different ways of processing the vibration signal in order to extract the data of interest.

Order tracking

There are numerous applications where the speed of the machine may vary during the vibration test:

Run-up tests
Coast down tests
Variable speed machines

- Mining industry (draglines, etc.)
- Elevators
- Vehicles (cars, trucks, trains, etc.)
- Cranes, etc.

Unless precautions are taken, the vibration spectra will be "smeared". Peaks are "broad" because during the test the frequency changed.

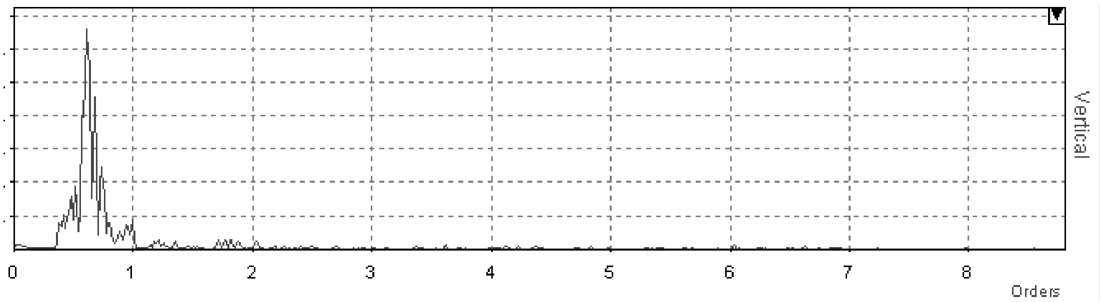


Figure 2-77

When you set the Fmax on the data collector, it determines how quickly it should sample the incoming analog signal according to the following equation:

$$T = T_s \times N = \frac{N}{F_s} = \frac{N}{2.56 \times F_{max}} = \frac{LOR}{F_{max}}$$

The time between each sample is constant.

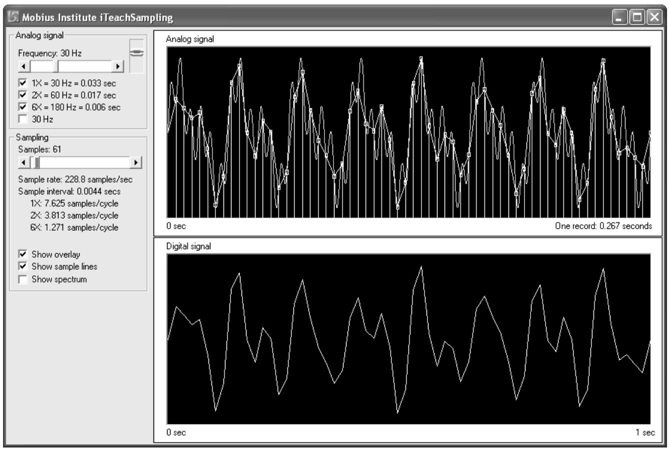


Figure 2-78

Example: Dragline in mining industry

A drag line represents a challenge for a vibration analyst:

- Load/speed varies dramatically.
- Very important machines, so vibration analysis is important.
- Special repeatable test set up to enable vibration monitoring.
- Speed varies during the test.

Figure 2-79

The following is an example of data from a gearbox. The speed and load changed during the test.

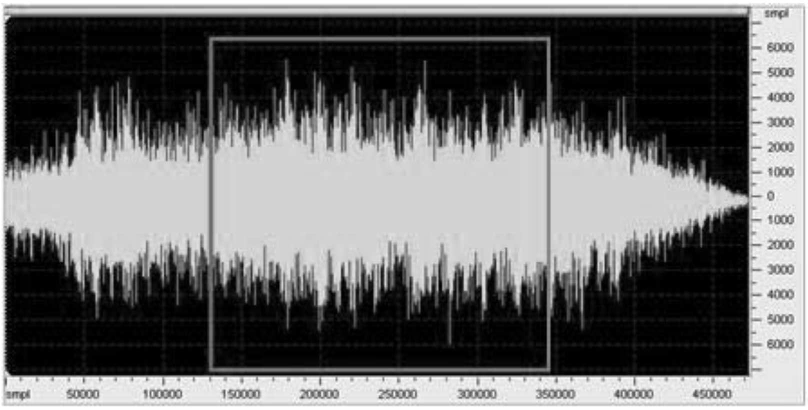


Figure 2-80

There are a number of possible solutions.

Solution 1

Stagger the averaging process

- Perform the test once (run the machine) and collect as many averages as possible
  - There may only be time to collect 1 or 2 averages with the speed and load constant
- Repeat the test and collect additional averages
- It may require two or three tests before you have sufficient data

Solution 2

Normalize the spectra:

- Measure the speed as each spectrum is collected
- Adjust each spectrum to make them all look like they were tested at the same speed
  - First measurement 25 Hz
  - Second spectrum 2% faster
    - Adjust all peaks by 2%
  - Third spectrum 1.5% slower
    - Adjust all peaks by 1.5%
- There will still be smearing in each spectrum
- Resonances (and external vibration) still a problem

Solution 3

Vary the sample rate according to the instantaneous speed of the machine.
Use a tachometer signal to indicate the machine speed.
Collect vibration data simultaneously.
Three alternatives:

- Use a tracking ratio synthesizer to generate pulses used to trigger the collection of each sample.
- Use software to post-process the data.
- Use a shaft encoder.

Variable sample rate

Based on our familiar equation:

$$T = T_s \times N = \frac{N}{F_s} = \frac{N}{2.56 \times F_{max}} = \frac{LOR}{F_{max}}$$

If speed was 25 Hz, Fmax was 1000 Hz and LOR=800
T = 800/1000 = 0.8 seconds   N = 2048
Ts = 0.8 / 2048 = 0.00039 seconds between samples
Speed = 25 Hz   Period = 0.04 seconds
Therefore ~100 samples per cycle (0.04/0.00039)

The following is an example of constant sample rate and a constant speed machine (frequency does not change):

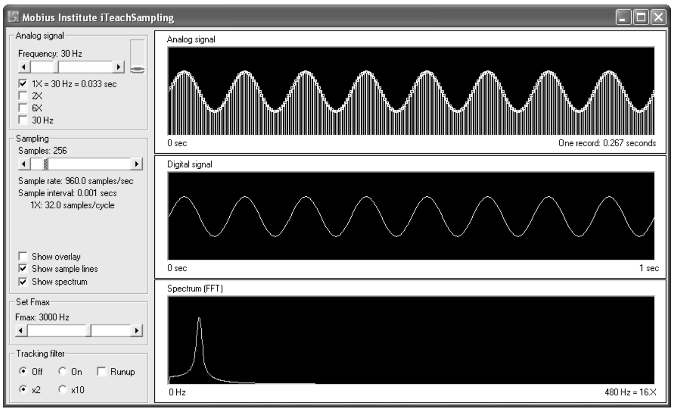


Figure 2-81

Now the speed is changing, but we are again using a constant sample rate. The spectrum is smeared.

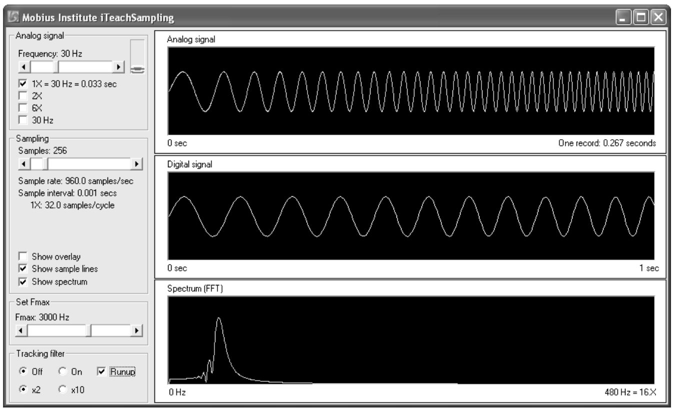


Figure 2-82

By varying the sample rate, the digital waveform "looks" like the machine was not varying in speed. The x-axis is no longer frequency – it is in orders.

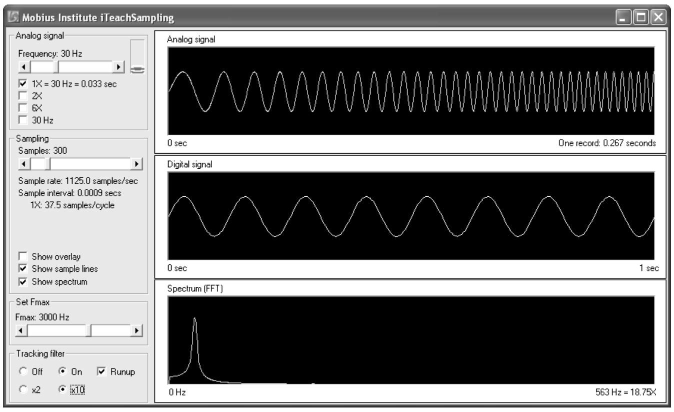


Figure 2-83

Sources of vibration that vary along with the running speed (1X, 2X, bearing frequencies, etc.) will stay in a fixed location on the x-axis. Other sources of vibration will move along the x-axis, for example resonant frequencies, electrical frequencies, etc. You can see the difference in the two spectral maps below.

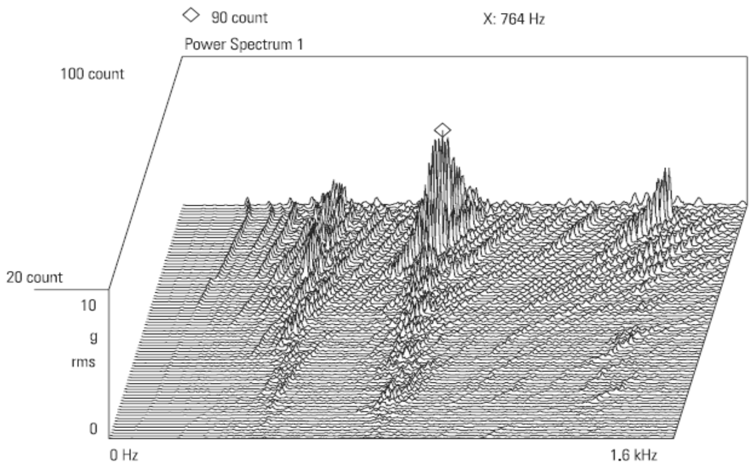


Figure 2-84

The figure above is a spectral map with no order tracking. The plot below is the same data with order tracking.

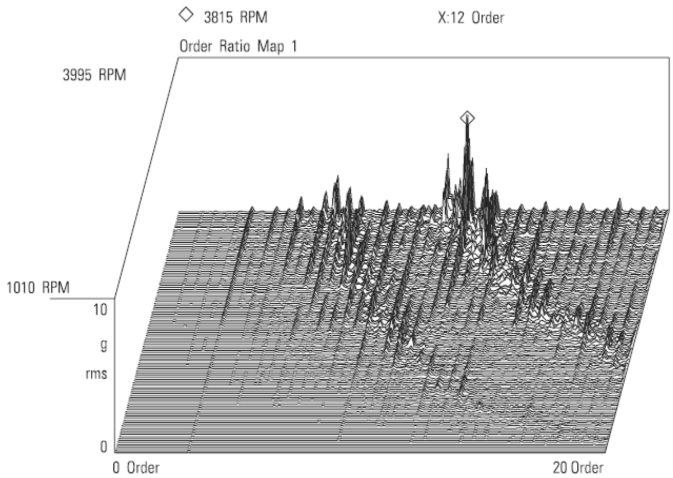


Figure 2-85

Tracking ratio synthesizer

Earlier we showed that if speed = 25 Hz, Fmax = 1000 Hz, and we use 800 lines, then we have 100 samples per cycle. But the tach signal is one pulse per cycle.

As the speed varies, we would like to maintain approximately 100 samples per second.

A tracking ratio synthesizer will take the tach as input and produce a pulse stream at a much higher frequency – e.g. 100 pulses per cycle.

The data collector or spectrum analyzer only collects a sample when it receives a pulse. This is commonly called "order analysis". The tracking ratio synthesizer is an expensive device and you need a data collector or analyzer that supports that type of measurement.

Tracking ratio synthesizers can also produce an output that is at a different ratio of running speed. In gearbox testing, the TRS can produce a pulse stream at the rate of the intermediate shaft, for example.

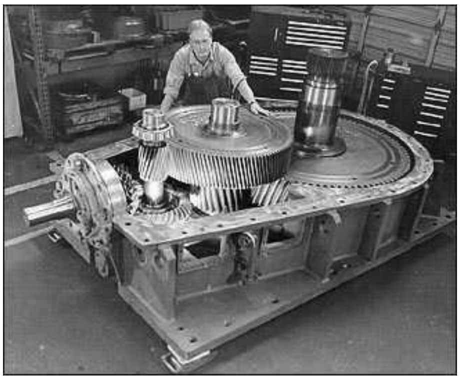


Figure 2-86

Tracking ratio synthesizers often use a "phase-locked-loop" and filters. They have to track the 1X speed as the machine varies in speed. The biggest problem is that they fail when the speed varies quickly.

Machine running up to speed or during coast down.
Dragline swinging the bucket.

The following is the tachometer signal from a machine that is varying in speed.

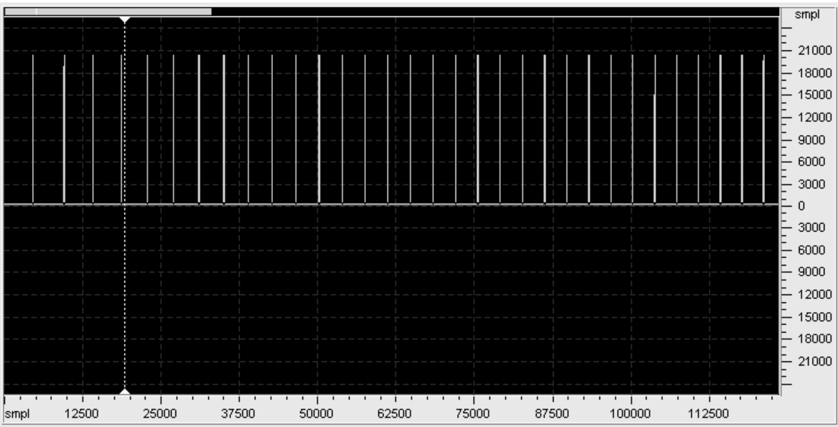


Figure 2-87

Produced by Sumico Technologies

Shaft encoder

A shaft encoder is a device that generates a large number of pulses per revolution. It can be a physical device, like a gear with a displacement probe monitoring the teeth. And it can be an electronic device that turns with the shaft.

For example, if the device generated 360 pulses per revolution, we would have one pulse per degree of rotation. (The data is in the "angle domain".)

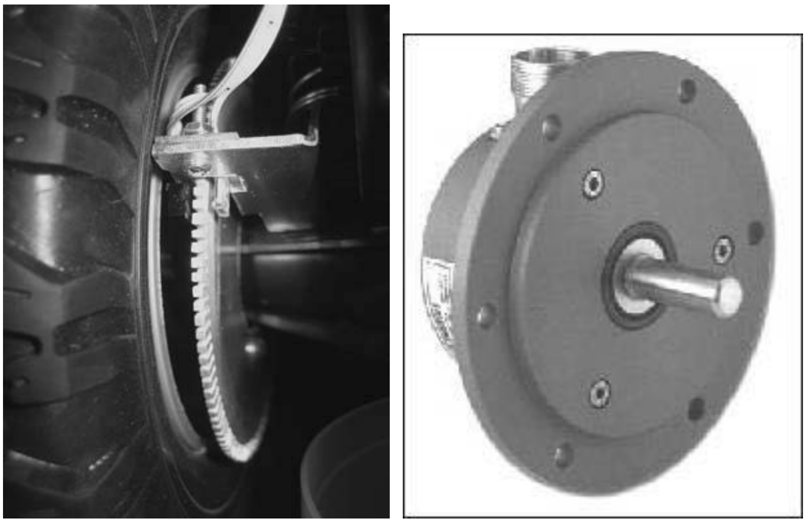


Figure 2-88

The signal is used to control the sample rate of the analyzer/data collector.

Note: unless a separate measurement of speed is made during the test, the data has no speed reference (it is not possible to relate a peak to a frequency in Hz or CPM).

The digital approach

Modern systems take a different approach:

Sample the vibration signal(s) and tachometer signal at a *very* high rate.
Move through the data cycle by cycle:

- Examine the tachometer signal to determine the speed of the machine.
- Digitally "re-sample" the vibration data accordingly.

Modern systems can do this in real time. Many systems perform the analysis after the data has been recorded.

The spectrum is displayed in orders (for example 0 – 10 x running speed). As mentioned earlier, fixed frequencies will "move" across the spectrum:

- Twice line frequency
- Resonances
- External vibration

It is typical to view this data as a waterfall plot or spectral map.

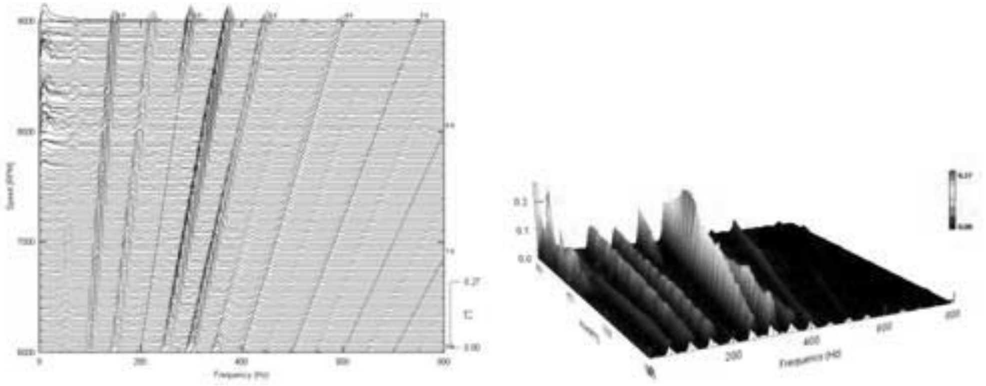
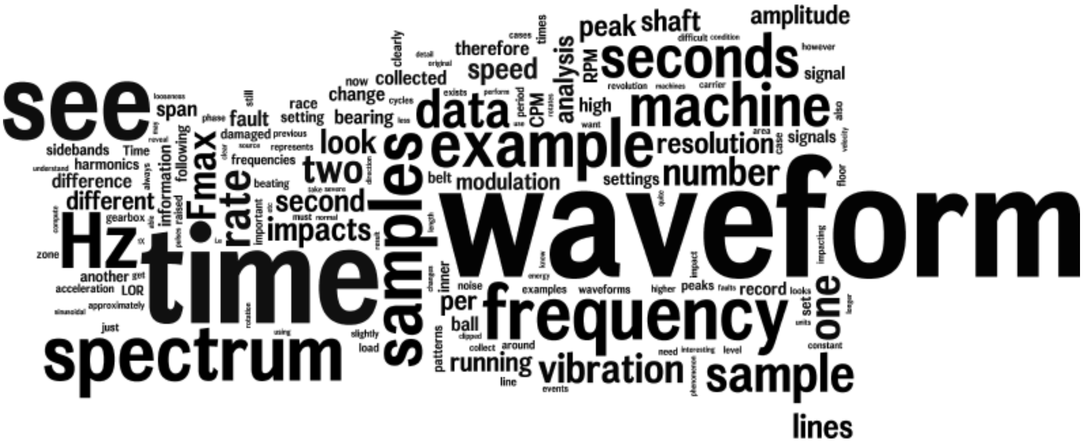


Figure 2-89

The spectral map and waterfall plots show how the vibration changed with time or speed. They remove the smearing issue related to fixed frequency sources of vibration.



Chapter 3

Time Waveform Analysis

Objectives:

- Describe the relationship between resolution, sampling and record length
- Describe optimizing the data collector configuration for time wave form analysis
- Describe waveform patterns such as beating, modulation and transients
- Describe common faults that can be detected using the time waveform such as looseness, belt damage and cavitation
- Describe the application of time waveform analysis to gearboxes

This chapter looks at vibration analysis from the perspective of the time waveform. In Category II students focus primarily on diagnosing machine faults based on patterns in the vibration spectrum. Data collector configurations and data presentation were also focused on optimizing the spectrum. In this chapter we will look at cases where analyzing the time waveform as the primary data type is useful and we will discuss ways to optimize data collection setups for time waveform analysis.

Time waveform analysis

Time waveform analysis is an important analysis tool. Time waveform data should be collected on every route, and it should always be considered when attempting to diagnose machine faults.

Time waveform analysis is a very important skill:

You will better understand the machine vibration
You will diagnose faults that are difficult to diagnose with spectra alone
It is *very* important when analyzing gearboxes and bearings

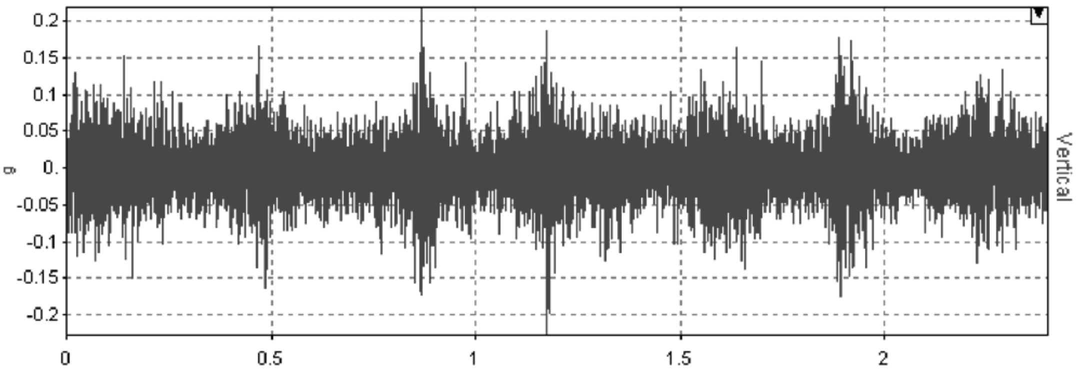


Figure 3-1

Let's get started by looking at two examples. This example is the time waveform from a machine that is out-of-balance. The time waveform has a strong "sinusoidal" pattern, and thus the spectrum as one strong peak in the spectrum corresponding to the running speed of the machine. The time waveform and spectrum are in velocity units.

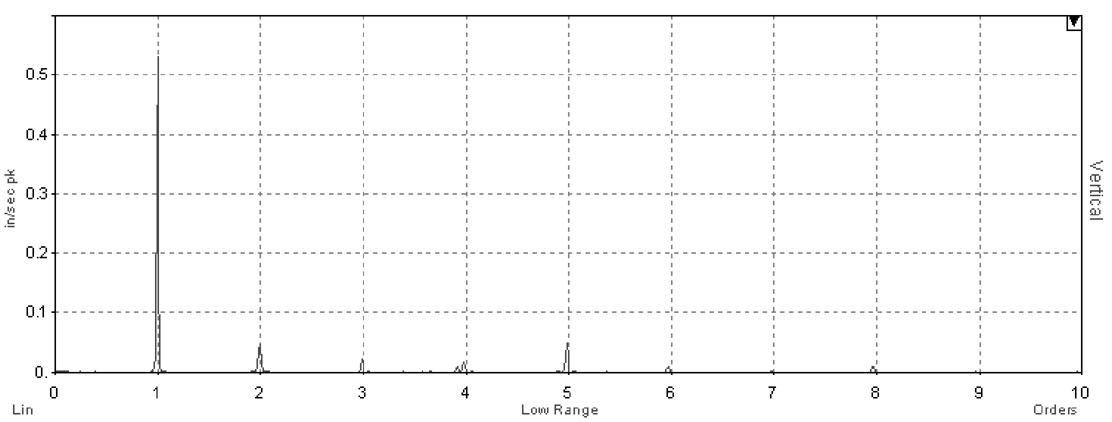


Figure 3-2

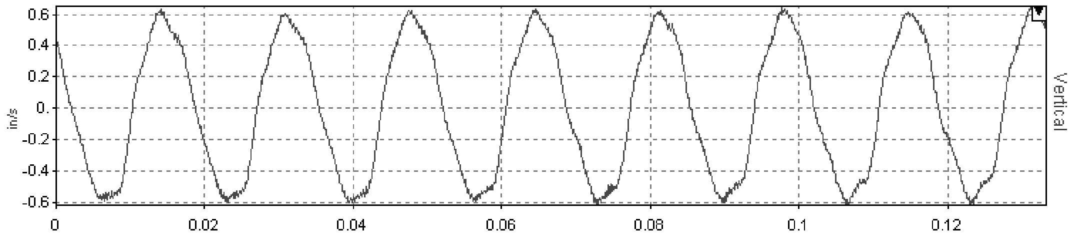


Figure 3-3

The second example is from the drive-end of a motor driving a fan via a belt. The belt is damaged, so once per revolution (of the belt – not the shaft) there is a transient as the damaged area passes through the sheave/pulley. The time waveforms in both of these cases tell us a great deal. But in the second example, the time waveform told us much more than the spectrum.

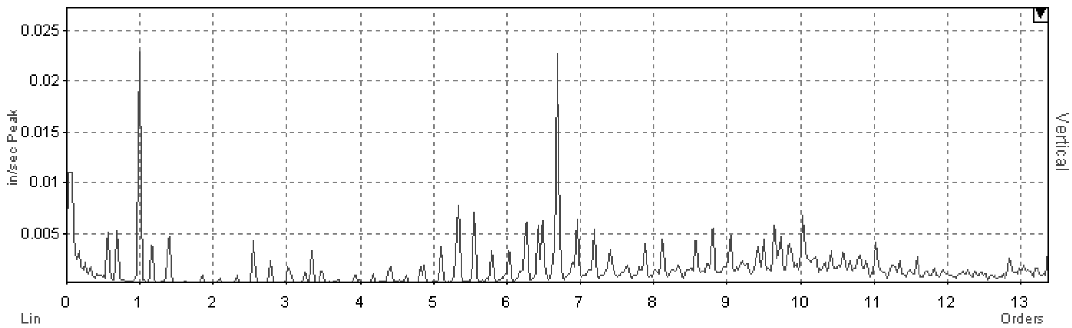


Figure 3-4

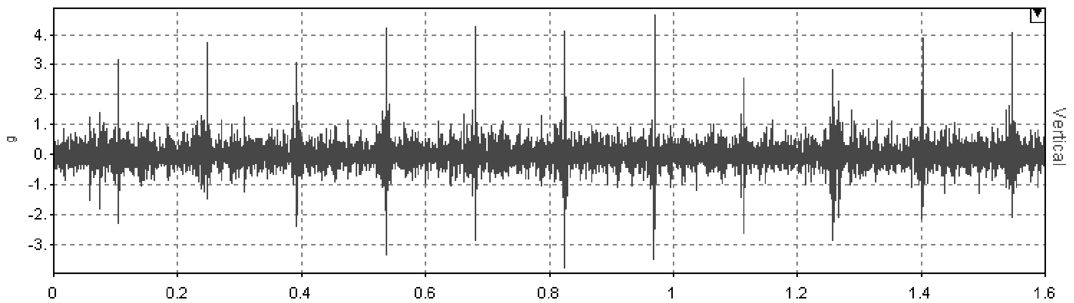


Figure 3-5

Spectrum analysis versus time waveform analysis

In the previous examples we have demonstrated that time waveform analysis can reveal patterns that may be difficult to interpret with a spectrum alone. Here is another example. The spectrum from this cooling water pump does not appear to have high levels.

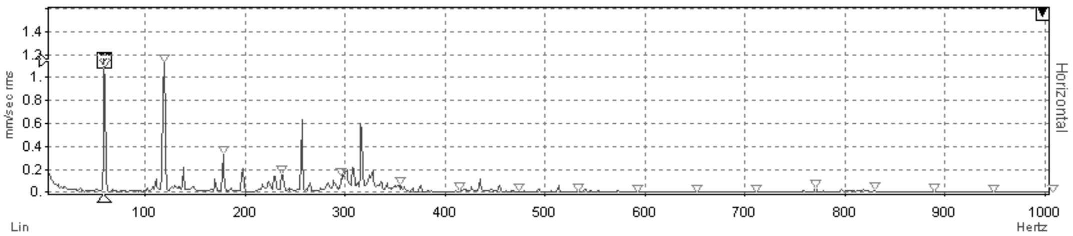


Figure 3-6

When we look at the time waveform, we see that there are very high G levels.

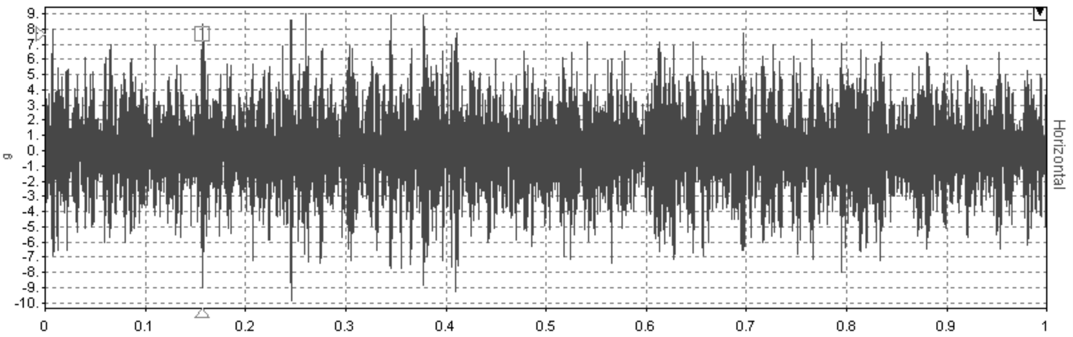


Figure 3-7

When we change the spectrum to units of G's we can see that there is a source of high frequency, high acceleration vibration.

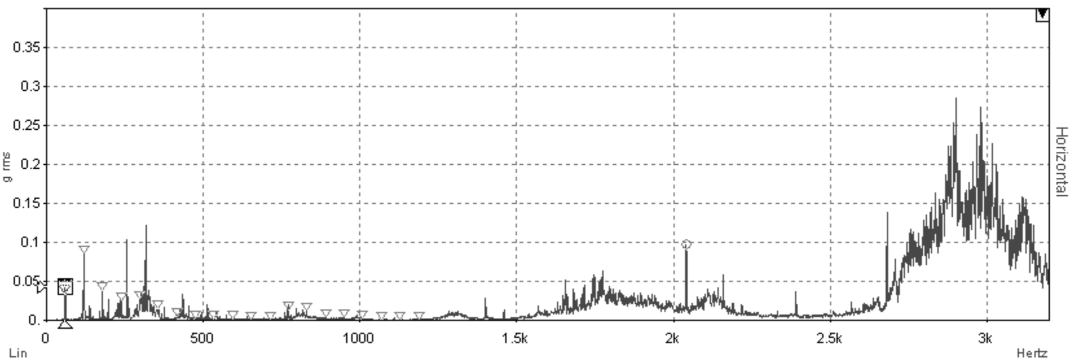


Figure 3-8

Time waveform settings

What we see on the data collector or computer screen is not the original, pure analog waveform that came from the sensor. We have to digitize it so that computers can store and manipulate it.

When you look at a time waveform you can see a limited number of samples (vibration amplitudes used to create the waveform graph). And depending upon how quickly the readings were taken, you may see a fraction of a second of time, or many seconds.

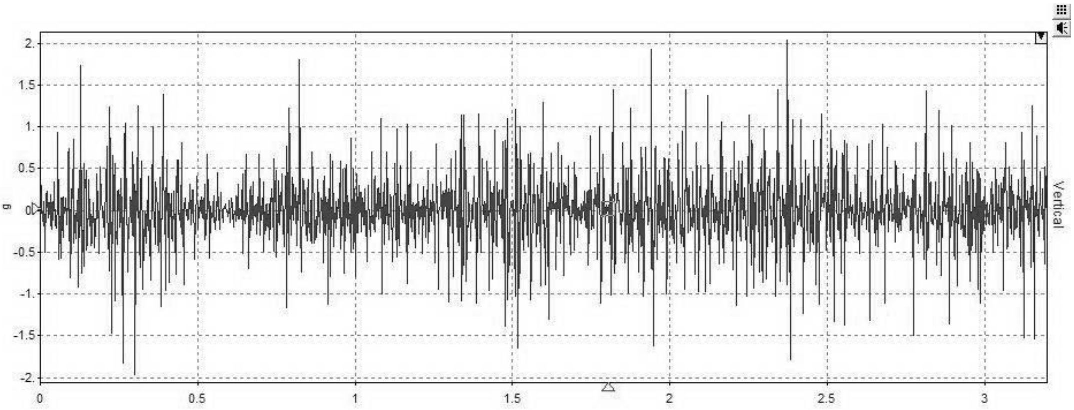


Figure 3-9

Here are the questions you have to answer yourself:

What did the machine do in that time?
How many times did the shaft rotate?
Could we tell if there were any short-duration impacts?
Could we see beating or modulation?

We have three parameters under our control: the sample rate (the number of samples per second), the length of the time record, and whether we integrate the signal (when using an accelerometer).

If we were to sample the waveform a billion times per second, and do so for one minute, we would have a very clear picture of how the waveform changed as the shaft turned and the balls rolled around the raceway, etc., but our data collector would run out of memory. Or we could sample the waveform at just 20 times per second, for only a tenth of a second – but the waveform would reveal no information at all. Obviously, the ideal setting is somewhere in between.

Here we have two time waveforms. This waveform represents 1024 time samples with 0.3 seconds of data. The time waveform does not reveal anything of great interest. It is clear that the machine is not running perfectly smoothly, but it is difficult to learn a lot more from this waveform.

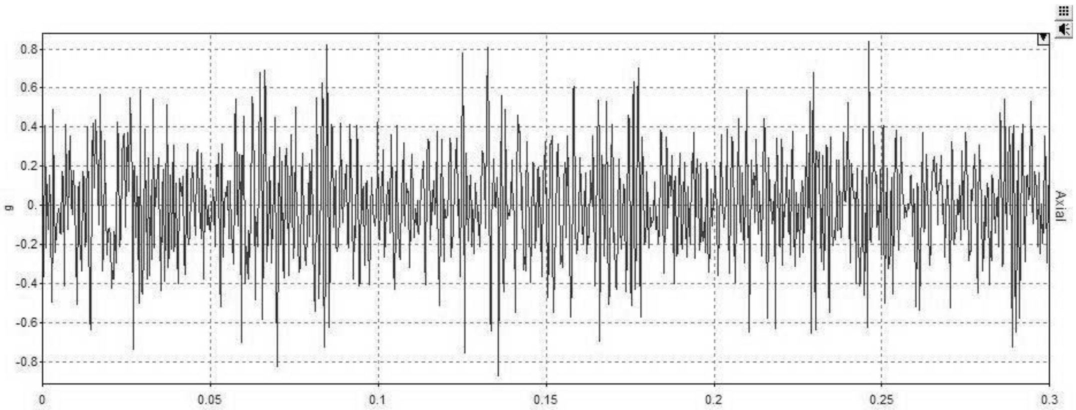


Figure 3-10

The next waveform is from the same machine; however it represents over 4.5 seconds of data. The sample rate (the number of samples per second) was the same, but this time we collected 16384 samples. (Spectrum settings: Fmax = 1,333 Hz or 79,980 CPM and 6400 lines of resolution.) The second waveform allows us to see what happened over a longer period of time. This pump was cavitating (random bursts of energy).

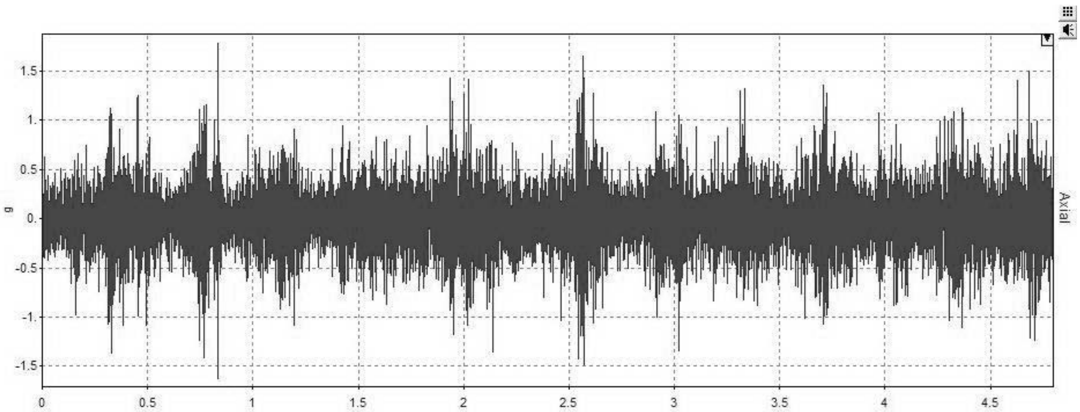


Figure 3-11

So that you get a good feel for the effects of sample rate and the total number of samples collected, we will now go through some more examples.

In this example, we have sampled the data four times more slowly, and collected four times less data than we did in the previous example. As a result, we still have over 4.5 seconds of data, but only 4096 time samples.

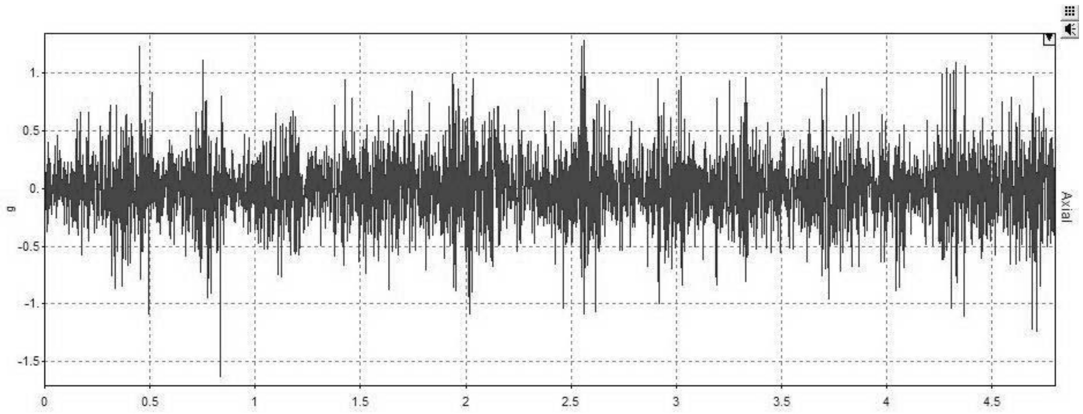


Figure 3-12

You can still see that there are random bursts of energy, however there is clearly less resolution. (The previous example was collected using an Fmax of 333.25 Hz (19,995 CPM) and 1600 lines of resolution.)

In the following example we have again reduced the sample rate by a factor of four, and collected a quarter of the number of samples (1024 samples). Therefore we still have over 4.5 seconds of data, but it is getting much harder to see the random bursts of energy. (This data was collected using an Fmax of 83.3 Hz (4999 CPM) and 400 lines of resolution.)

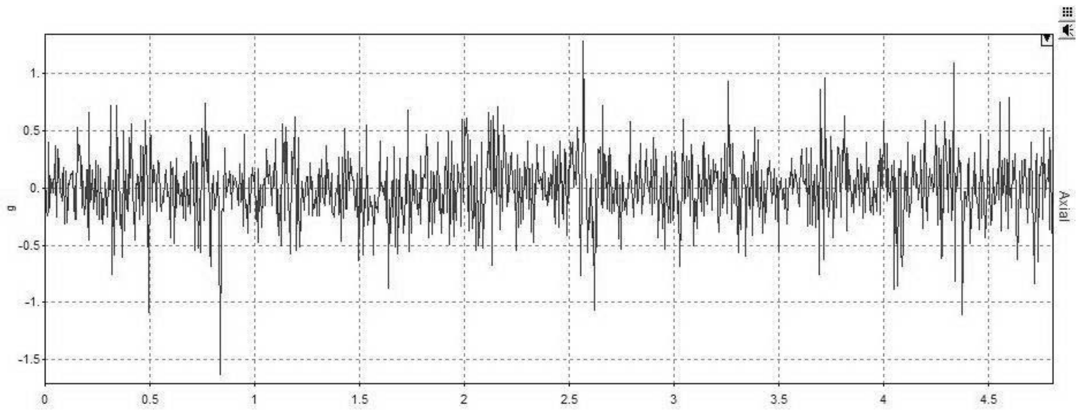


Figure 3-13

So you can see that the quality and usefulness of the waveform is affected by the sample rate, which is controlled by the Fmax setting, and the number of samples in the waveform, which is controlled by the spectral lines of resolution setting (# samples = LOR x 2.56).

We will now look at another example of a time waveform that demonstrates a fault condition in a bearing.

In this waveform, the Fmax was set to 2000 Hz and LOR to 800 lines, therefore we have 2048 samples.

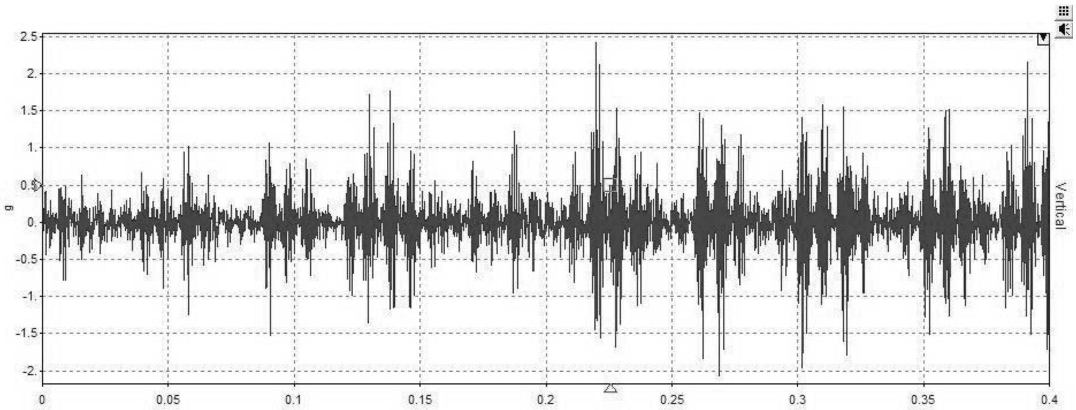


Figure 3-14

You can see the impacts in the bearing. The machine speed was approximately 30 Hz, therefore one rotation takes 0.033 seconds (1/30), so this waveform represents approximately 10 rotations of the shaft.

One interesting point is that while you can see the impacts, you can also see modulation.

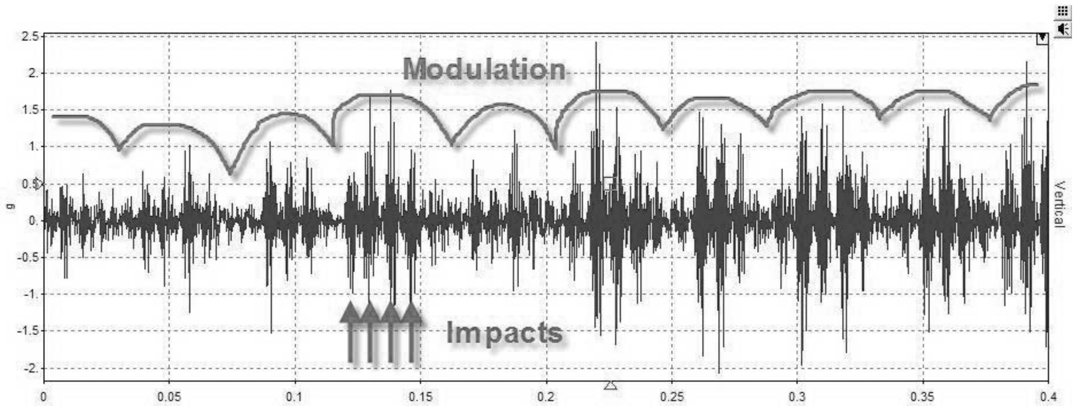


Figure 3-15

And secondly, here is another recording from the same machine with the same settings. You can still see the same impacts and modulations, but you get quite a different picture of severity.

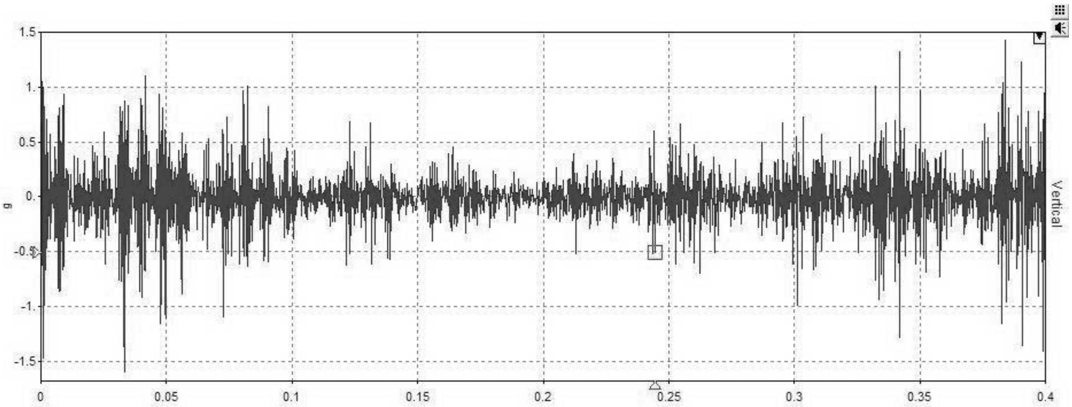


Figure 3-16

And one more example, again we have the same machine, same settings, just a split second later.

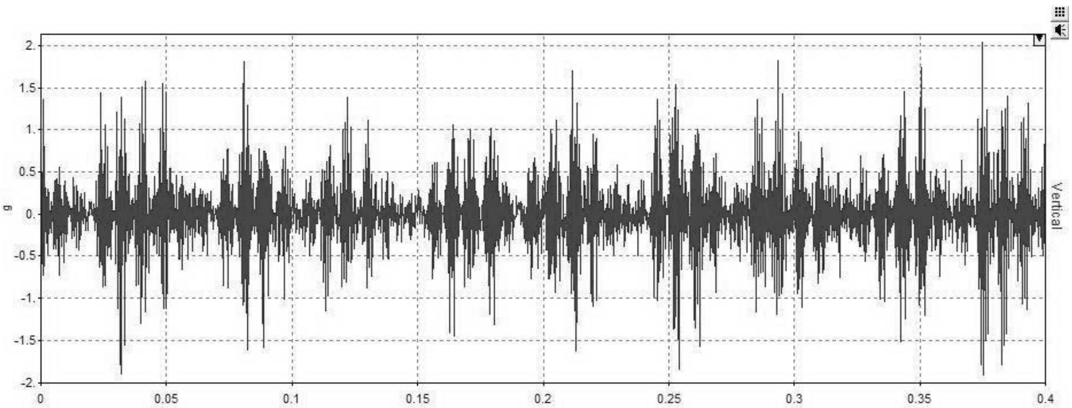


Figure 3-17

Seeing these waveforms you would have to assume that the vibration amplitude is rising and falling 'slowly'; it could be modulation as the cage rotates.

Now we will use different settings. The Fmax is still 2000 Hz, but the LOR has been increased to 1600 lines.

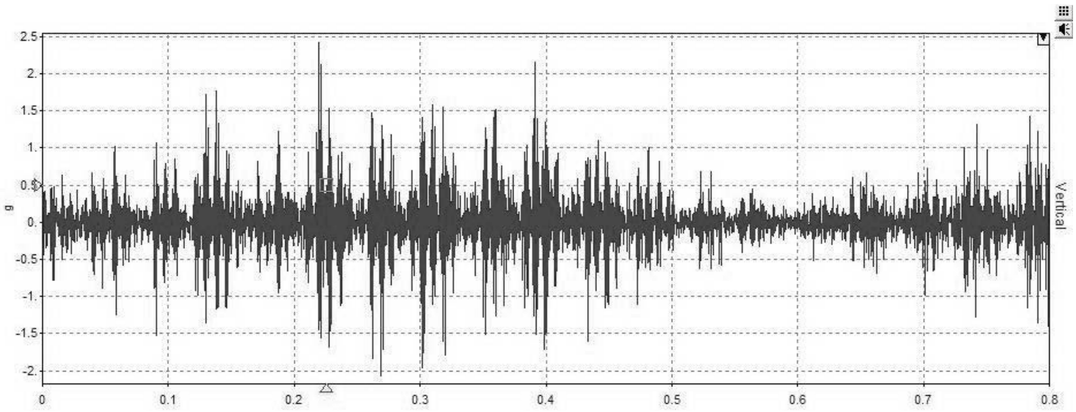


Figure 3-18

You can see 0.8 seconds of data instead of 0.4 seconds as we had previously. The additional source of modulation is becoming more apparent.

In the following example we have the same sample rate (based on Fmax = 2000 Hz) but now we have 3200 lines of resolution. You can see the modulation more clearly now.

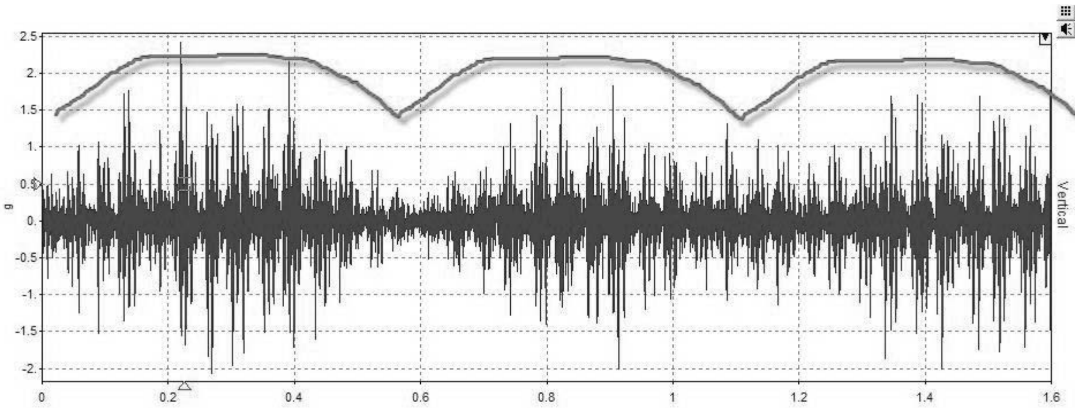


Figure 3-19

Of course, when we keep the sample rate constant, the number of samples per impact or shaft rotation remains constant. When we change the LOR setting, it only changes the total number of samples we collect, and therefore the number of times the shaft rotates. In most software programs you can graphically zoom in on the higher resolution (e.g. 3200 line or 8192 sample) data to see the same detail that we could see in the lower resolution data (e.g. 800 line or 2048 sample).

Now we will change the sample rate. In most analyzers, this is achieved by changing the Fmax setting. For reference, we will repeat the waveform with 800 line and Fmax = 2000 Hz.

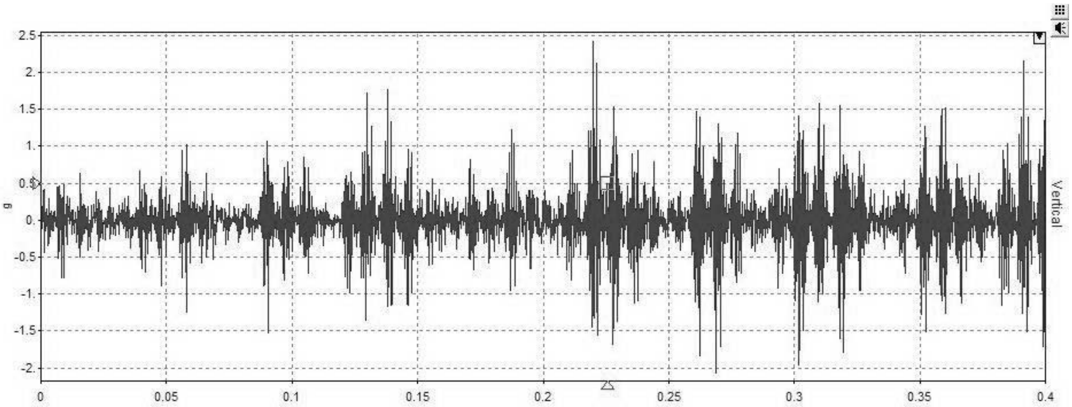


Figure 3-20

In the following time waveform, we have 800 line (2048 samples) and an Fmax of 1000 Hz (half the sample rate). Because we are sampling more slowly, but still collecting the same number of samples, we have 0.8 seconds of data in the time record.

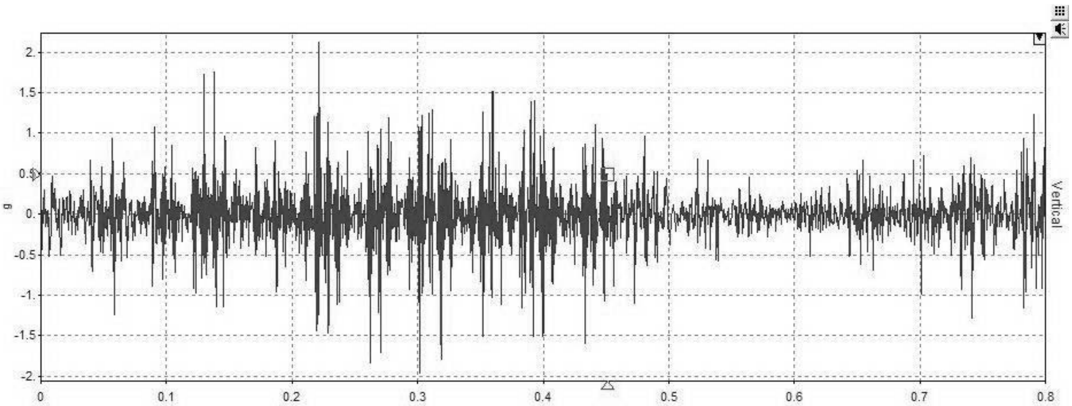


Figure 3-21

If you compare the first of the previous waveform with the original 800 line, Fmax = 2000 Hz measurement, you can see the same impact, but the second measurement does not provide the same level of detail.

Note: The measurements you are viewing are not repeated tests taken on the machine with different settings. Instead they were generated by "re-sampling" the original data – i.e. re-processing the waveform to simulate a different sample rate.

In the following example we have set the Fmax to 250 Hz, and again used 800 lines (2048 samples). We have 3.2 seconds of data, but now it is very difficult to see the impacts or the modulation.

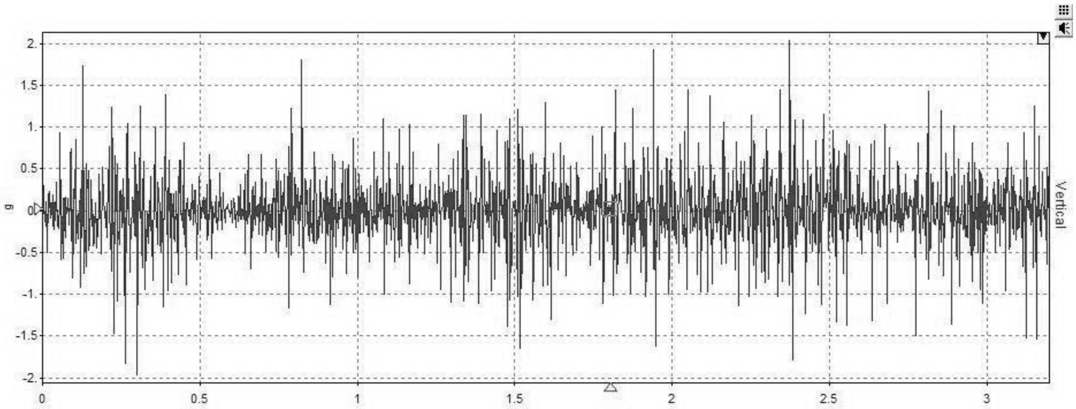


Figure 3-22

Compare the previous waveform to the following waveform that was collected with Fmax set to 2000 Hz (i.e. the original data) and 6400 lines of resolution (i.e. 16384 samples of data in the record).

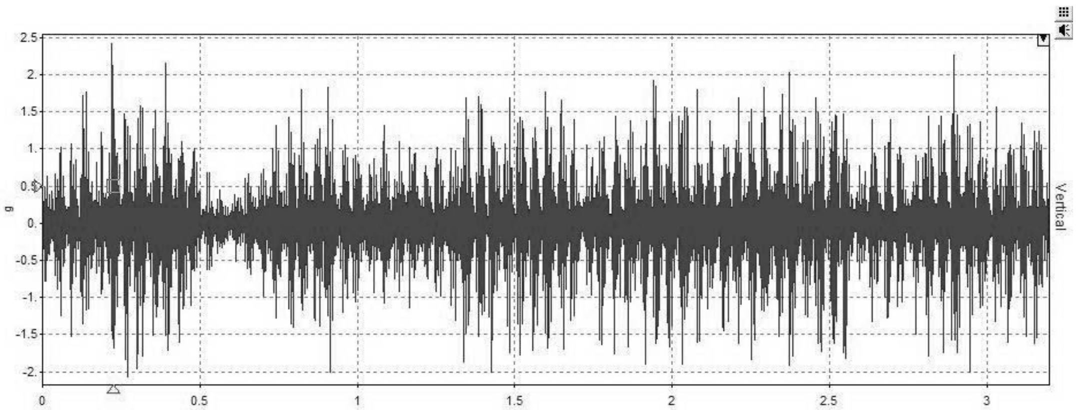


Figure 3-23

There is far more detail, and we can now graphically zoom in any portion of the waveform to better understand the impacts and modulation.

We will take a look at one more example. The following waveform has 2048 samples, representing 1.6 seconds of time. If you look closely you may see that the amplitudes look lower at each end of the waveform; slightly higher in the middle.

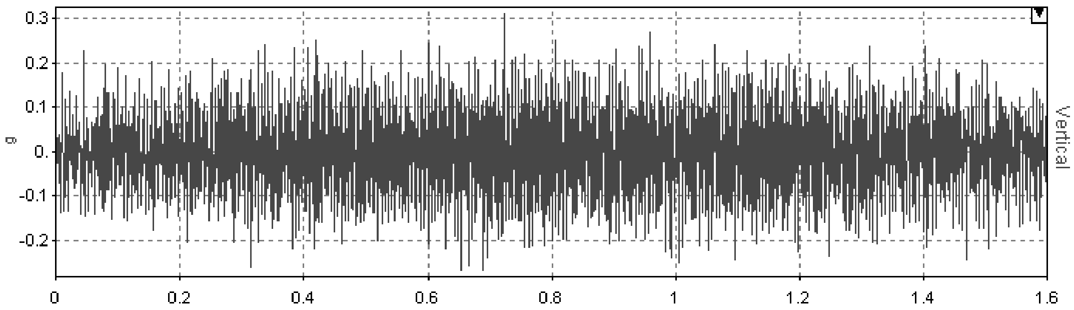


Figure 3-24

This data above represents just part of one of the pulsations in the longer waveform, as shown below (the area from the red box). The following waveform was acquired with the same sample rate (Fmax setting), but it has 16384 samples (LOR = 6400).

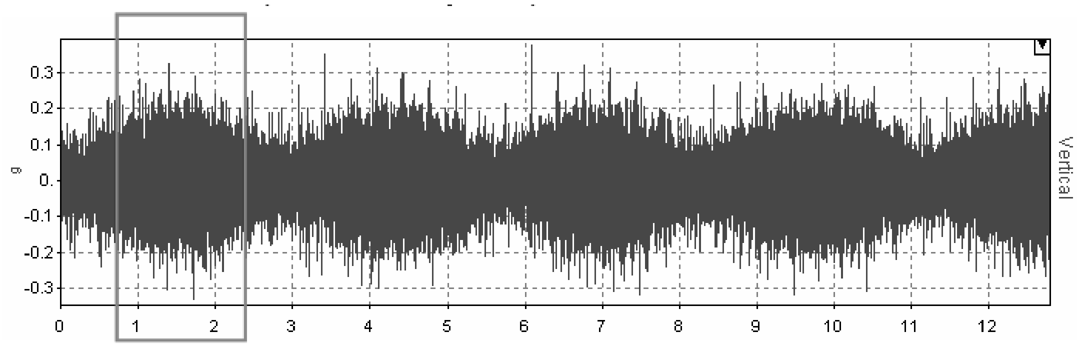


Figure 3-25

You can see that the beating is obvious in the second waveform. However, it is unlikely that you would collect data that requires 12.5 seconds to collect as part of a normal route. You probably would have (hopefully) heard the beating while you were visiting the machine.

All of these examples demonstrate how the number of samples and the sample rate (which dictates the number of seconds covered by the time waveform) can affect what the waveform looks like, even with the same source of vibration. The settings used to collect the time waveform are therefore very important.

Measurement setup

There are ultimately two settings that dictate the format of the time waveform: the number of samples and the time length of the time record.

The length of the time record is determined by the sample rate (samples per second).

Record length = sample per second x number of samples

Some data collectors (and accompanying software) allow you to set the sample rate and number of samples directly. However in most cases you must choose from options normally associated with the spectrum: LOR (lines of resolution) and Fmax.

The Fmax controls the sample rate:

Sample rate = 2.56 x Fmax

And the number of samples in the time waveform is determined by the lines of resolution setting:

Number of samples = 2.56 x Lines of resolution

The number of samples is always a "power of two": 2048, 4096, 8192, etc.

400 lines -> 1024 samples

800 lines -> 2048 samples

1600 lines -> 4096 samples

3200 lines -> 8192 samples

The time span (number of seconds in the time window) is therefore a ratio of these values:

Time span = Lines of resolution / Fmax

But what should the time span be?

We can attempt to quantify this (Catlin, 1987). If we consider a machine with a rotation rate of period T, then we need a time span of T/100 in order to view the details of high frequency events (transients). For example, a machine running at 1800 RPM (1500 RPM) we would want a frequency span of .000333 seconds (0.0004 seconds).

If we wish to see multiple high frequency events in the time waveform, we need a span of T/10. For example, a machine running at 1800 RPM (1500 RPM) we would want a frequency span of .00333 seconds (0.004 seconds).

If we wish to capture one cycle of the shaft, ideal for performing balancing and phase analysis, we would have a time span of T. For example, a machine running at 1800 RPM (1500 RPM) we would want a frequency span of .0333 seconds (0.04 seconds).

To capture multiple cycles of rotation, we would sample at 10T. For example, a machine running at 1800 RPM (1500 RPM) we would want a frequency span of .333 seconds (0.4 seconds). We have a sample of data to illustrate this point. The machine runs at 1776 CPM (29.6 Hz) and the period is therefore 1/29.6 or 0.034 seconds.

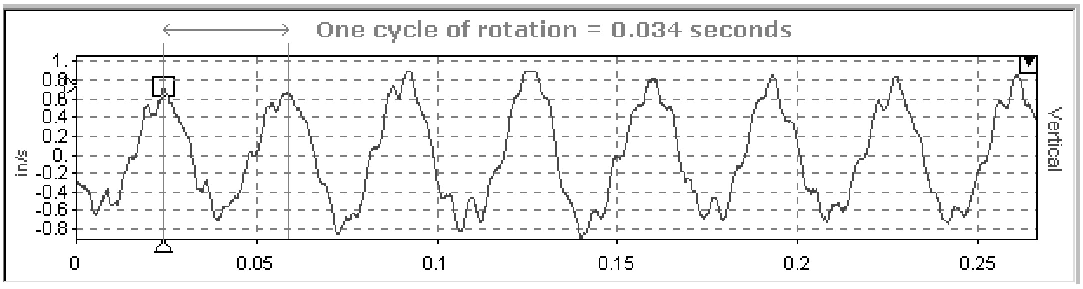


Figure 3-26

Black liquor DIL #1

To capture a larger number of cycles in order to see changes over time – in particular to be able to note phase relationships within the waveform, you need approximately 10T. Although this sample is not a full 10T, it does illustrate some of the patterns you can see when you have a larger number of cycles.

And finally, to view longer term trends and low frequency changes as a result of modulation or beating, you should set up a time span of approximately 100T. Although this sample is not a full 100T, it does show the benefit of having a longer time record.

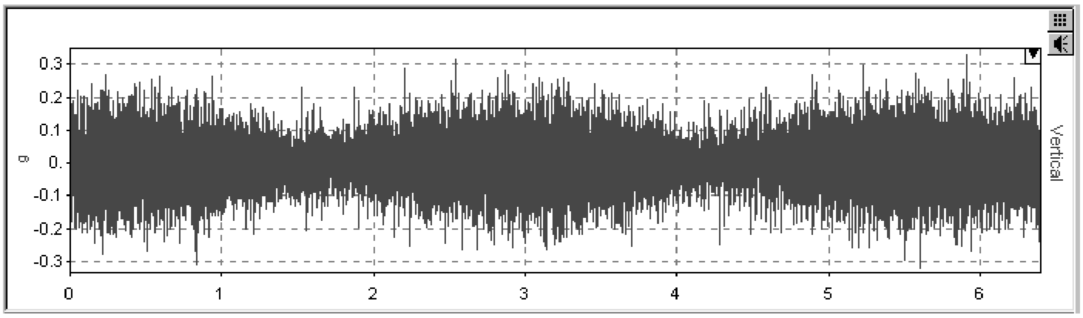


Figure 3-27

Computing the correct settings

Let's use an example of an 1800 RPM machine. 1800 RPM = 30 Hz. That means T is 0.034 seconds (1/30).

If we want 6 rotations, then we need to watch for (6 x 0.034) 0.2 seconds.

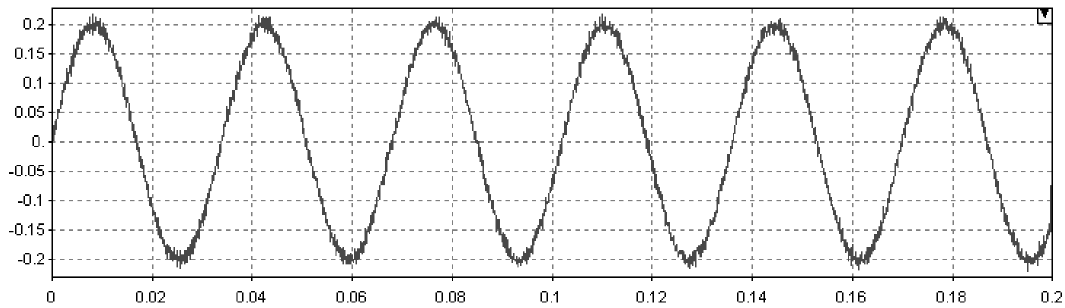


Figure 3-28

If we select 4096 samples in the waveform (recommended) which corresponds to a 1600 line spectrum, we need to sample the machine at (4096/0.2) 20,480 samples per second.

That sample rate corresponds to an Fmax setting of (20480/2.56) 8,000 Hz. So, to collect the desired time waveform, we would set LOR to 1600 lines, and the Fmax to 8 kHz.

There is a tool on our Web site that can help you to perform this calculation in the future:

http://www.mobiusinstitute.com/calculator.aspx

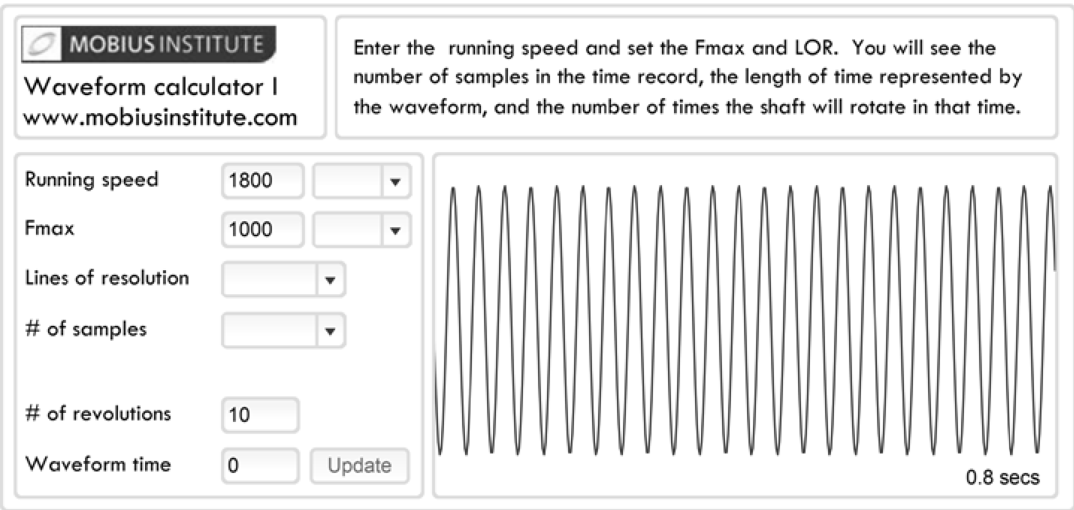


Figure 3-29

If you set the time waveform with a long time record it becomes more and more difficult to see the detail of any impacts (this example is in velocity, not acceleration). In this example, the machine speed was 1755 and the Fmax was set by default to 20 x machine speed (600 Hz) with a resolution of 800 lines (2048 samples). The time record was 1.36 seconds – which represents 40 shaft revolutions.

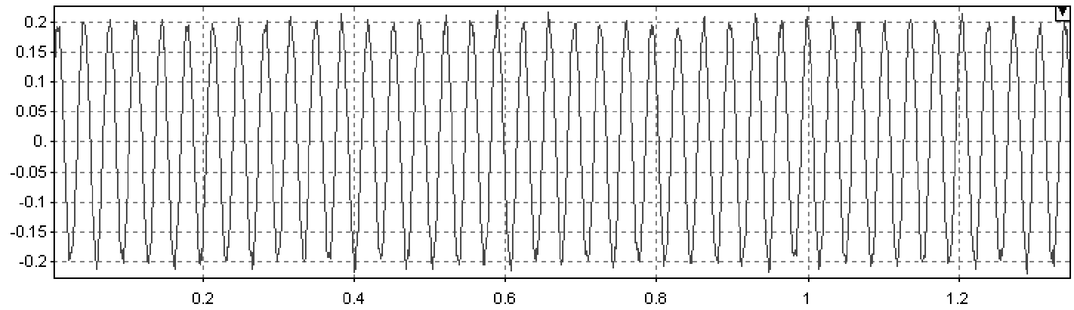


Figure 3-30

Vibration units

The other issue we have to consider is whether to look at the time waveform in units of acceleration, velocity or displacement. A lot could be said on this topic, but in most cases you should use acceleration (e.g. g's). The acceleration waveform will highlight (retain) the important high frequency impacting information. If the signal is "integrated" to velocity, the higher frequency information is reduced or lost, and double integrating to displacement will remove the remaining high frequency information. (The rules are different for journal bearing machines – displacement is the unit of choice.)

These two signals come from the same position and axis of the same machine. You can easily see that the high frequency information is far more obvious in the acceleration waveform (and spectrum).

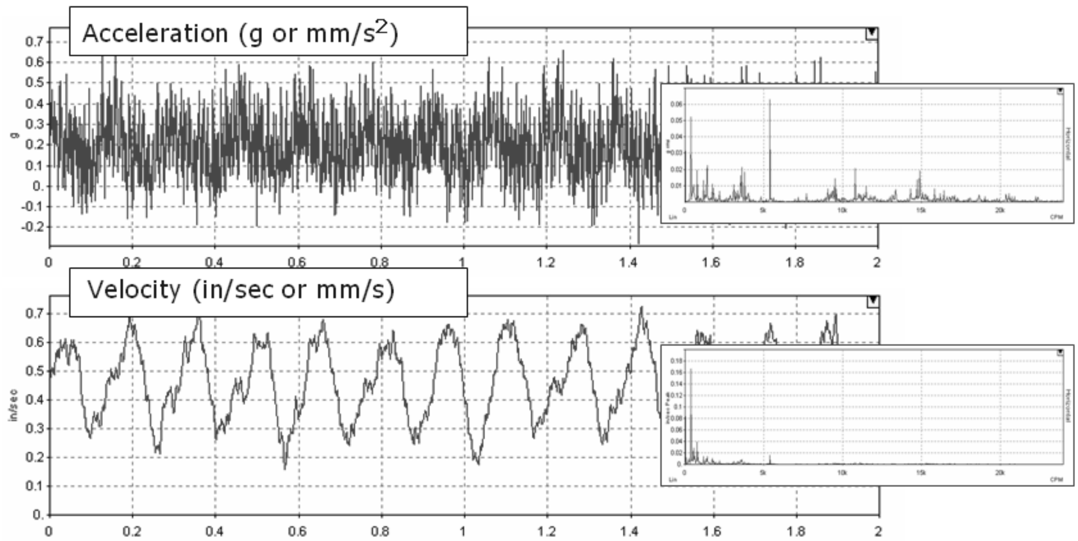


Figure 3-31

Time waveform analysis techniques

When analyzing time waveforms there are a few terms you must be familiar with: the peak, rms, peak-to-peak, period and frequency. These are all described in this slide, and in the vibration fundamentals section. The key issue is that if you measure the time between two "events" in a time waveform, and then compute the inverse (reciprocal), you will compute the frequency. With that information you can look in the spectrum for the corresponding peak, and more importantly, you can compare it with what you know about the machine: speed, bearing forcing frequencies, belt rate, etc.

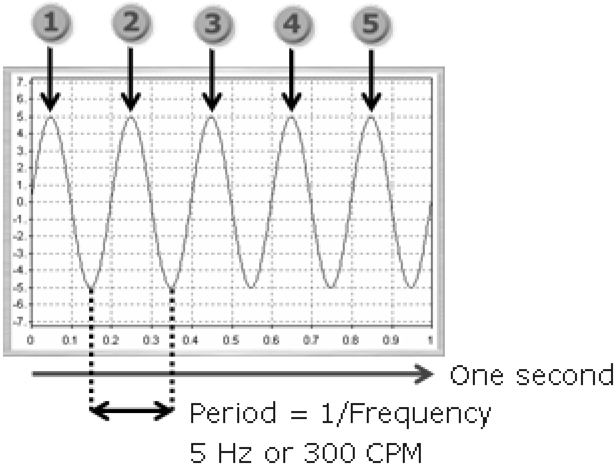


Figure 3-32

Let's take a simple example first. Here we can clearly see a strong sinusoidal signal.

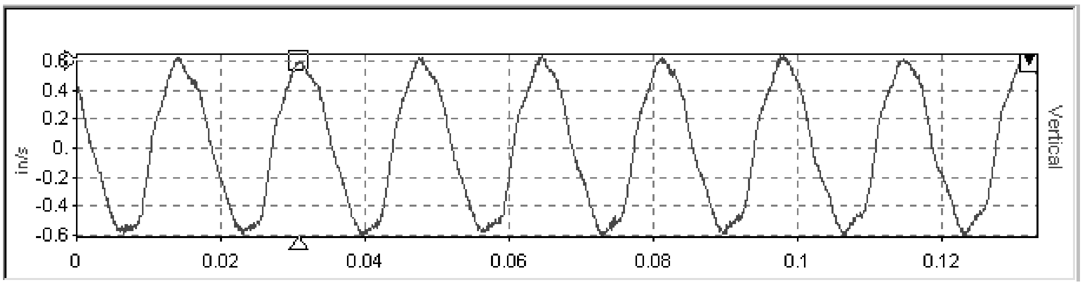


Figure 3-33

If we place the cursor on the time waveform we will get a time relative to the beginning of the record. We don't really care what that time is, however it does allow us to measure the time difference between two events.

Let's take the time at the top of two cycles. We can see that the samples are at 0.0141 seconds and 0.0308 samples. The delta is 0.0167 seconds. We know that the frequency is the inverse of the period, so the frequency must be 1/0.0167 or 59.88 Hz or 3593 CPM.

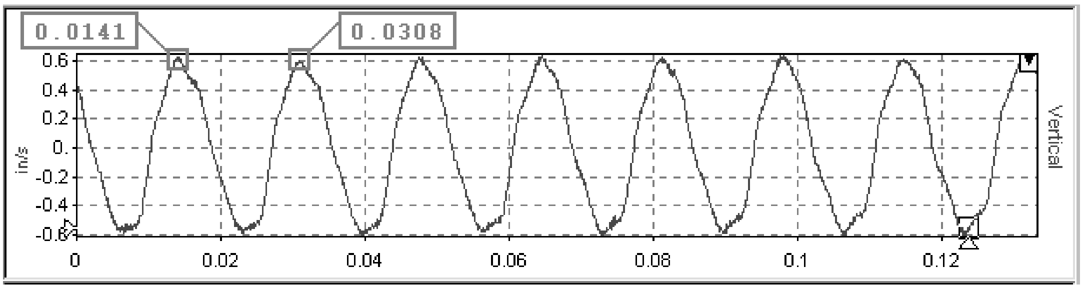


Figure 3-34

The machine in questions rotates at 3593 CPM. If we now look at the spectrum, we can see that there is a clear peak at this frequency, just as we would expect.

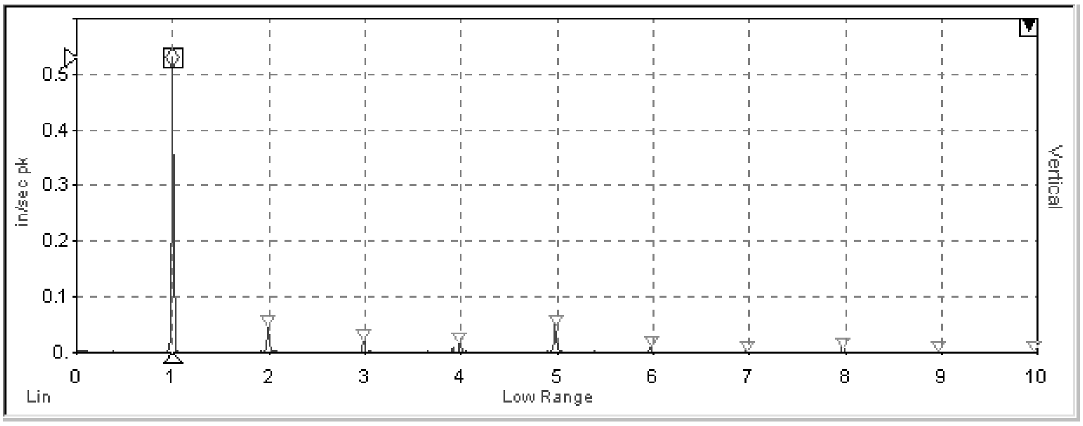


Figure 3-35

Let's take a look at a slightly more complicated waveform. If we revisit a sample we looked at a moment ago, we can see that there are pulses in the time waveform. If we look at the time of any two pulses, and then compute the difference, we therefore have the time between the two events.

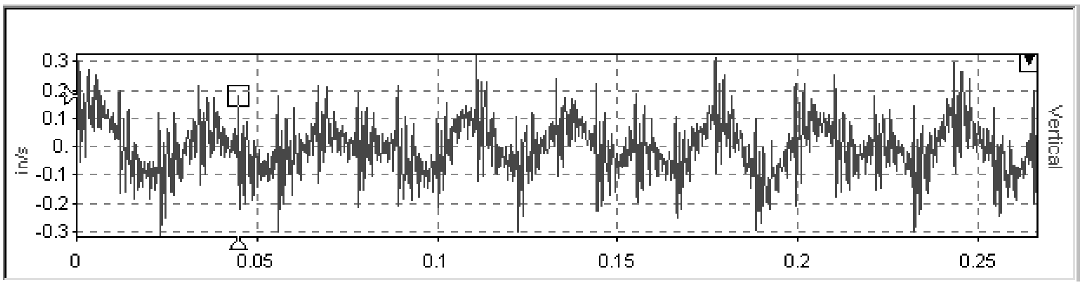


Figure 3-36

In this case the difference is 0.0109 seconds. This would suggest that the period of this signal is 0.0109 seconds. The frequency is therefore 1/0.0109 or 91.74 Hz (5504 CPM). The running speed of this machine is 1776 CPM, so this signal is 3.099X the running speed.

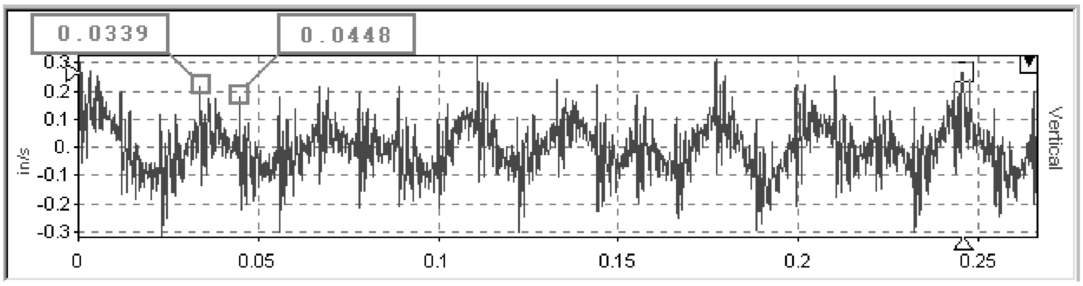


Figure 3-37

In the spectrum we also find a peak at this frequency. In this case we had verification – there was a peak in the spectrum at the same frequency. This is not always the case, and that's why this method is so powerful.

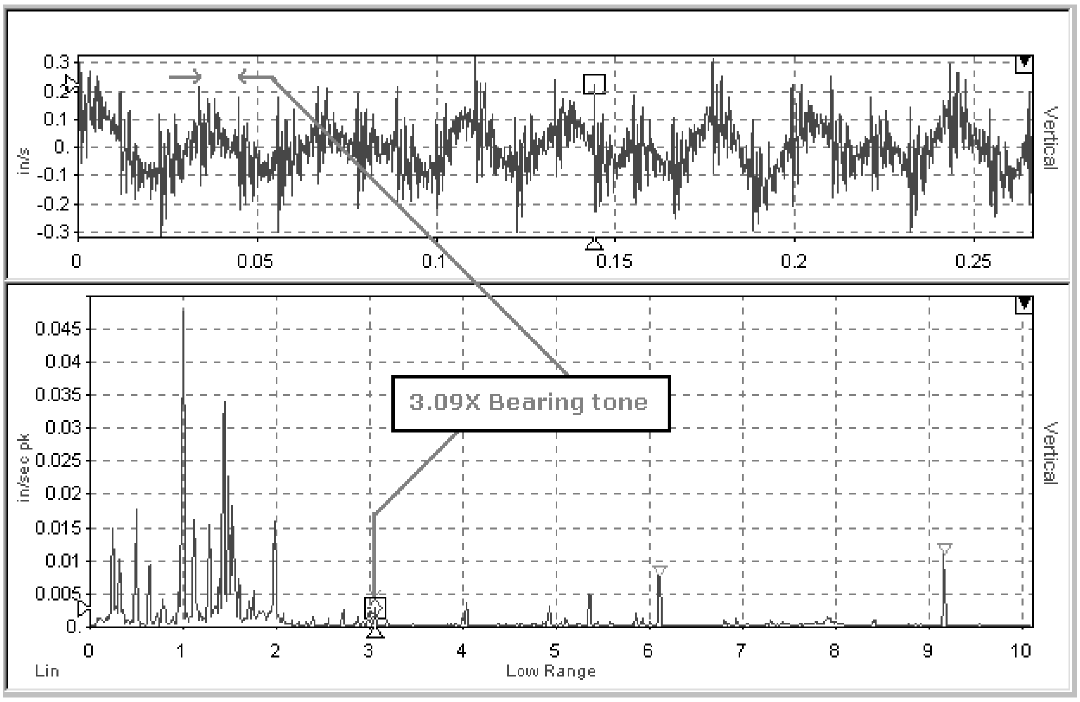


Figure 3-38

In this example, we can see that the time between the pulses is 0.1441 seconds, which relates to 6.94 Hz, or 416 CPM. The running speed of the machine was 1792 CPM, so this frequency is approximately 0.23X.

The distance between the main shaft and the input sheave of the gear was 6" (152 mm). The diameter of the motor sheave was 1.34" (34 mm), and the diameter of the gearbox sheave was 2.64" (67 mm). If we use a formula to compute the linear speed of the belt (known as the belt rate) we see that the belt rate is 0.23 times running speed.

So, we know for sure that the impacts are related to the belt.

$$Belt\ rate\ = \frac{2\ \times Driver\ \phi}{Driver\ \phi + Driven\ \phi + \frac{4}{\pi} \times \sqrt{\left[\left(\frac{Driven\ \phi - Driver\ \phi}{2}\right)^2 + \ Distance\ b/n\ Centers^2\right]}}$$

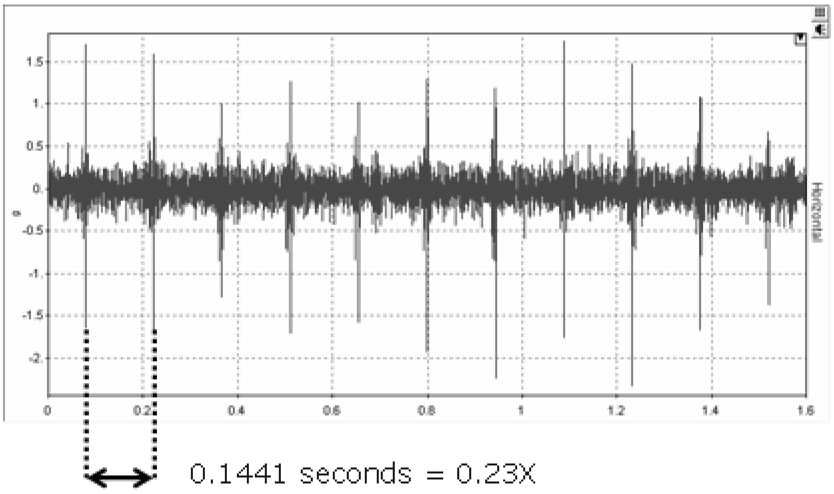


Figure 3-39

Recognizing vibration patterns

In order to perform time waveform analysis it is VERY helpful if you understand the relationship between mechanical phenomenon, signals, the time waveform, and the spectrum. Hopefully when you see sinusoidal patterns, transients and pulses in the waveform (and see harmonics, sidebands or a raised noise floor in the spectrum), you will know why and how they got there, and you can relate that back to what is happening inside the machine.

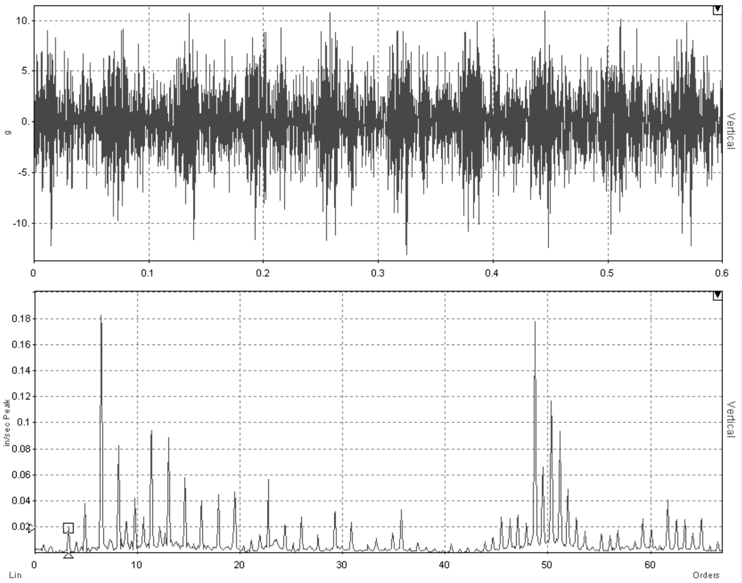


Figure 3-40

If we have a single frequency, the time waveform is very simple, and the spectrum has a single peak. As we add new signals, with different amplitudes, frequencies and phase values, the time waveform becomes more complex, and the spectrum has additional peaks. In example 2 and 3, the frequency and amplitude of the signal components are the same, however the phase relationship is different. Note that the composite waveform looks quite different, whereas the spectrum looks the same.

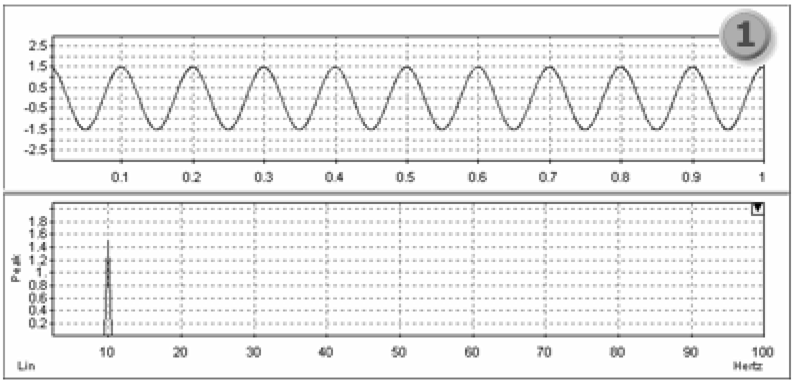


Figure 3-41 - Simple 10 Hz signal

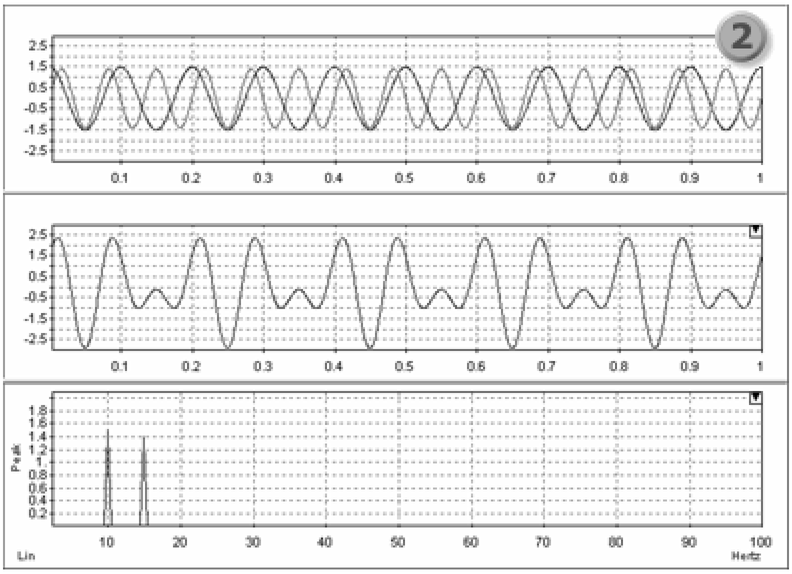


Figure 3-42 - 10 Hz + 15 Hz with 0° difference

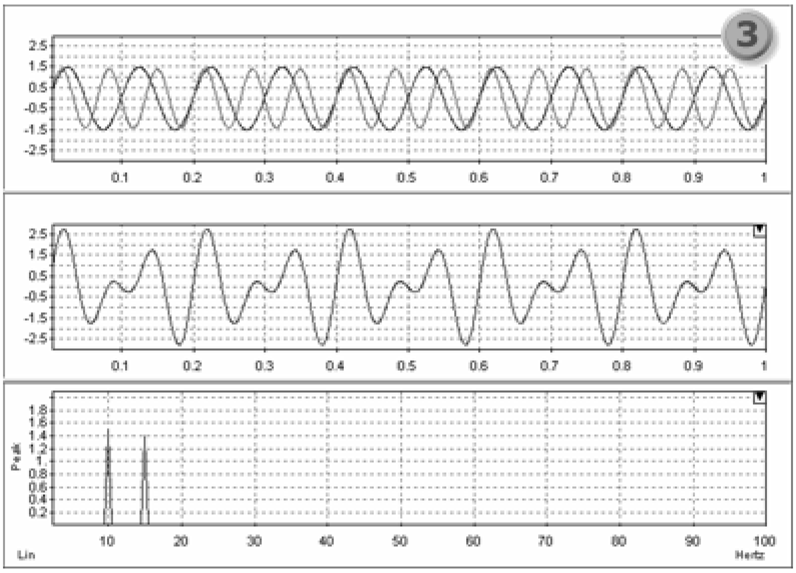


Figure 3-43 - 10 Hz + 15 Hz with 90° difference

Beating

There are a few special cases we should try to understand.

The first is where you have two signals (sources of vibration) with very similar frequencies (less than about 4 Hz difference). In this example the two frequencies are 200 and 202 Hz. While we can see the two peaks (resolution permitting) in the spectrum, the time waveform looks quite interesting. In short, the two signals move in and out of phase, so at one point in the cycle they cancel each other out (if they are the same amplitude – in this case they are not), and half a cycle later they add together. The result is that we hear a throbbing (or beating) sound at the frequency of the difference between the two signals (2 Hz in this example). This can happen when two machines are located close together, with one running at a slightly different speed.

Note that you will also often see a "sum" and "difference" frequency in the spectrum. The "sum" frequency is equal to the sum of both frequencies: 200 + 202 = 402 Hz. The "difference" is the difference between them: 202 + 200 = 2 Hz.

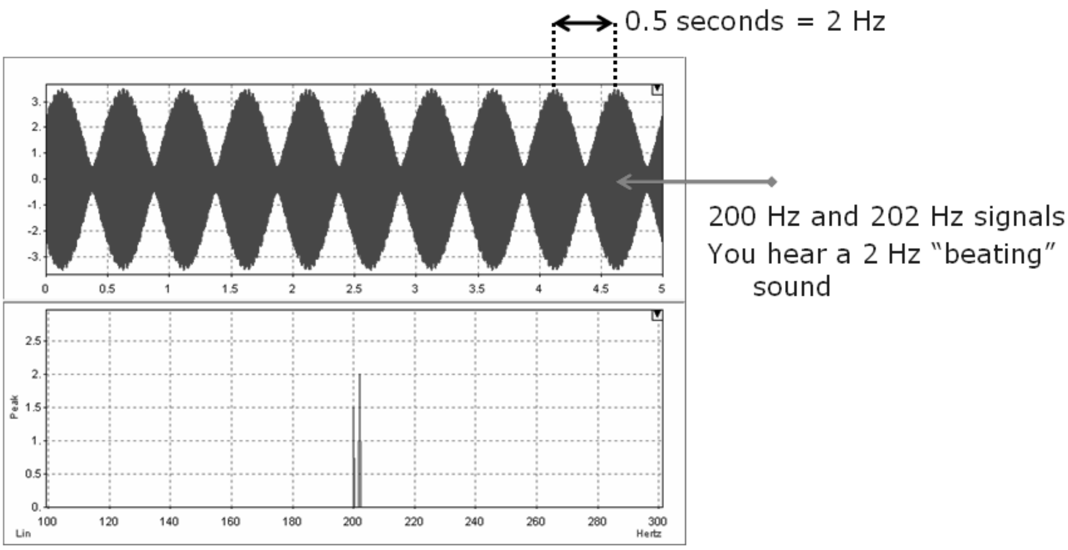


Figure 3-44

Modulation

The next phenomenon to look at is amplitude modulation.

This occurs when one signal (called the carrier) changes in amplitude periodically – i.e. at another frequency. The time waveform looks similar to the beating example, however the underlying mechanism is quite different, and the spectrum is different. We typically see the carrier frequency, and we see sidebands. For example, if the carrier was 200 Hz, and the amplitude of the signal was varying at 20 cycles per second, we would see a peak at 200 Hz, with sidebands at 180 Hz and 220 Hz (and often also at 140 Hz, 160 Hz, 240 Hz, etc.).

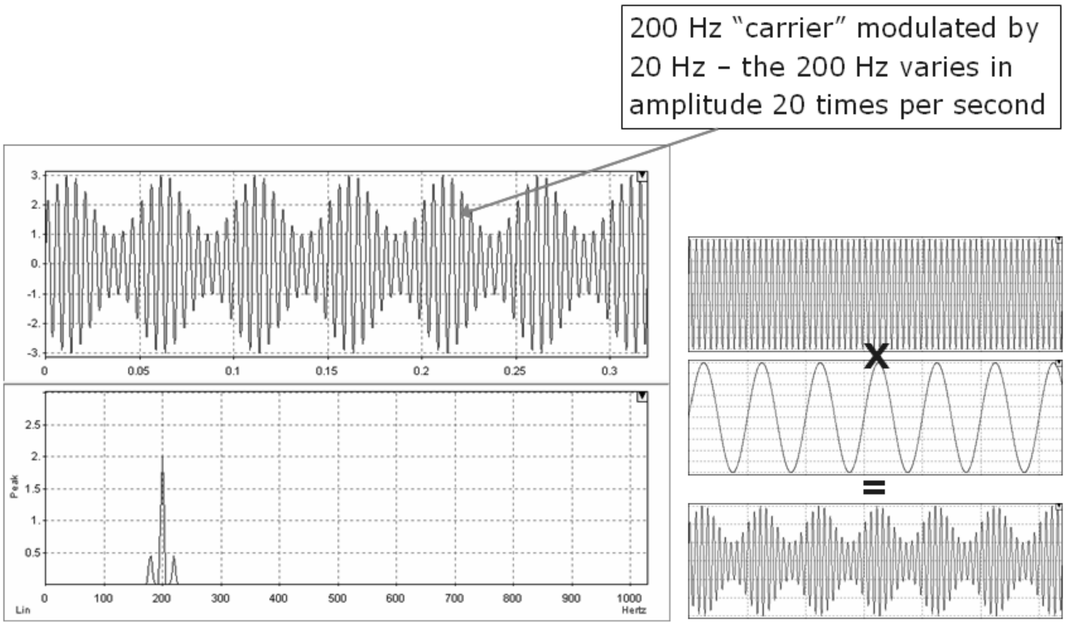


Figure 3-45

Bearing faults and amplitude modulation

Why are we discussing modulation? Well, we often see it when diagnosing common fault conditions, especially rolling element bearings and gearboxes.

Imagine, if you will, a ball rolling around inside a bearing – a ball with some damage. Now, as it rolls around, the damaged area impacts the inner and outer race, which generates vibration. The frequency of the vibration can be calculated, as it depends upon the size of the ball and the diameter of the inner and outer race. But the amplitude of the impacts is not constant – and therefore the vibration is not constant. You see, as the ball rolls into the load zone of the bearing, the forces are greater, reaching its greatest level when it is in the center of the load zone. But as it rolls out of the load zone, the force of the impacts is reduced, reaching its minimum level when the ball is on top of the inner race (out of the load zone).

The result is that the amplitude of the vibration rises and falls periodically. The "carrier" frequency is the ball spin frequency, and it is modulated by the cage frequency (the rate at which the ball moves around the bearing). The spectrum would have a peak at the "ball spin" frequency, and there would be "fundamental cage" or "FT" sidebands.

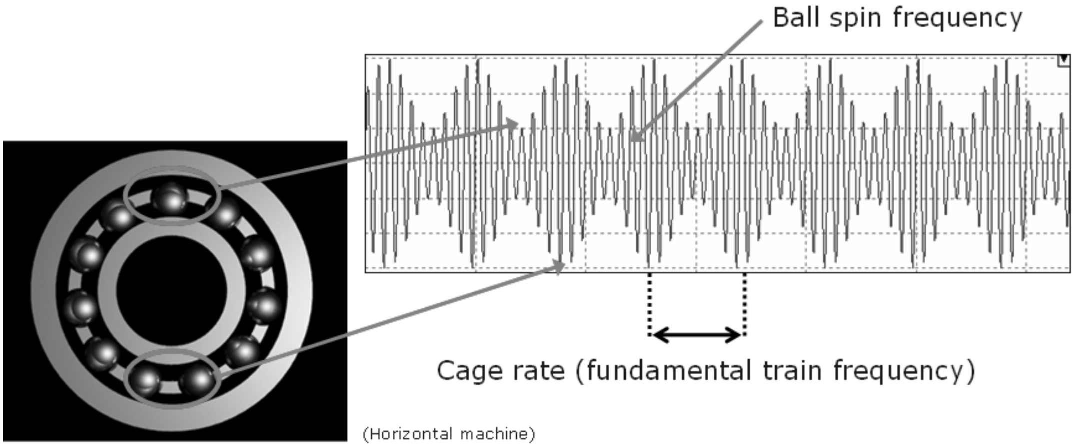


Figure 3-46

Similarly, if there was a defect on the inner race of the shaft, as that defect comes into contact with balls/rollers in the load zone the vibration is higher than when it impacts out of the load zone. This time the carrier is the "ball pass inner race" or "BPI" frequency, and the modulating frequency is running speed (as the inner race rotates once per shaft revolution). The spectrum would have a peak at the "ball pass inner race" frequency, and there would be "running speed" or "1X" sidebands.

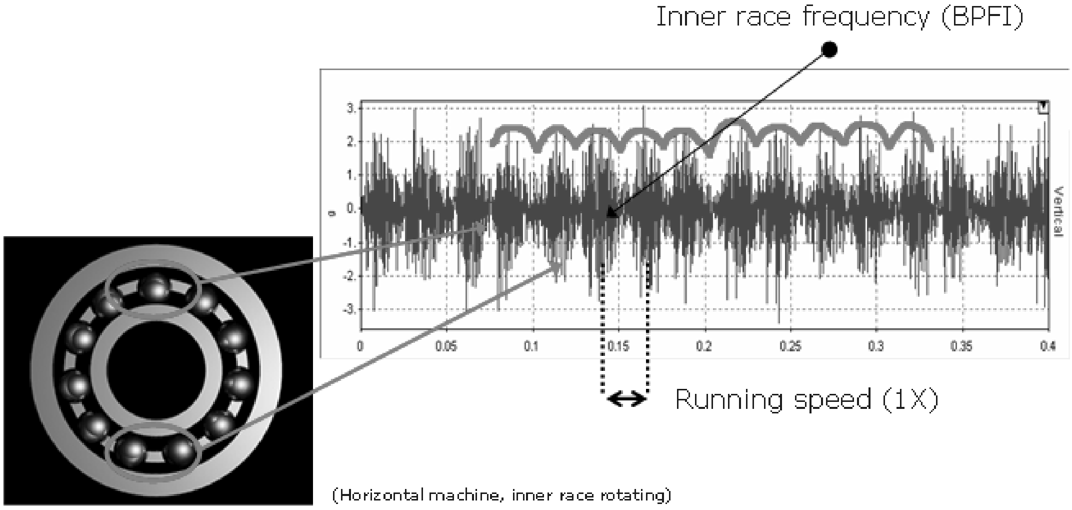


Figure 3-47

Here is another classic case of amplitude modulation. The two gears are meshing together, but they are slightly misaligned.

You can see two modulating patterns, one for the larger, slower turning shaft/gear, and another lower amplitude pattern for the smaller, faster turning pinion gear. The spectrum also had 1X sidebands around the gearmesh peak. The spacing between the sidebands corresponds to the speed of the slower output shaft.

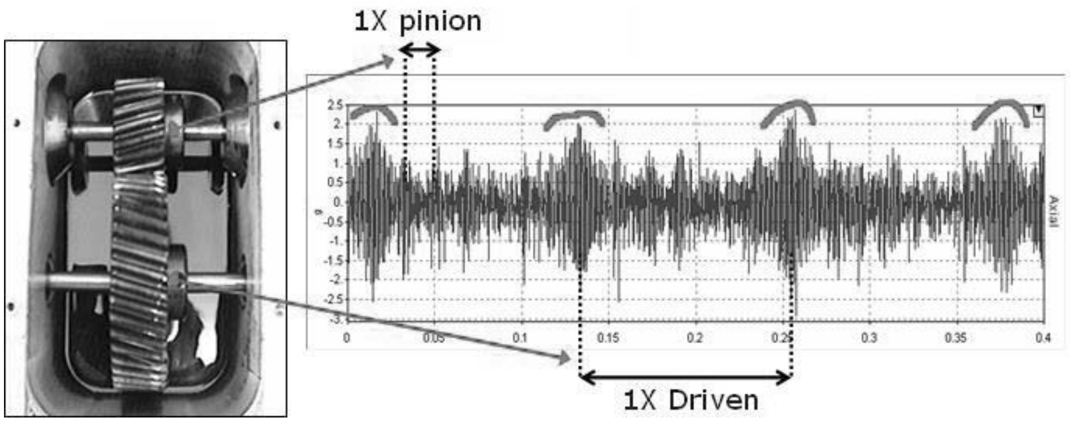


Figure 3-48

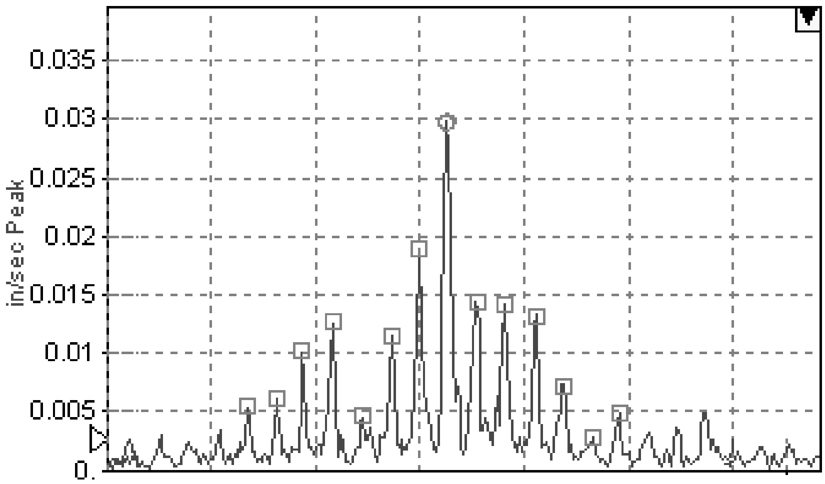


Figure 3-49

"Non linear" clipped vibration

Another very common phenomenon is "non-linear" vibration, or truncated waveforms. If movement in one direction was restricted, then instead of a classic sinusoidal signals, you instead find that the waveform is clipped. In this example we have two 100 Hz signals, one clipped, and one normal. The normal sine wave of course has just a single peak in the spectrum, but the clipped waveform has harmonics – peaks at 100 Hz, 200 Hz, 300 Hz, and so on.

In more severe cases, the harmonics will appear at 1/2 the frequency, or 1/3 the frequency. We'll see some examples shortly.

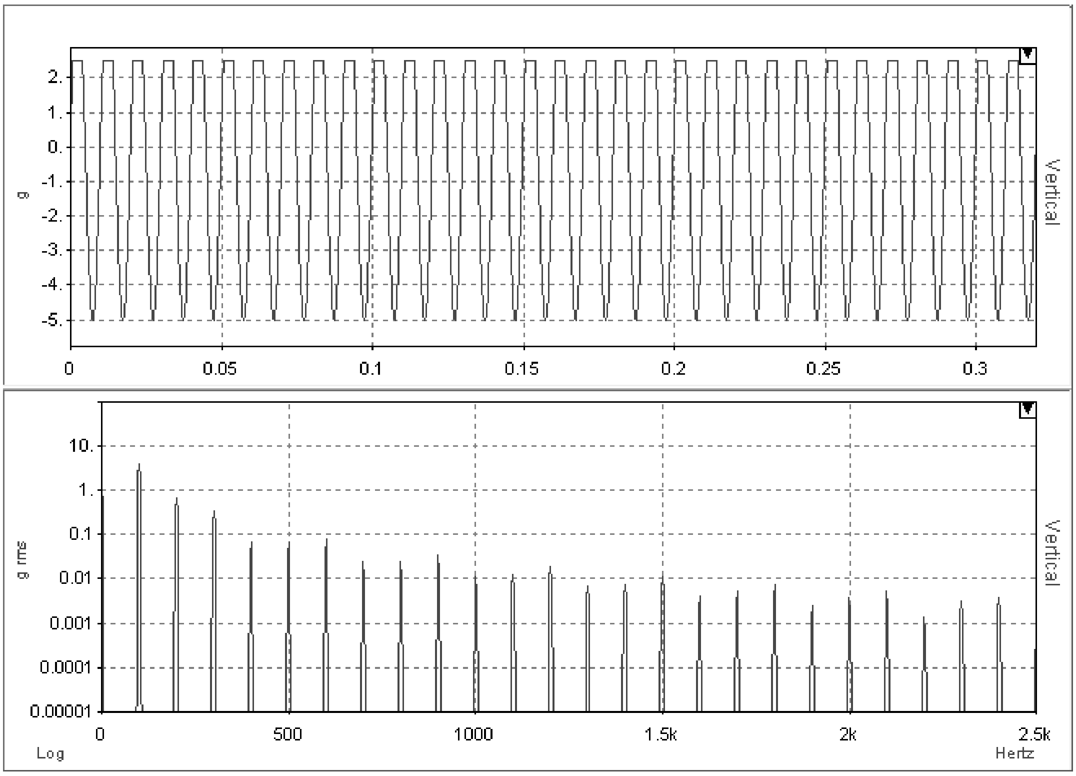


Figure 3-50

Impacting

And yet another common phenomenon in when studying vibration from machines is impacting. Impacts can occur as defects on rollers strike the inner and outer race of a bearing; impacts occur as loose parts rattle; and there are lots of other examples. If we were to strike a machine and look at the waveform we would see one pulse or transient, and the spectrum would have energy at all frequencies. The spectrum does not have any peaks, because there are no repeating patterns in the waveform, but the noise floor has been raised. So, if impacting occurs we would expect the noise floor of the spectrum to be raised.

In this example, there was actually an electrical spike in the waveform. You can also see the raised noise floor, and the classic "ski-slope" at the beginning of the spectrum. If you see this ski slope, you can always assume that there was a severe impact or some other kind of transient, and you should repeat the test and look for the reason for the ski-slope (there are a few possibilities).

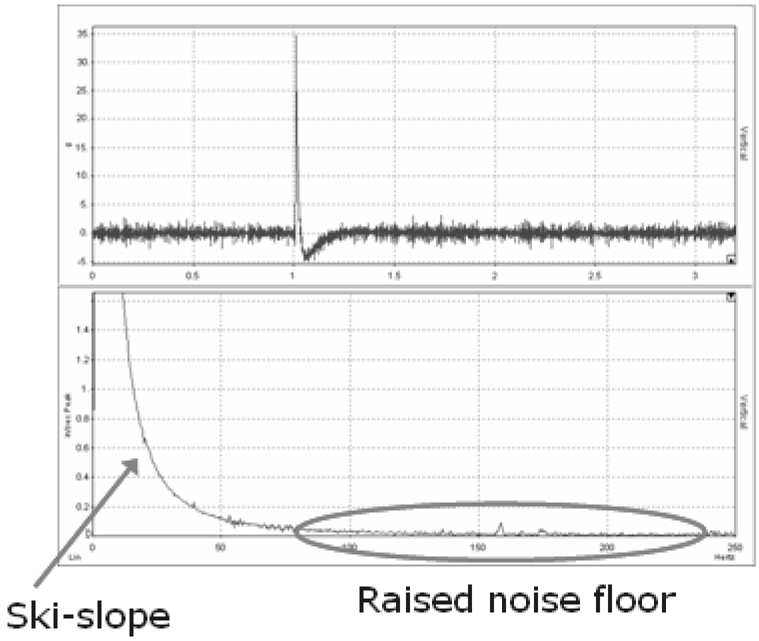


Figure 3-51

Rotating looseness

If the impacts are repetitive, for example the impact occurs once per revolution (as in the case of rotating looseness), not only do we see the raised noise floor, we see two other characteristics. First, because the impacting is repetitive, we see a peak in the spectrum at the repetition rate. If there was one impact per revolution, there will be a peak at 1X (running speed) in the spectrum. Second, because the impacting is "non-linear", we see harmonics in the spectrum (as discussed in 'understanding signals').

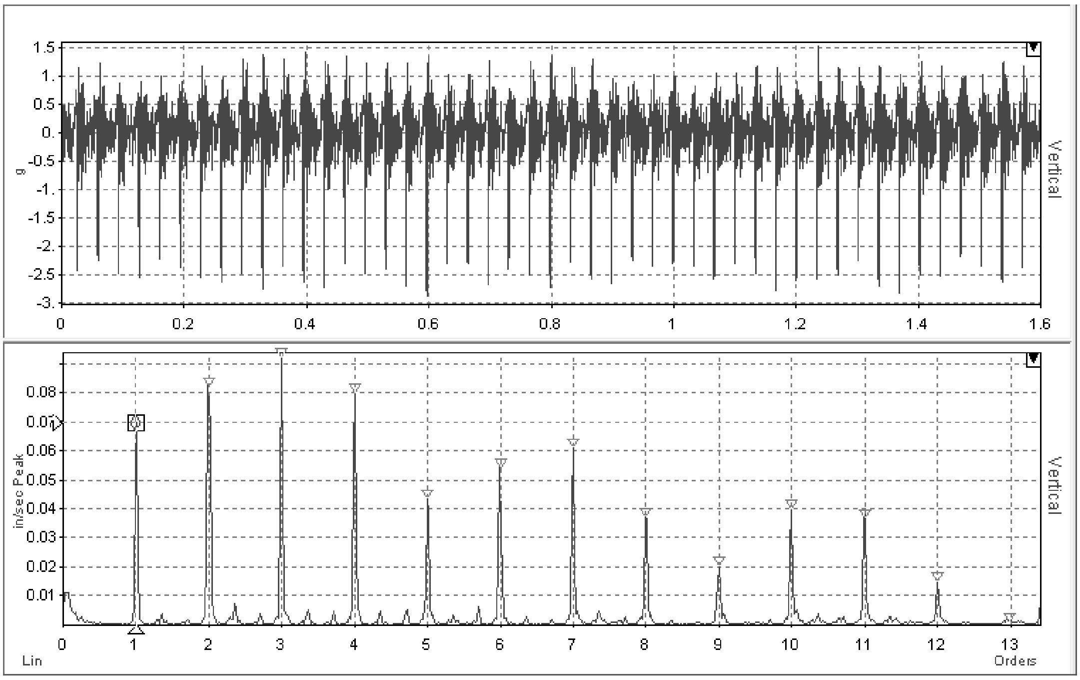


Figure 3-52

Let's assume that a pedestal bearing is loose, and as the shaft turns it is rattling up and down. If the hold-down bolts were very loose, and there were no impacts, there would be no harmonics – just a high 1X peak (in the horizontal direction) as it rocked back and forth. But we do have severe impacts, so we get lots of harmonics. You can clearly see the impacts, and you can see the harmonics in the spectrum (in fact, you can see that there are actually 1/3 harmonics, a sure sign of severe looseness).

There is something else interesting about this data. Do you notice that the negative going peaks (0.6-1 G) are greater than the positive going peaks (0.2-0.4 G)? That means that the bearing had slightly greater freedom in one direction than the other, which is consistent with the looseness diagnosis.

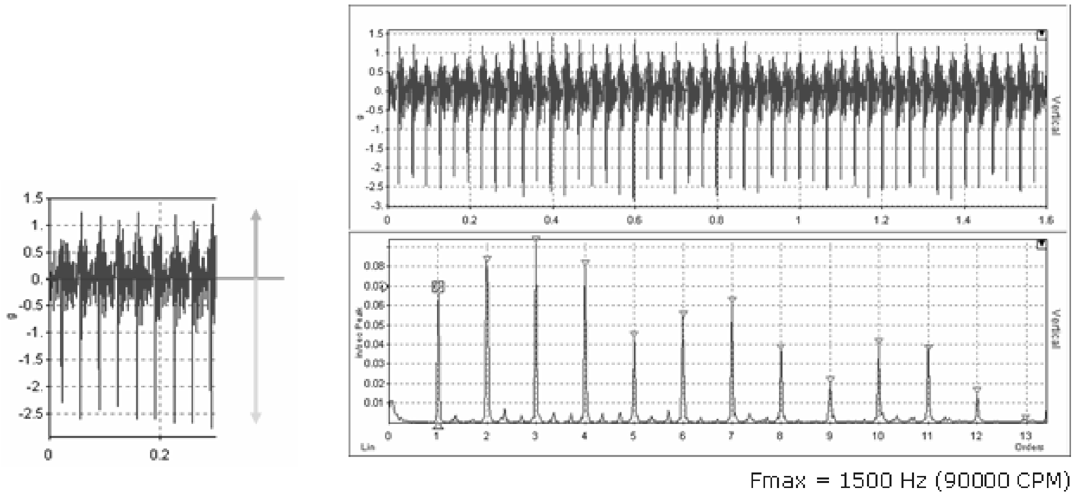


Figure 3-53

Cavitation

This is an interesting example: cavitation in a pump. If you could only listen to this data (which you can do in our training product), you would be able to hear a grinding, gravelly sound. If you look closely to the pulsations in the waveform you will see that they are not evenly spaced. The source of the vibration is not periodic. You will see in the spectrum that the noise floor has been raised, especially at the higher frequencies.

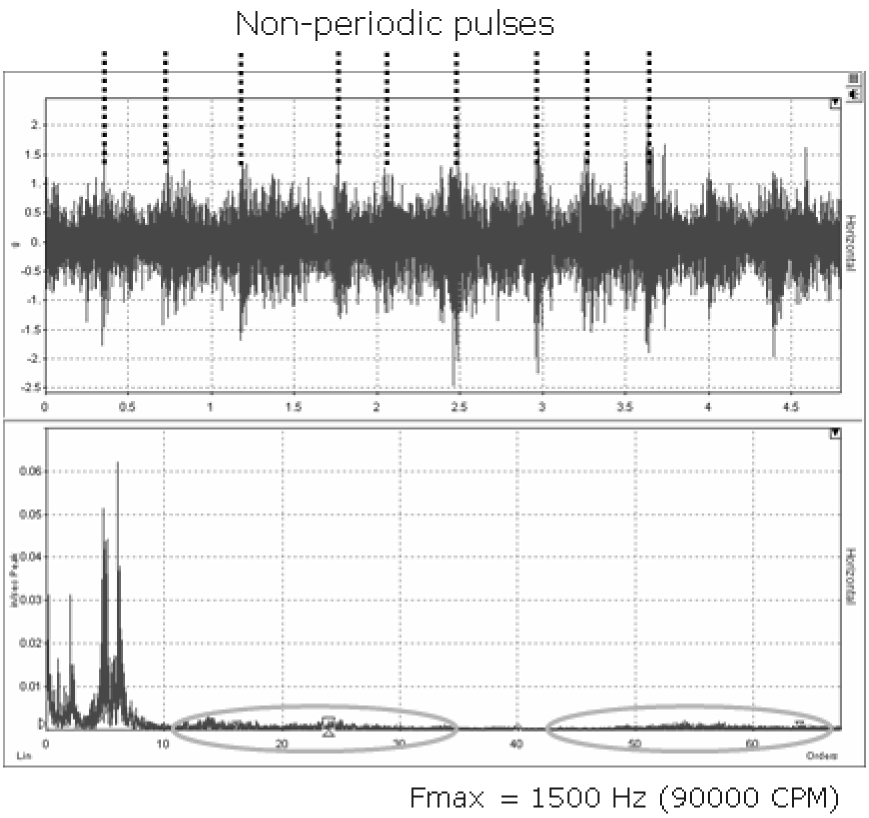


Figure 3-54

Measurement directions and storage

Remember that machines vibrate in three axes. When performing normal spectrum analysis it is always important to have an axial, vertical and horizontal reading. The same is true with your time waveform readings. Because some faults only show up in certain axes, or they can be best understood by comparing the level or pattern across axes (or between different points on a machine), you should perform the same comparisons with your waveform data. It does not require any extra collection time, and you should have sufficient memory capacity in your data collector.

If you do have limited memory, then you must decide which axis will provide the best data. Think carefully about the type of fault you are looking for, and consider in which direction the key forces.

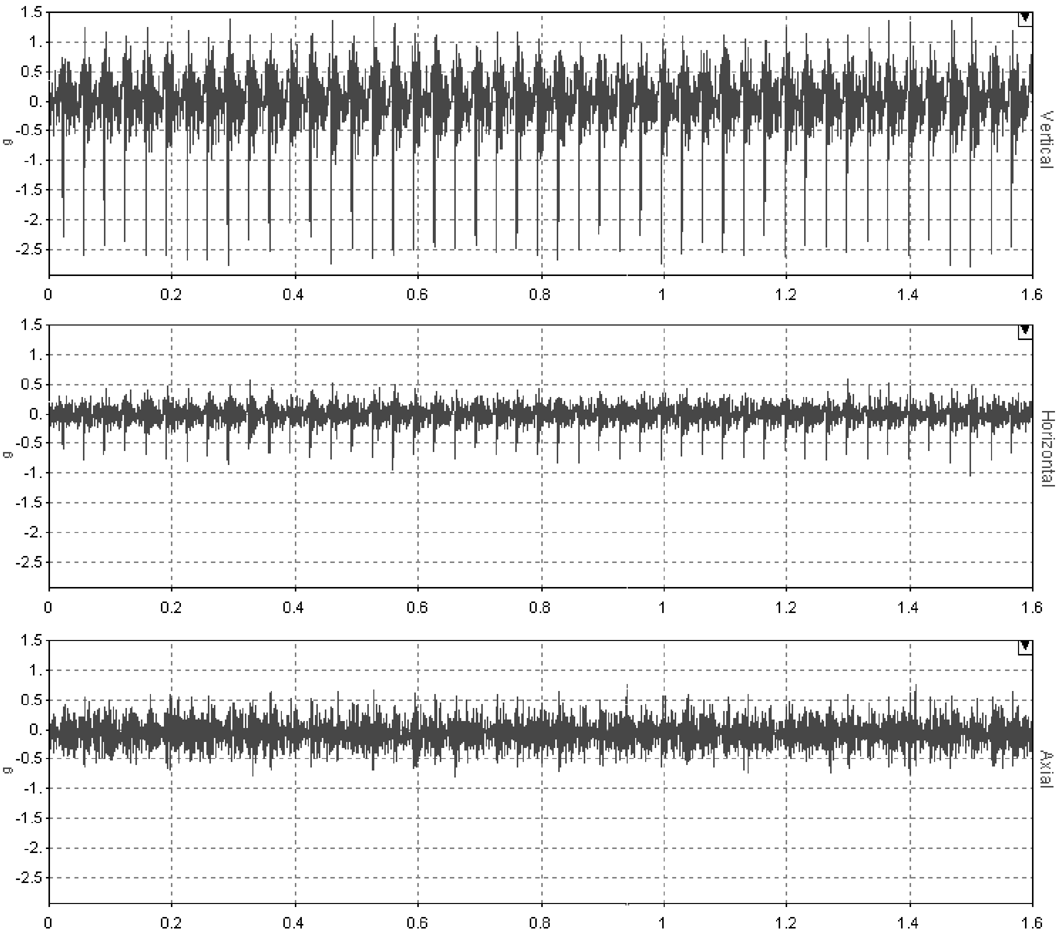


Figure 3-55

Gearbox fault analysis

If you have a gearbox then you must acquire and analyze time waveforms – it is as simple as that. While there are spectral patterns that you can look for, there is no doubt that the time waveform reveals the clearest information.

We discuss the details in the gearbox analysis section, but there are broadly two types of gearbox fault:

If there is misalignment or a bent shaft, for example, the vibration will be relative constant as each shaft turns (although the forces will change periodically) – therefore the information in the spectrum will better represent the fault condition. The gearmesh frequency (and or 2xGM and 3xGM) will change in amplitude , and the sidebands around the gearmesh peaks will change in amplitude.

If there is damage on one or more teeth, then the vibration will be different when that tooth is in mesh. The change in vibration will be momentary. If one tooth is damaged (cracked, broken, or worn), then the vibration will change once per revolution. The spectrum will change, but it will change in less predictable ways. The time waveform, if collected correctly, will reveal the exact nature of the fault condition. You may be able to tell how many teeth are damaged, and using techniques like time synchronous averaging, you will be able to perform this analysis even in complex gearboxes.

The following spectrum is from a gearbox. It does not indicate that a fault condition exists.

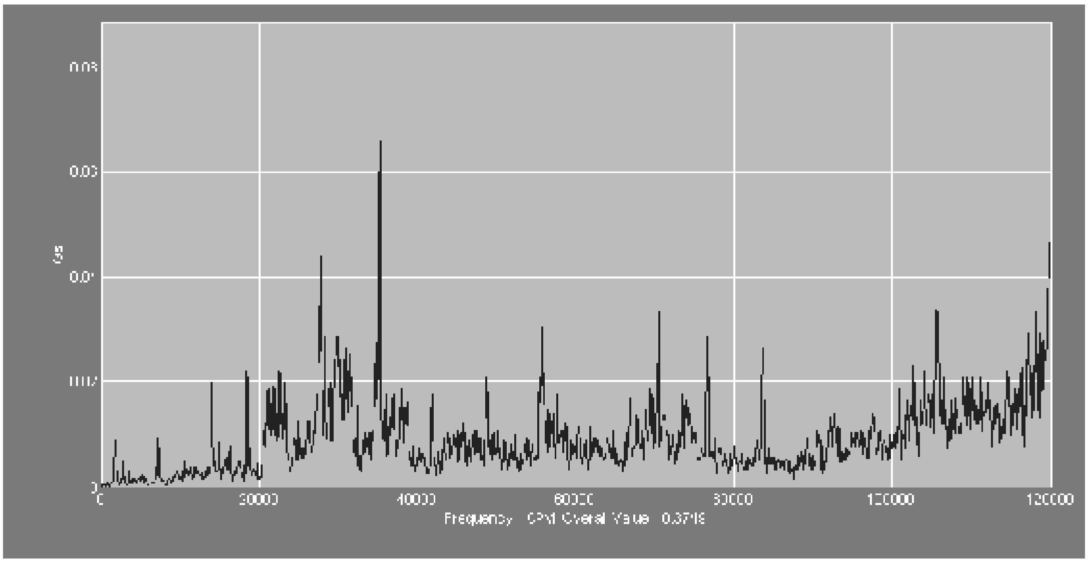


Figure 3-56

When you look at the time waveform, it is clear that a fault exists.

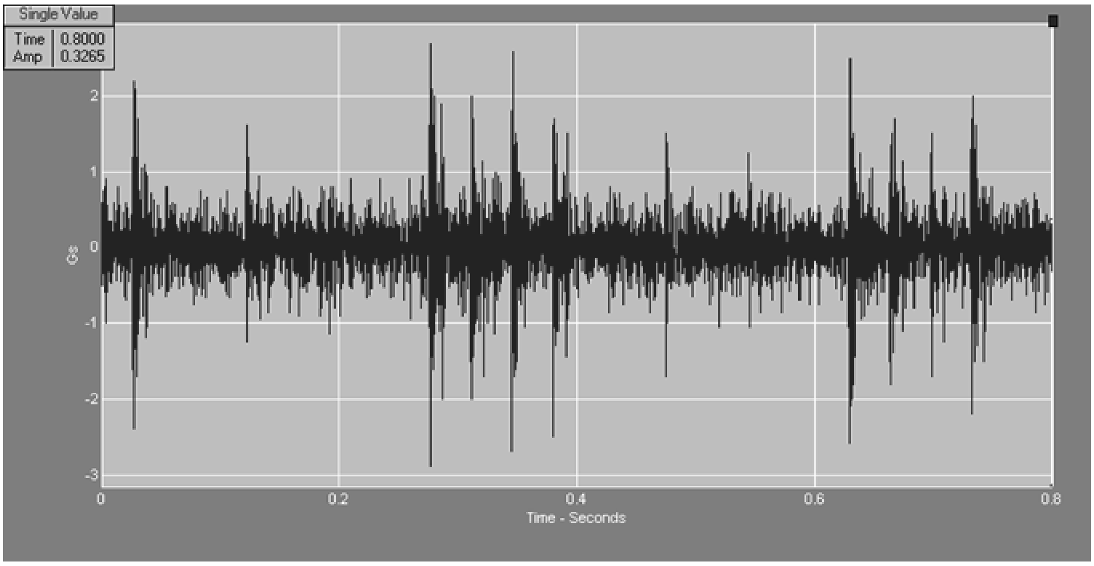


Figure 3-57

The time waveform clearly indicates that a fault exists, and even that the damage is confined to one area. Although the graphic is very poor, the green waveform was taken from the machine when it was repaired, and the blue waveform was collected when the gearbox was damaged.

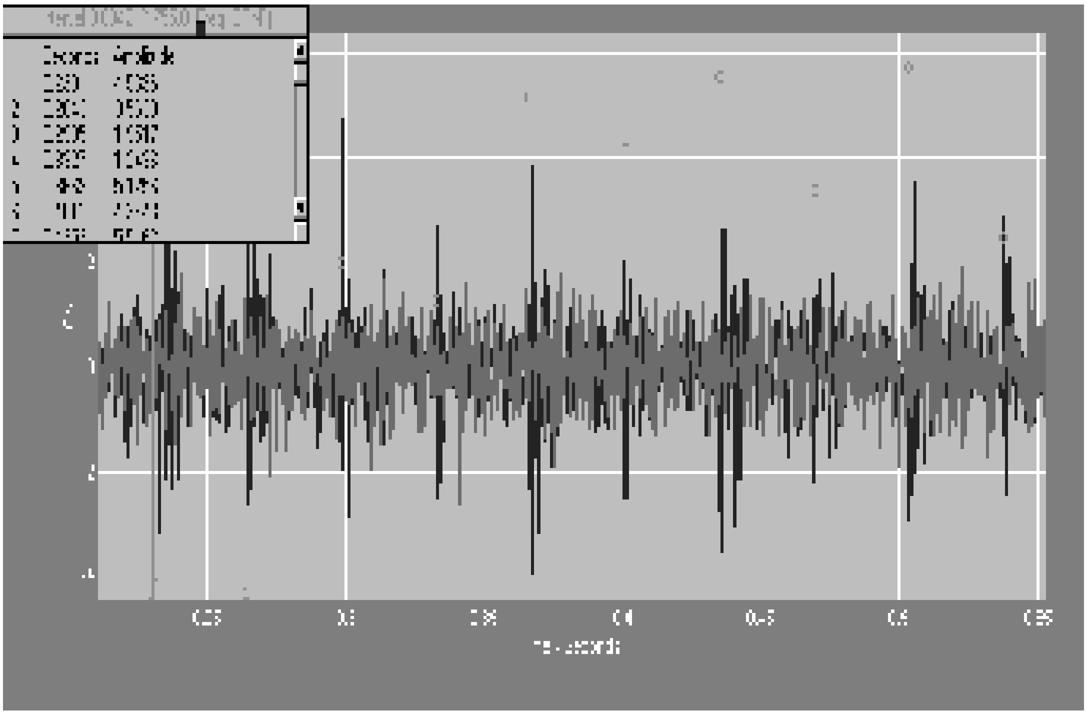


Figure 3-58

Here we have another example where the spectrum does not indicate that a problem exists, whereas the time waveform clearly indicates that one of the gears is damaged in one area.

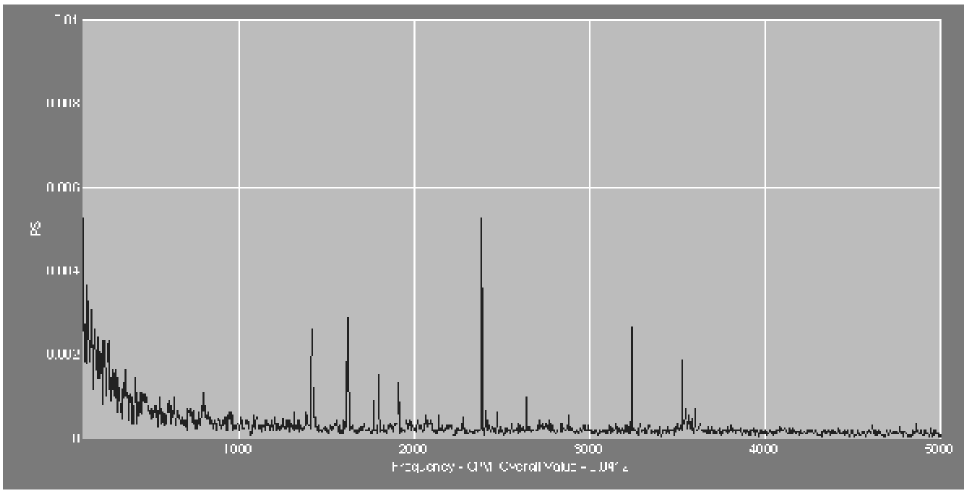


Figure 3-59

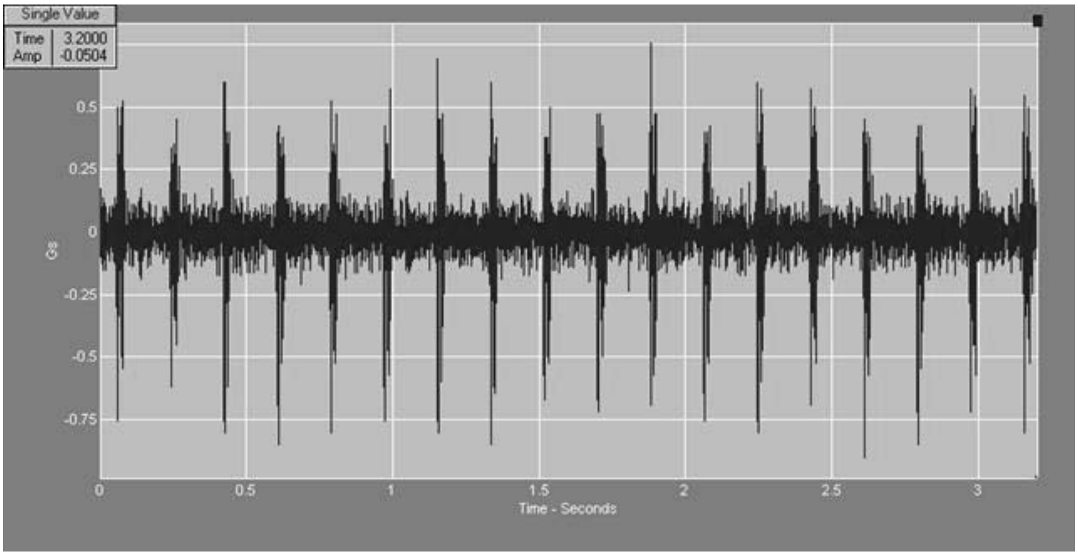


Figure 3-60

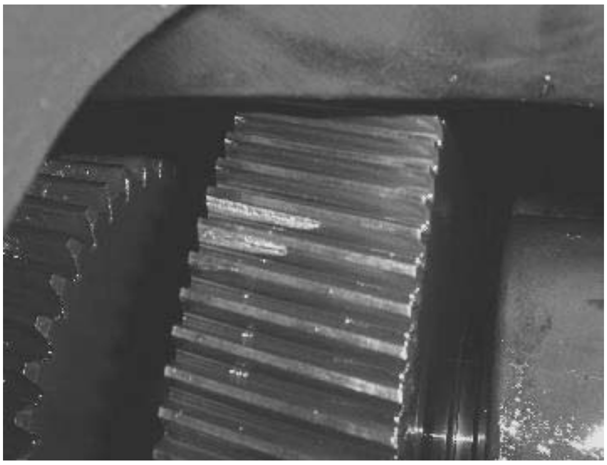


Figure 3-61

And here is one more example. The waveform clearly indicates that a fault exists, yet the spectrum does not provide the same level of information.

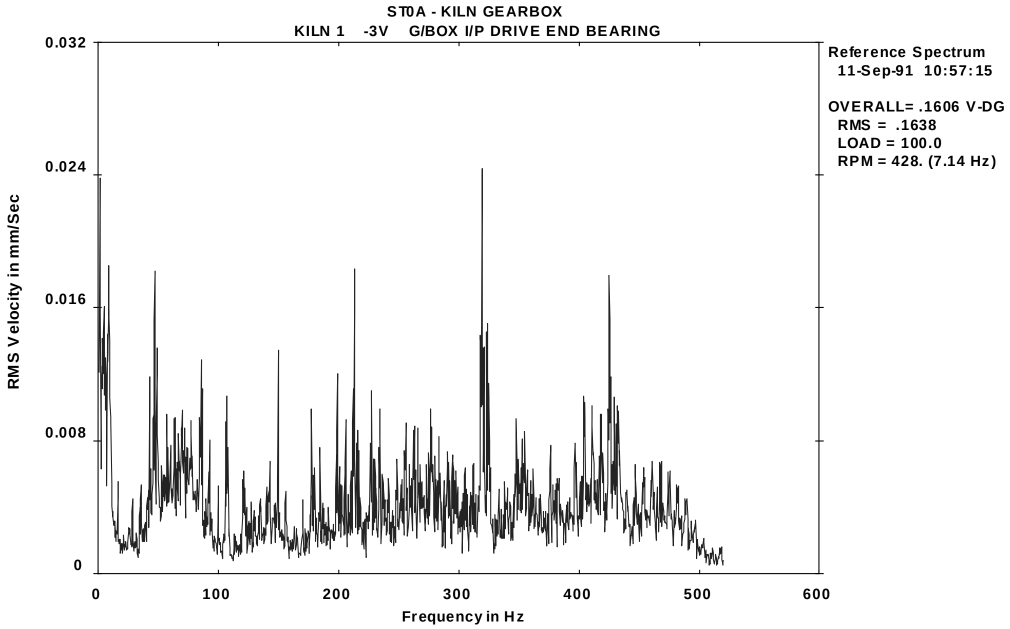


Figure 3-62

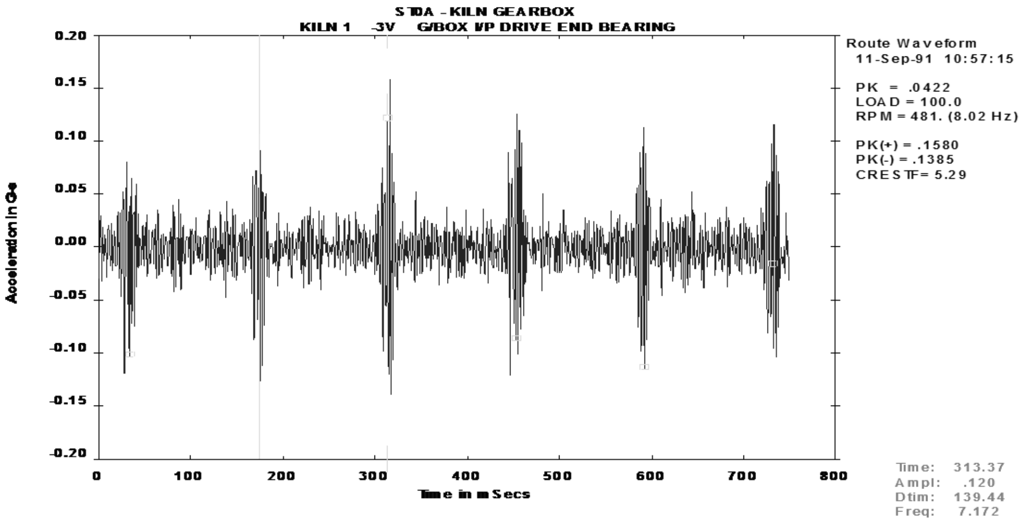
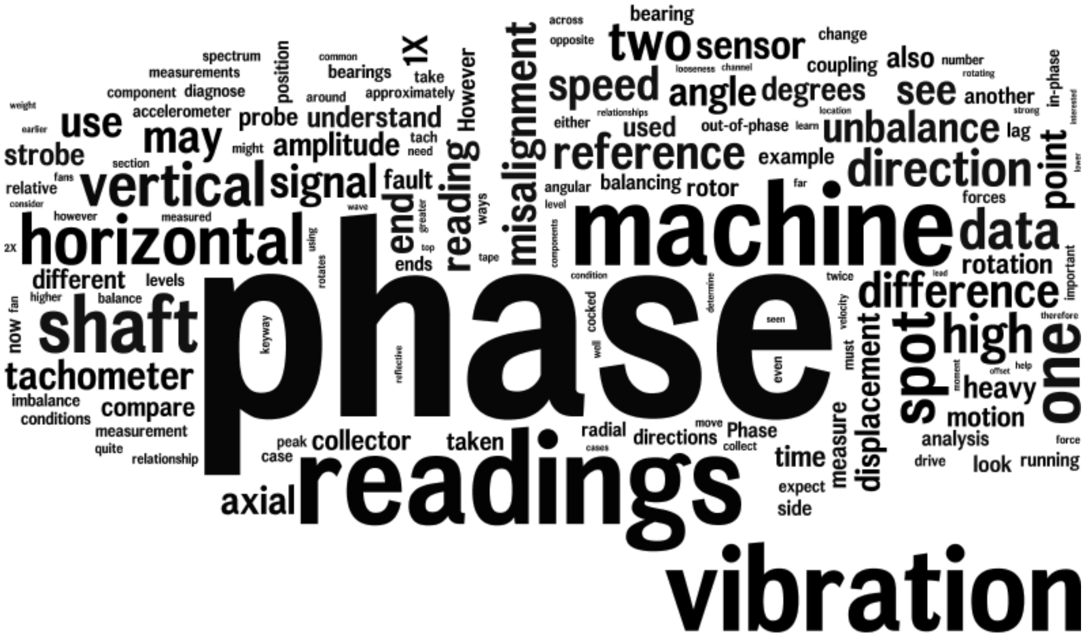


Figure 3-63



Chapter 4

Phase Analysis

Objectives:

- Describe the concept of phase in detail including leading and lagging, measuring phase and the relationship between phase and integration.
- Describe the concepts of heavy spot, high spot and phase lag.
- Review the use of phase in diagnosing common machine faults

This chapter discusses all aspects of phase from measurement to applications and introduces the concepts of high spot, heavy spot and lag. A good understanding of phase will be important when we discuss structural measurements in coming chapters. We also review the phase relationships associated with common machine faults to serve as a review of important Cat II topics.

Introduction

For so many vibration analysts, life revolves around the spectrum. If the fault is not obvious in the spectrum then the fault will not be detected. And in many cases, one fault condition may be confused with another because spectral data alone is used. Vibration analysis is not easy, and it is important that we diagnose faults correctly. If you master phase analysis, your ability to diagnose faults correctly will be enhanced greatly.

Phase analysis is a very powerful tool. The perception may be that phase measurements are difficult to collect, or possibly that they are difficult to understand. Some may even believe that phase measurements do not offer any useful information. They are wrong.

The aim of this section is to show that phase measurements are neither difficult to collect nor difficult to understand. And by the end of this section, I hope you will understand the benefits of phase in machine fault diagnosis.

We will start by revising phase, and understand the terms phase lead and lag. We will learn how system resonances can affect the phase that we measure. We will learn how different types of transducers affect phase readings. We will learn about different ways to collect phase readings.

We will then look at how we can interpret the phase readings. We can use some simple rules that will help us to diagnose unbalance, misalignment, bents shafts, cocked bearings, and looseness faults.

We will then go even further. We will then study how we can utilize phase data to better understand the dynamic movement of a machine. We will start with some simple methods, but then look at the operating deflection shape "ODS" method in far greater detail, as well as modal analysis and finite element analysis.

By the end of this section you should feel very comfortable with the types of measurements that are required, and with the various methods that can be used to utilize the data in order to diagnose fault conditions and improve the reliability of your machinery and other plant assets.

What is phase?

We need to step back and make sure that we understand phase. Having performed single- or two-plane balancing you will be familiar with phase, but if you predominantly study vibration spectra (and ignore time waveforms), phase may be one of those concepts that you "sort of understand".

Phase is all about timing

Phase is all about the relative timing of related events. Here are a few examples:

When balancing we are interested in the timing between the heavy spot on the rotor and a reference point on the shaft. We need to determine where that heavy spot is located, and the amount of weight required to counteract the rotational forces.

When we look at fault conditions such as unbalance, misalignment, eccentricity, and foundation problems, we are interested in the dynamic forces inside the machine, and as a result, the movement of one point in relation to another point.

We can use phase to understand the motion of the machine or structure when we suspect a machine of structural resonance, where the whole machine may be swaying from side to side, twisting this way and that, or bouncing up and down.

So, phase is very helpful when balancing (there are other ways to perform a balance, but most data collectors utilize phase), and essential when trying to understand the motion of a machine or structure. But phase is also very useful when trying to diagnose machine fault conditions. If your attitude is "the vibration levels are high – it needs to be overhauled", then you probably don't care about phase. But if you want to make an accurate diagnosis, and correctly distinguish between faults such as unbalance, misalignment and bent shaft, then phase is an essential tool.

Phase fundamentals

When a shaft rotates you can measure the vibration at the frequency of rotation and you will see a sine wave. The vibration level will be dictated by a number of factors, but let's just focus on the forces due to unbalance.

Let's use an eight-bladed fan as our reference machine. There is a gold coin attached to one of the blades which generates the unbalance force. We see a sine wave as follows:

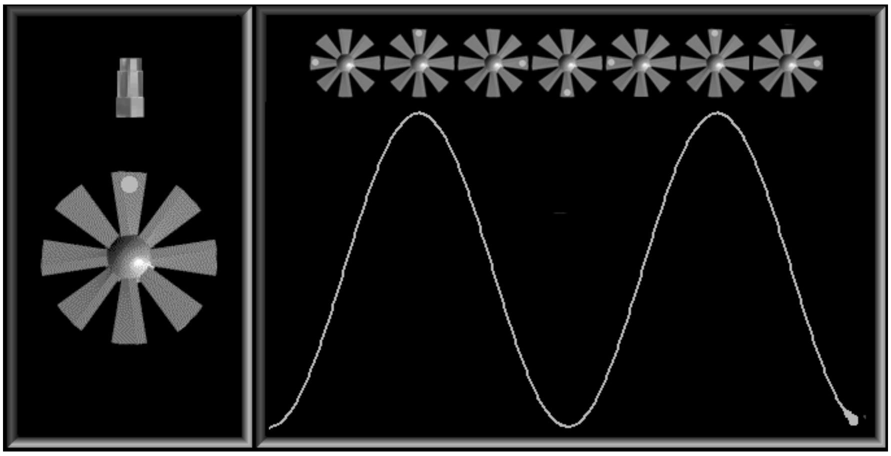


Figure 4-1

As you should already know, you get one rotation of the fan for each cycle (i.e. 360°) of rotation.

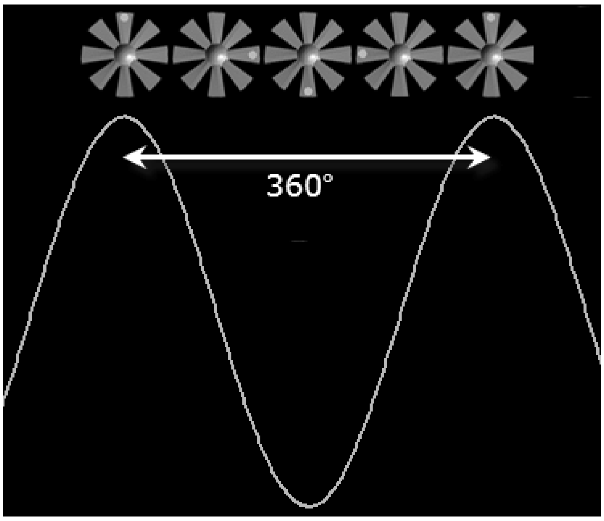


Figure 4-2

But this information by itself does not tell us very much. Phase is a relative measure, so we need to compare one source of vibration to either another source of vibration or a reference of some kind.

First we'll try to understand phase by comparing two sources of vibration. If we had two identical fans, each with coins on a blade (to generate an unbalance force), we would expect to see sine waves from each fan. If the fans were turning in perfect synchronization such that the coins were both at the 12:00 position at the same time, they would be said to be "in-phase".

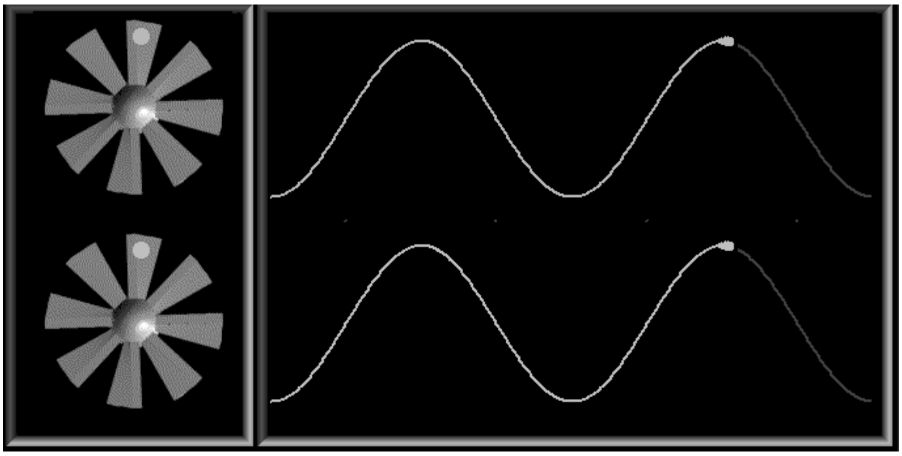


Figure 4-3

However, if one coin was at the top (12:00) when the other was at the bottom (6:00), they would be "180° out-of-phase". Why 180°? Because one rotation is 360°, so half a rotation is 180°.

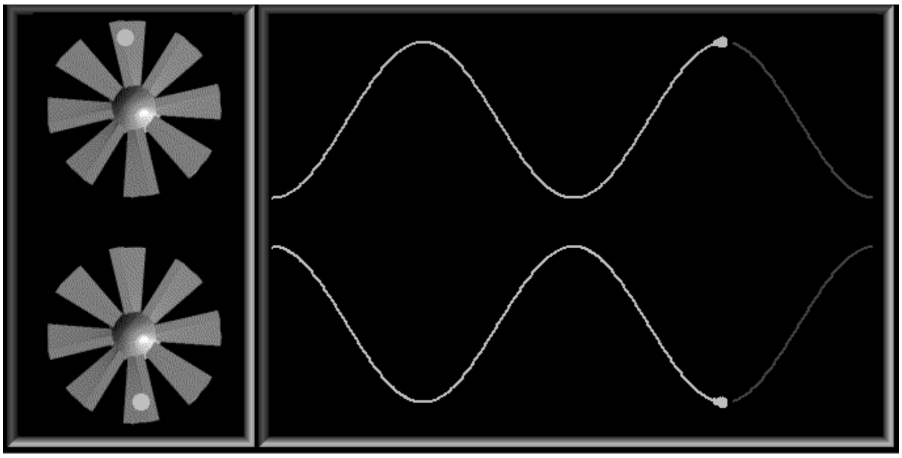


Figure 4-4

And if one coin was at the top, and the other was a quarter of a rotation around, they would be 90° (or 270°) degree out of phase.

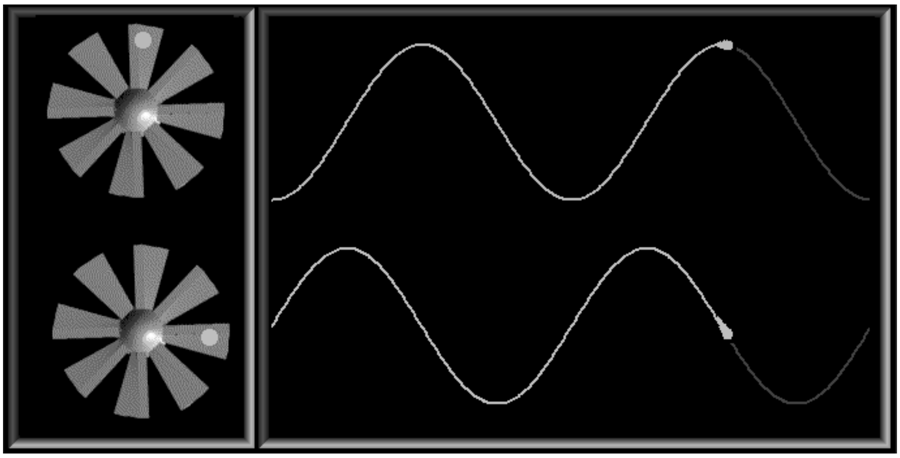


Figure 4-5

In reality, the phase angle between the two fans could be anything from 0° to 359°.

In this particular example (of two fans), we don't really care what the phase angle is, but there are many situations where we do. So we have two different ways to determine the phase difference.

Comparing two waveforms

First, if you look at the previous examples you can see two waveforms – of the same frequency (the fans are running at exactly the same speed). By comparing the two time waveforms we can see the time difference between them.

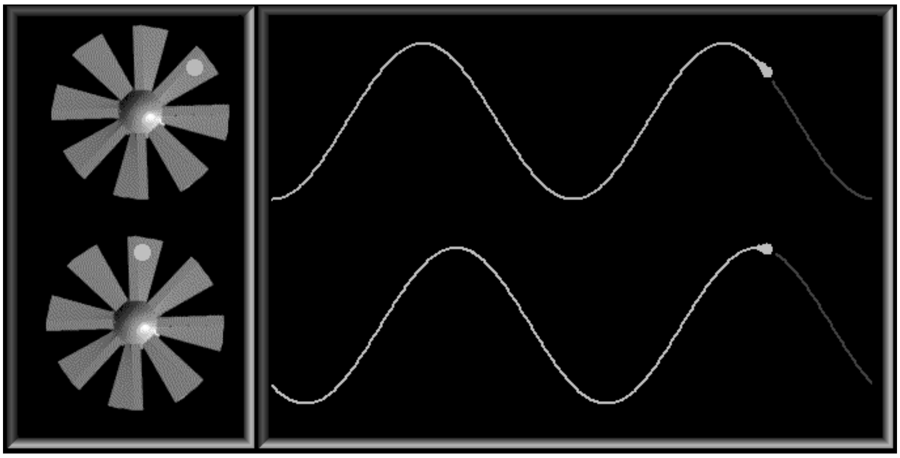


Figure 4-6

We can determine the phase difference by first measuring the time for one complete cycle (remember, one cycle is 360 degrees) and comparing that to the difference in time between the waves.

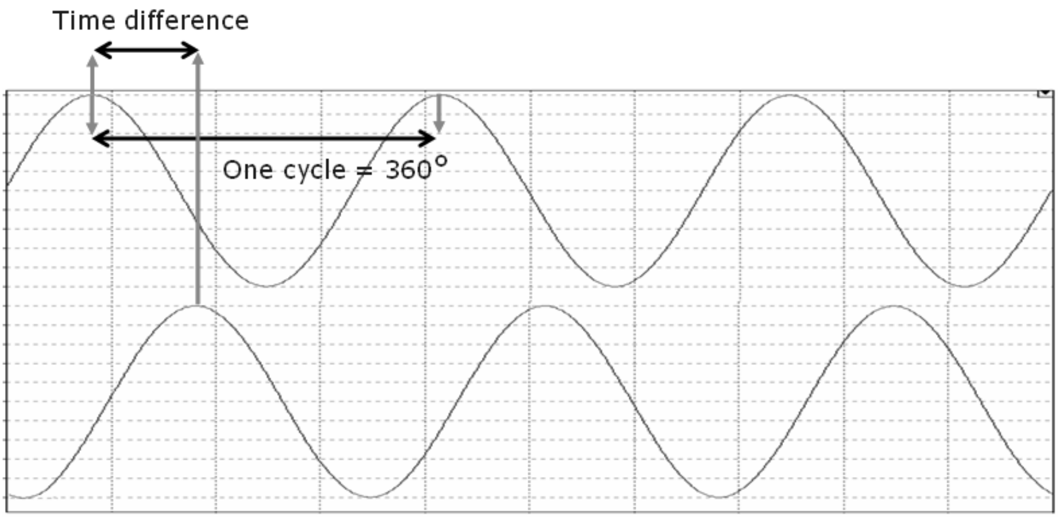


Figure 4-7

So, we measure the time for one cycle, and compare that to the time difference between the same points (in the case above, the point that the wave is at its peak) in both waveforms.

Using a reference

If we were to walk up to the fan (with the coin attached to one blade) and take a vibration measurement we might expect to see a simple spectrum with one dominant peak, and a time waveform that looked like a sine wave – assuming that everything else within the machine is in perfect condition.

If the amplitude was high enough we might assume that the fan was out of balance. But can we tell where to place the balance weight to counteract the out-of-balance force (in this example a gold coin)? No, we can't. We need a reference so that we can compare the point in the rotation where the vibration was highest (the heavy spot) to a known point on the shaft. (For the moment we will assume that we are dealing with a rigid machine and no phase lag, so the heavy spot is at the high spot.)

The answer is to compare the vibration signal to a reference from the same machine. We can use a tachometer reference. The two most common methods are to place a piece of reflective tape on the shaft and then use a photo cell or laser to generate a pulse each time the shaft rotates, or to use a displacement probe opposite a keyway. Each time the keyway passes the tip of the displacement transducer, the measured displacement changes dramatically so the signal will have a step change.

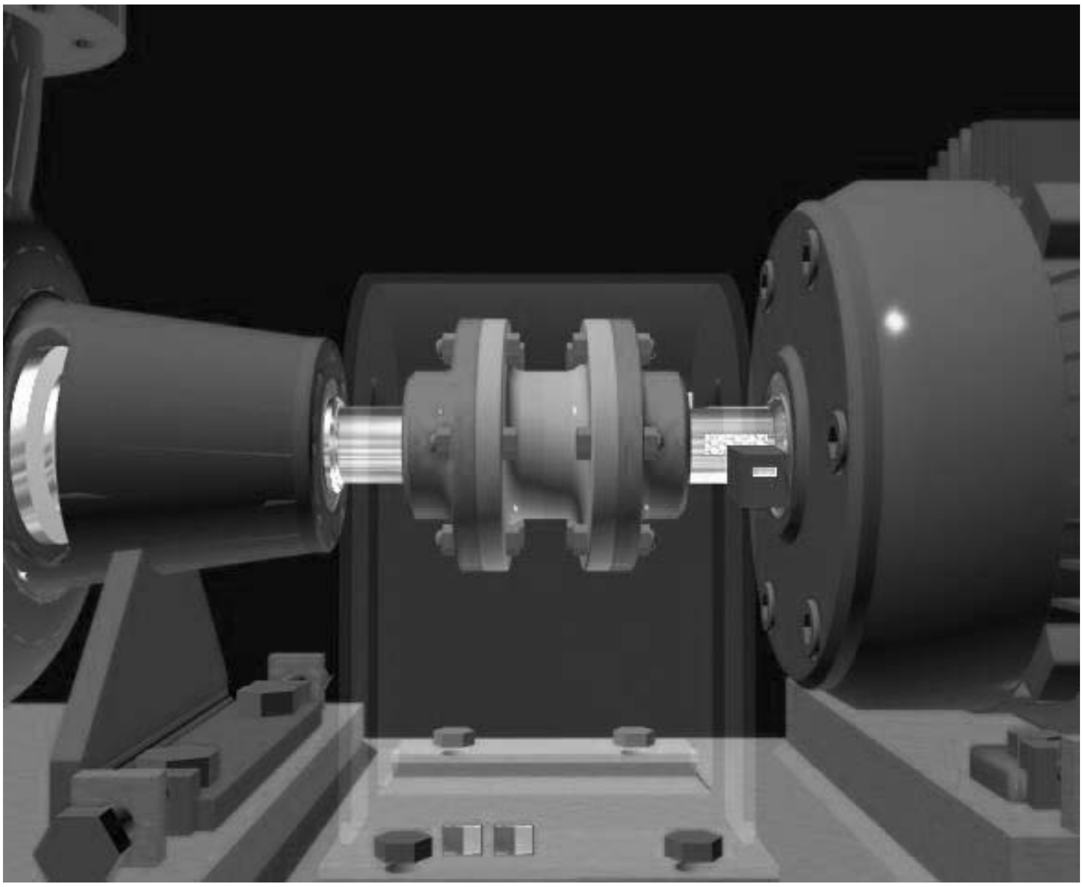


Figure 4-8

The result is a voltage signal that provides a "TTL" pulse once-per-revolution.

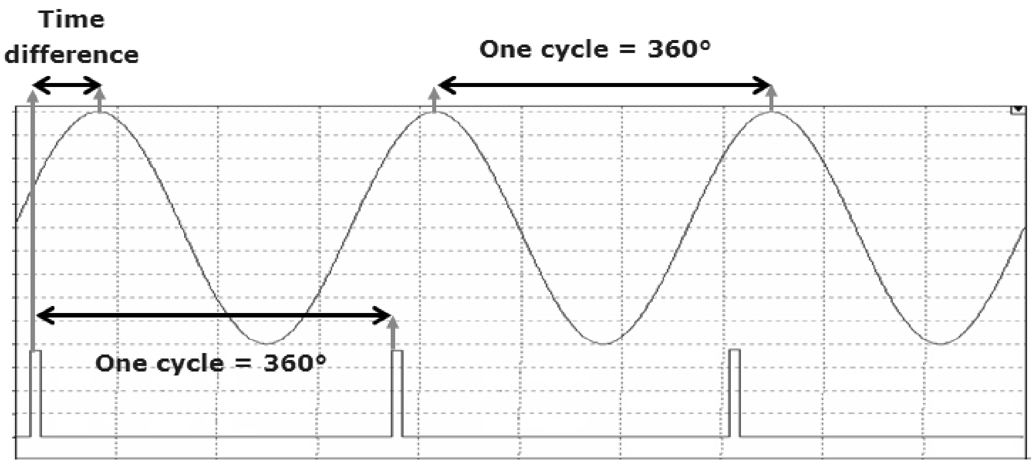


Figure 4-9

The time between pulses is the period of the machine speed. If the machine was rotating at 1500 RPM, which is 25 Hz, the time between the pulses would be 0.04 seconds (1/25 = 0.04).

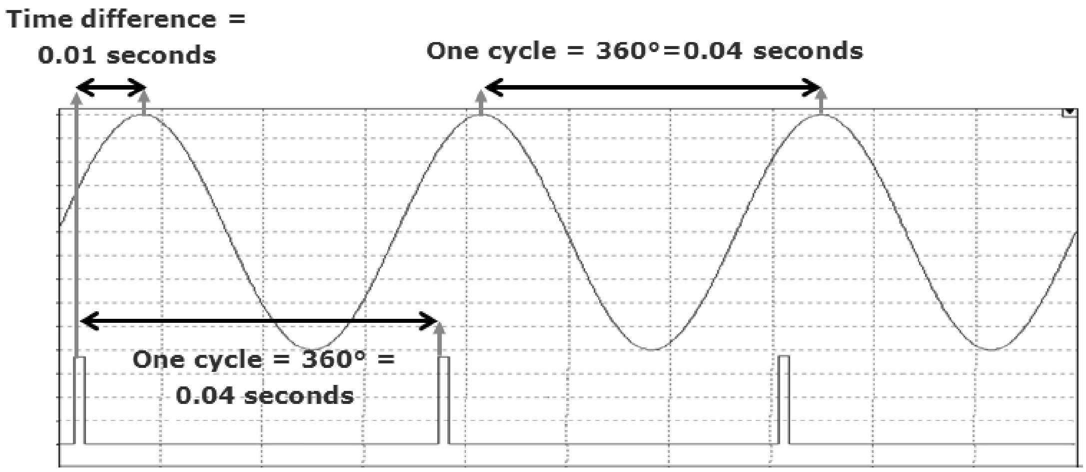


Figure 4-10

As before, we can compare the vibration from the machine to the reference tach signal. The time between pulses is 0.04 seconds, and the time between the peaks of the wave would be 0.04 seconds. If there is 0.01 seconds between the pulse and the peak of the wave, then the phase difference would be 90°. ¼ of 0.04 seconds is 0.01 seconds. ¼ of 360° is 90°.

Fortunately the data collector has the electronics and software necessary to utilize tach signals or signals from accelerometers in order to determine the phase angle.

Leading and lagging phase

When discussing phase there is a pair of important terms we need to understand: leading phase and lagging phase. This term describes the relationship of one signal or event to another. If one event occurs before another event, then it "leads". Looking at this tachometer reference signal and the waveform, does the vibration signal lead or lag the tachometer signal?

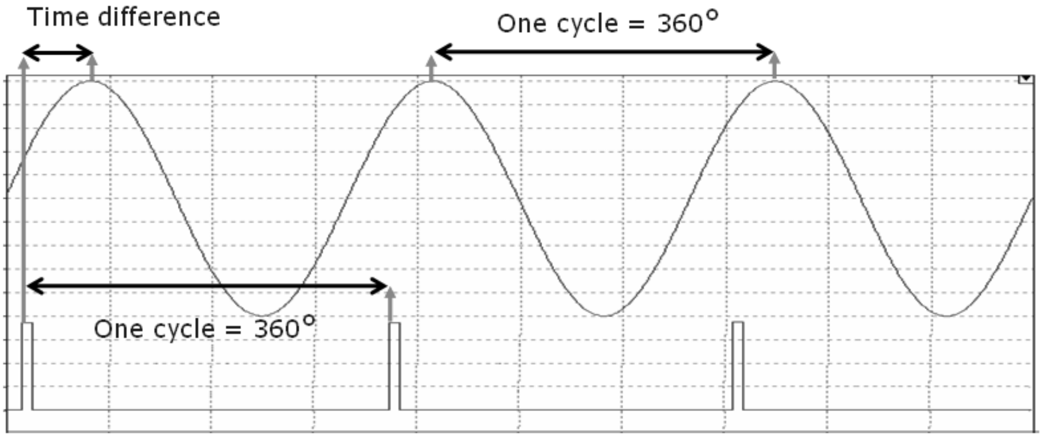


Figure 4-11

The vibration signal is lagging the tachometer signal. It may look like it is leading because it is in front of the tachometer signal. If it were a race it would be leading. But if you look at them in terms of the timing of the events, the tach pulse occurs a quarter of a rotation before the vibration signal reaches its peak.

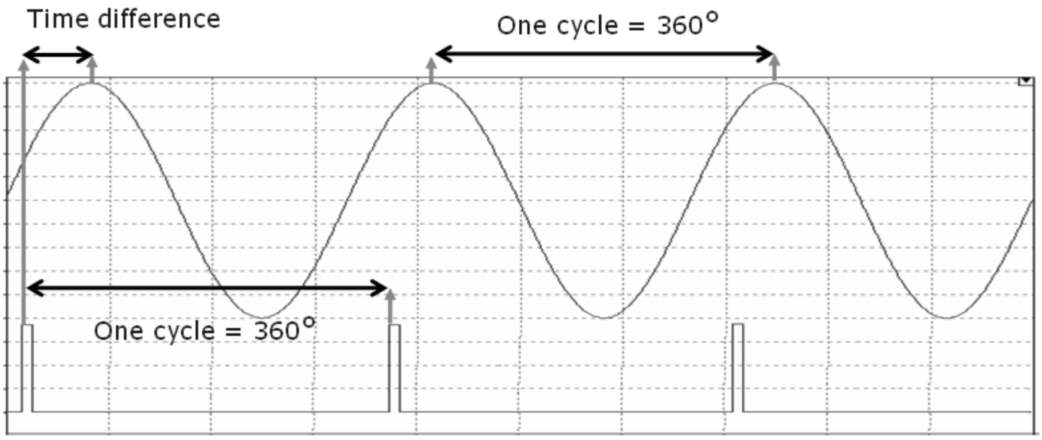


Figure 4-12

Let's have a closer look at the tachometer signal and the vibration signal. We would expect that as the speed of the machine is changed the phase relationship would not change.

We can vary the phase angle between the tach signal and the heavy spot on the rotor and we can see how the two signals shift past each other. We can see here that the vibration signal is leading when the peak occurs before the tach signal.

The effect of the type of transducer

Thus far we have assumed that the sensor was a displacement probe – i.e. it measured displacement. The probe is measuring the distance between the tip of the probe and the shaft. It is therefore sensitive to the position of the "high spot" – i.e. the point at which the shaft is closest to the probe.

Before we explore the term "high spot", let's first look at two other issues – the position of the transducer in relation to the phase reference, and the type of transducer.

In our example thus far the tach reference has been in line with the tachometer. In our virtual machine, they may be positioned as follows:

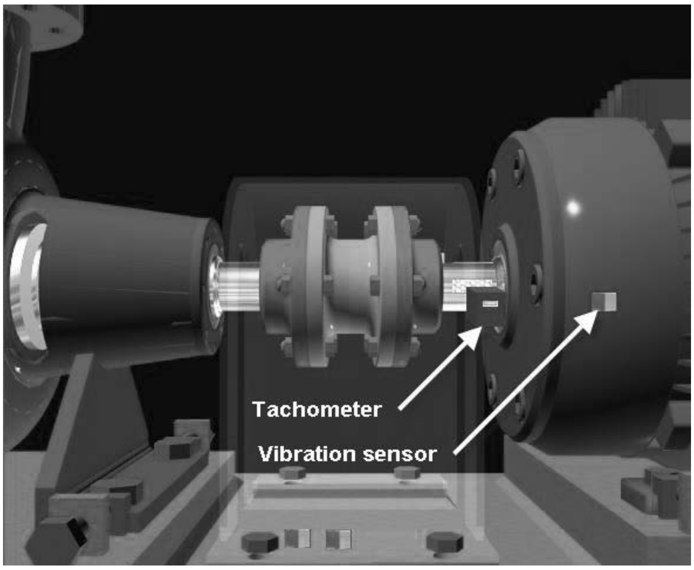


Figure 4-13

If they are in-line with each other, and the high spot happened to be located at the same angular position as the tachometer, then the phase angle reading would be zero degrees. There is no angle between the tachometer and the vibration signal. (This may well be the case if the source of out-of-balance vibration was a keyway and it was not cut correctly.)

But what if the vibration sensor was not located in line with the tachometer? What if the sensor were located 90 degrees (in the direction of rotation) around from the tachometer?

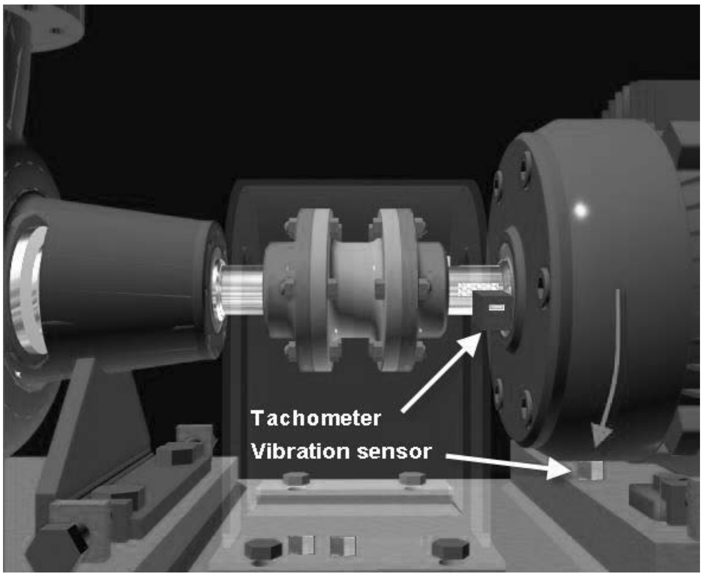


Figure 4-14

The phase angle would be different. The phase angle would be 90 degrees. The tach signal would lead the vibration signal by 90 degrees because the high spot would pass the tachometer a quarter-rotation before it reaches the point where the sensor is located.

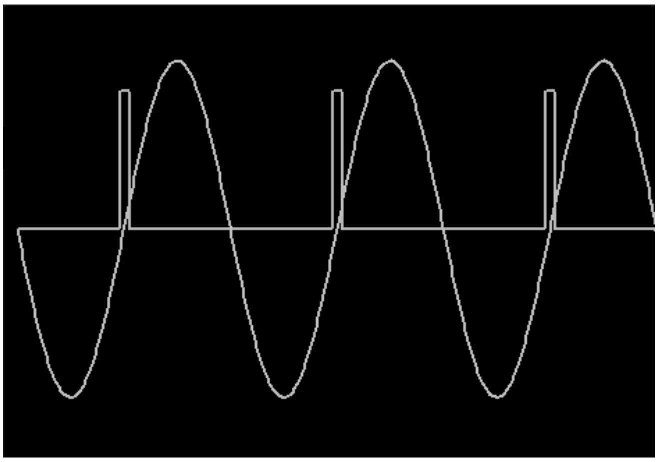


Figure 4-15

This is one very good reason why you should mount the sensor in line with the tachometer (or whatever you are using as your phase reference).

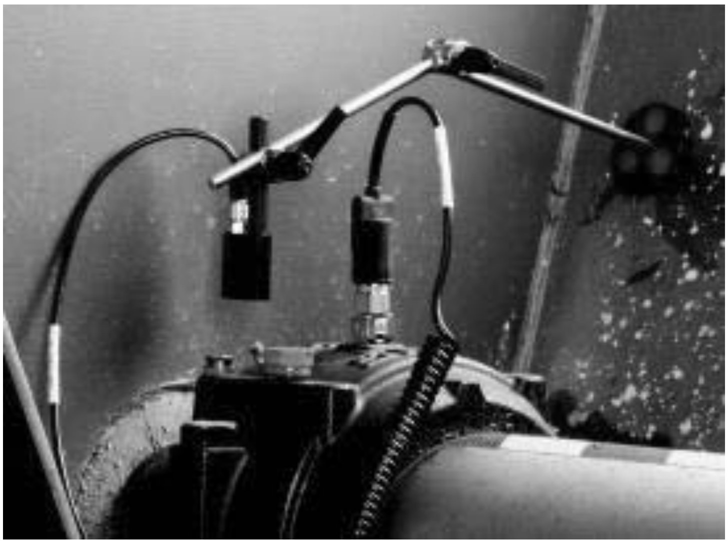


Figure 4-16

Vibration sensors

What would happen if we did not use a displacement probe? Would the phase reading be affected? Yes it would. Thus far we have used a displacement probe. As the high spot passed the displacement probe the sensor indicated the increase in vibration level. In the case of an out-of-balance shaft, the vibration pattern would be sinusoidal. And if the high spot happened to be in-line with the tachometer reference (our reflective tape), the phase reading would be zero degrees.

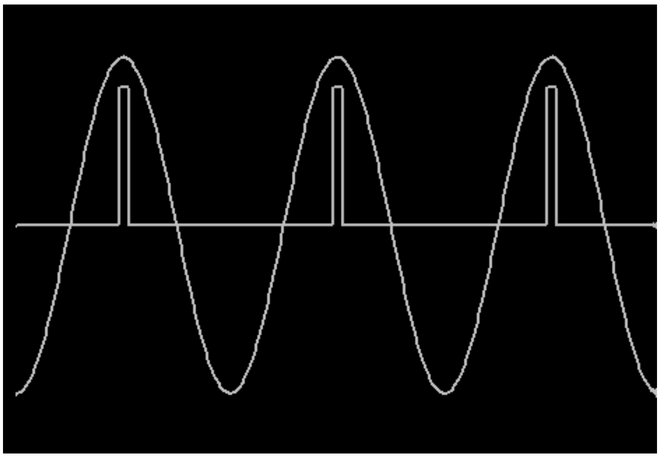


Figure 4-17

We learned earlier that the point of maximum displacement occurs at a different point in the cycle to the maximum acceleration and velocity:

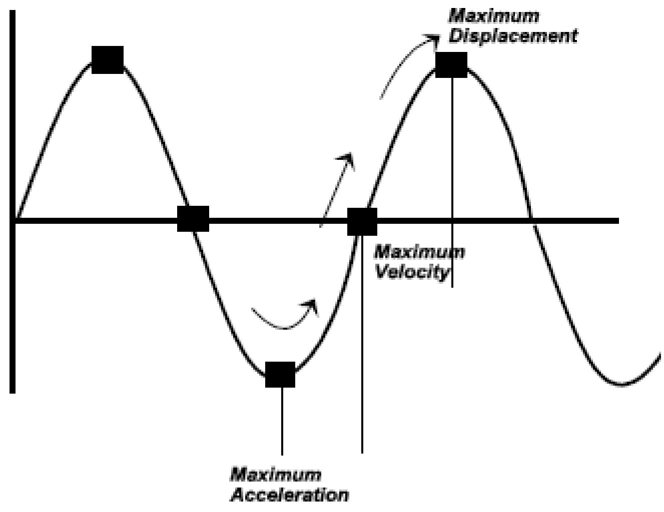


Figure 4-18

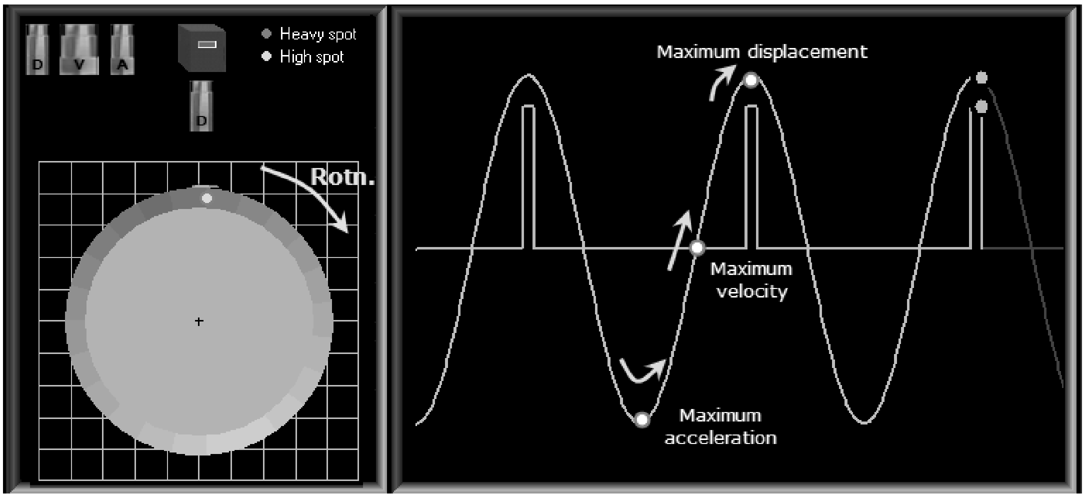


Figure 4-19

But now we will replace our displacement probe with a velocimeter. For the moment we will keep it simple and assume that the sensor purely measures the instantaneous velocity of the vibration and does not introduce any phase shifts.

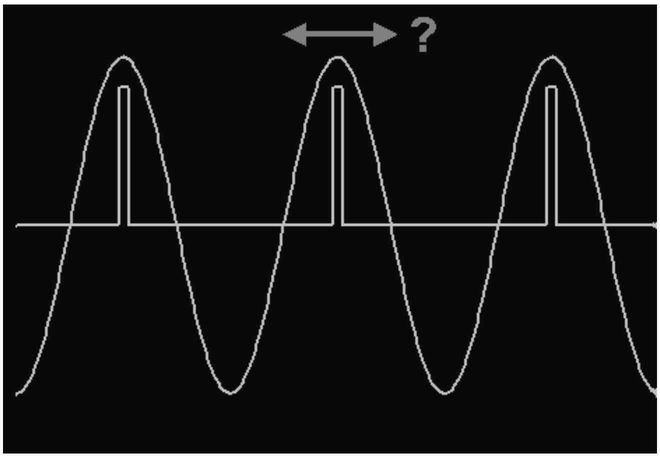


Figure 4-20

We do see a change in the reading. It now reads -90 degrees. The vibration waveform leads the reference tachometer signal by 90 degrees. We learned in the fundamentals that there is a phase relationship between displacement, velocity and acceleration. Velocity leads displacement by 90 degrees, so we will see that change in our reading.

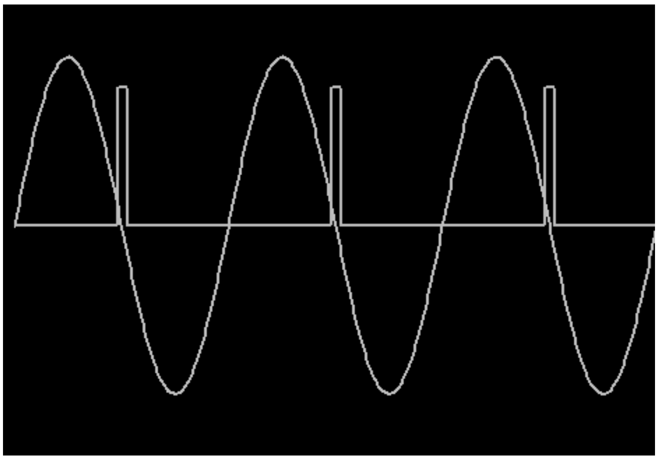


Figure 4-21

If we use a velocity sensor to measure vibration we could easily think that the heavy spot leads the reference spot by 90°. In this graphic you can see the actual position of the heavy spot (in-line with the reflective tape) and the apparent position of the high spot. We see this because at that point in the rotation of the shaft the velocity is at its greatest.

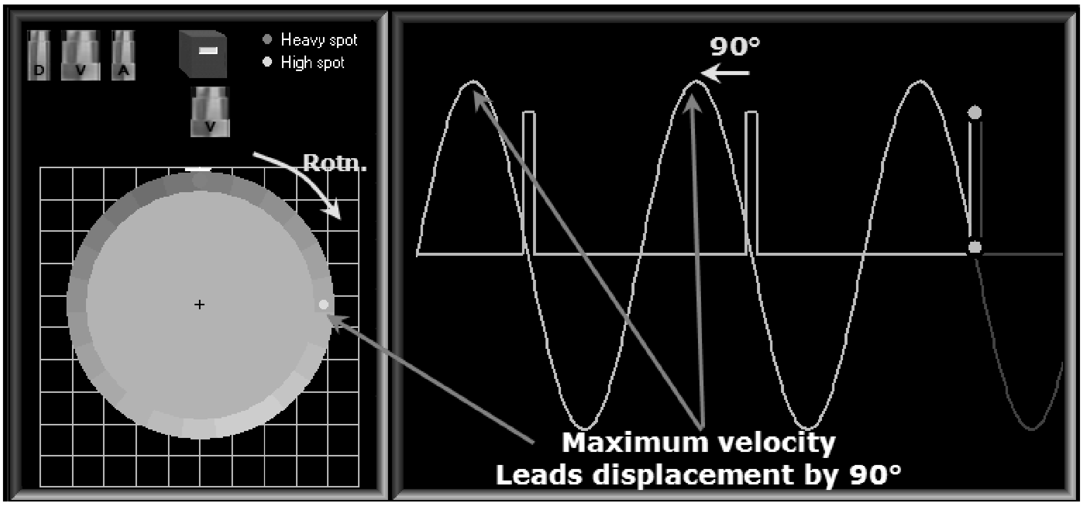


Figure 4-22

What if we now used an accelerometer? The phase is shifted by a further 90 degrees. It now reads -180 degrees. Acceleration leads displacement by 180 degrees.

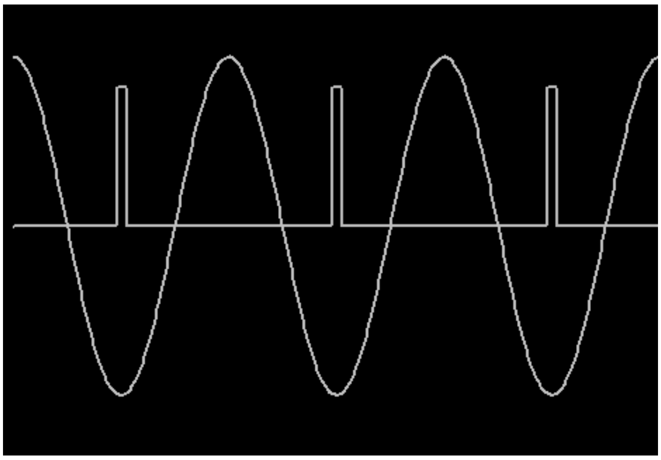


Figure 4-23

If we use an accelerometer to measure vibration we could easily think that the heavy spot is on the opposite side of the shaft to the phase reference (it leads by 180°). In this graphic you can see the actual position of the heavy spot (in-line with the reflective tape) and the apparent position of the high spot. We see this because at that point in the rotation of the shaft the acceleration is at its greatest.

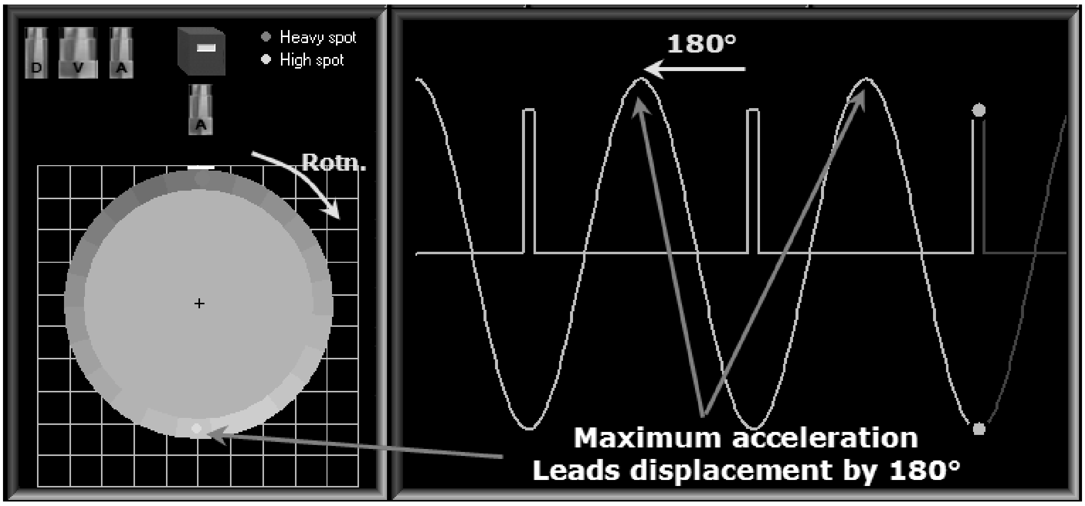


Figure 4-24

So, the vibration transducer used does affect the phase reading. But do we care? If we are performing vibration analysis or we are studying resonances, and all we are trying to understand is how one part of the machine is changing in relation to another part of the machine, then we actually do not care. We'll learn more about this later, but in brief, we don't care because all of these phase shifts do not change as long as the speed of the machine and the instrumentation do not change. We are not interested in absolute phase, just relative phase.

However, if we are performing machine balancing then we do care about the sensor; for two reasons. The first reason, which is a little off the topic, relates to sensitivity. A displacement probe measures the amount of displacement, which may be quite small. An accelerometer measures the vibration which is proportional to force (remember Force = mass x acceleration).

The more relevant reason right now relates to the selection of the position of the trial weight. You see, if we were to take a vibration reading (the "original run") and get a phase reading of 30 degrees, then we could potentially use that information to select the appropriate position for the balance correction weight. If the angle was 30 degrees (and we are using a displacement (proximity) probe), then we might place the correction weight opposite that point; at 210 degrees. If we used a velocity probe then we have to adjust the reading by 90 degrees first. And if we used an accelerometer then we either have to adjust the reading by 180. Is that correct?

Unfortunately, it is not necessarily correct. We have made an assumption about our readings which is not necessarily valid. We assumed that the reading from the displacement probe of the "high spot" is the same as the "heavy spot". When we are balancing a rotor we are trying to counteract the effect of the heavy spot, not the high spot. Unfortunately, they may not be the same.

"Heavy spot" versus "high spot"

Let's just step back a moment and make sure we understand the "heavy spot" and "high spot". The high spot has been defined as "the angular location on the shaft directly under the vibration transducer at the point of closest proximity." The displacement probe is sensitive to the high spot. The heavy spot has been defined as "the angular location of the imbalance vector at a specific lateral location on a shaft." As we will learn, they are not the same angular location on the rotor.

As the shaft rotates it responds to the centrifugal forces. This force can be represented as a vector that rotates at the running speed of the machine (the turning speed of the rotor). The unbalance force results in a measurable displacement and a force that is transmitted through the bearing housing which may be measured with a velocity sensor or an accelerometer. But now we need to take a closer look at what we actually measure versus the parameter we are trying to measure.

If you look at the output of the displacement sensor, the signal is a sinusoidal wave (once it has been filtered and slow-roll compensated). The highest point in the waveform is the "high spot". If we are measuring the vibration on a turbine, or paper machine roll, or another "flexible" rotor, then we have to be sensitive to where the machine is running in relation to the critical frequency.

At low speed (at a speed below approximately 50% of the first natural frequency, the shaft operates as a rigid rotor. It is responding to the radial centrifugal forces. In this case the angle of the "high spot" and "heavy spot" will be quite close. However, as the speed of the shaft increases, two things will happen. Depending upon the amount of damping, the vibration levels will increase significantly (remember, they will always increase in proportion to the square of the speed), but the phase angle between the high spot and heavy spot will increase. The high spot will lag behind the heavy spot.

At the resonant frequency (critical speed), the phase shift is 90 degrees. As the speed increases, the phase lag will increase to 180 degrees. The motion of the rotation changes above the critical speed. Now the shaft will attempt to rotate around the center of mass.

Most pumps, fans, motors and compressors that run at speeds up to 3600 RPM (60 Hz) are rigid rotors. So if a displacement (proximity) probe were used to measure the vibration, the high spot will not lag far behind the heavy spot, and you can assume that a trial weight should be placed opposite the phase reading. That is, if we read 30 degrees, we could place the mass at 210 degrees.

But what if we are using a velocity sensor, or more likely an accelerometer? Can we simply adjust the phase reading as described above? No, we cannot. You see, although we are not concerned about phase lag due to the shaft resonance, we have to deal with a mechanical timing lab, and lags in the sensor and electronics.

Mechanical lag

The centrifugal forces must be transmitted through the bearings, into the structure, and up to the vibration sensor. That takes time. The vibration travels at the speed of sound, which varies according to the medium. Although the time for the vibration to travel to the sensor is not great (and can be estimated), in proportion to the time it takes for the shaft to rotate, it could represent a few degrees.

Sensor lag

The vibration then must activate" the sensing element (crystal in the accelerometer for example), and there may be additional phase lags introduced at this stage.

Electronics

The signal being measured must then pass through various electronic components which can also introduce phase lags and phase shifts. These can be considerable.

The bottom line is that the phase lag between the actual heavy spot and the high spot recorded in the analysis system can be considerable, especially when accelerometers are involved.

How important is this phase lag?

At this level of your training it is important that you understand the information we have just covered. You should understand the relationship between the recorded phase measurement and the actual heavy spot on the shaft. Remember, even if we are not performing machine balancing you will often be measuring phase at the running speed which is most likely due to the residual imbalance or misalignment.

However, it is true to say that whether you are performing balancing, or you are trying to understand the motion of the machine, you will be taking multiple phase readings and making comparisons. In the case of balancing you will be looking at the change in phase angle as a result of the addition of a trial weight. In other situations you will be using phase to see how one point on the machine or structure is moving in relation to another point. We have now seen that as long as we do not change the speed of the machine we can safely perform these comparisons.

Vector representation

Phase data can be used in a number of ways. A common way to represent the phase and amplitude readings is via a vector diagram. There are numerous applications for vector diagrams, but they are very popular when balancing a machine without balancing software.

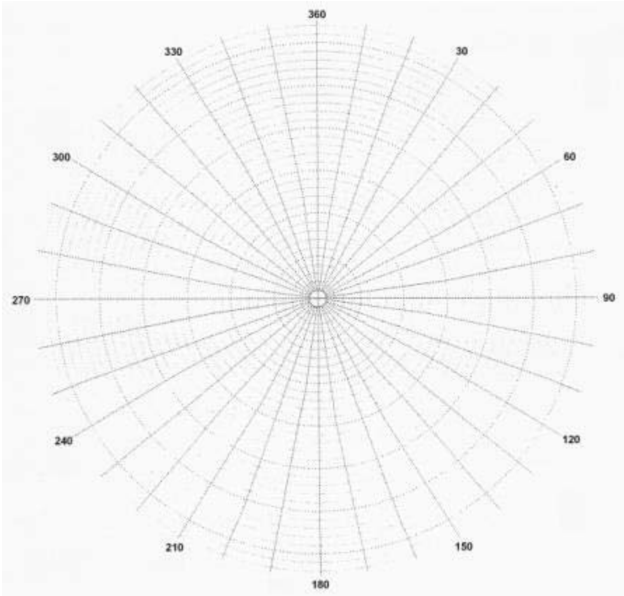


Figure 4-25

The distance from the center is amplitude, and the angle is phase angle. In the example, we have specified that each ring is 1 mil of vibration, and the reading was 5.5 mils at 45°.

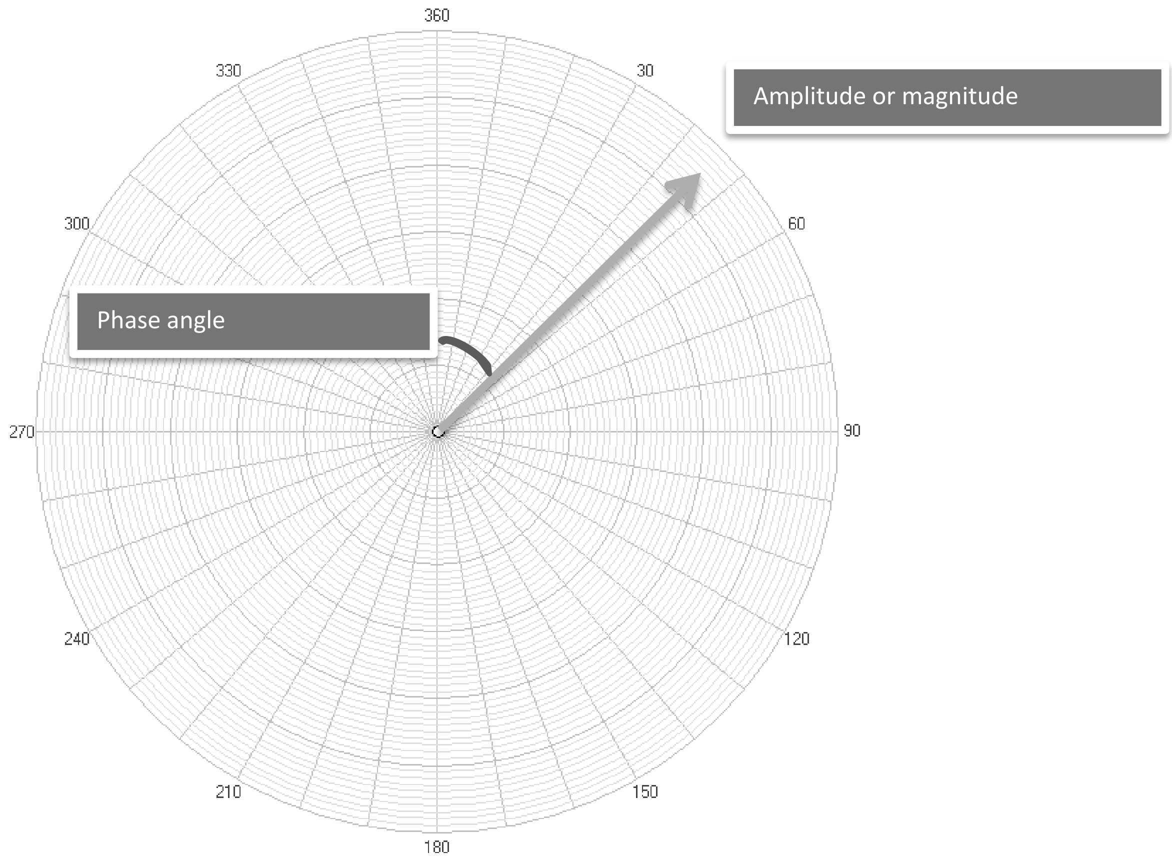


Figure 4-26

We will discuss vectors and polar plots in greater detail in the balancing section.

Phase convention

It is very important that you understand the phase convention being used. You have to consider your point of reference: the phase angle relative to the tach reference (photocell or displacement probe) or the phase angle relative to a fixed point on the machine.

Most vibration data collectors report "true phase" – the phase angle ***increases opposite to the direction of rotation*** from the reference mark. This is known as the "rotating protractor" convention. As the shaft rotates (counter clockwise in this example) you can visualize the protractor also rotating. The tach reference is watching from above, and the angles appear to increase as the shaft rotates (from the perspective of the photocell/displacement probe).

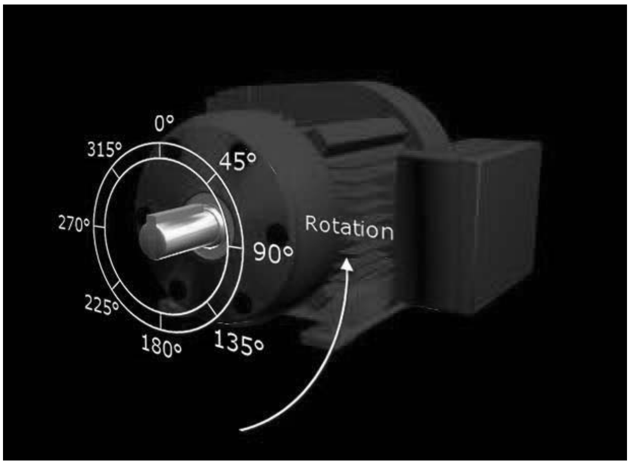


Figure 4-27

If you use a strobe to measure phase, you are observing how the keyway (or some other mark on the shaft) moves as the vibration sensor is moved (or how it moves during the balancing procedure). In this case you use the "fixed protractor" convention. This method counts ***increasing phase angles in the direction of rotation***.

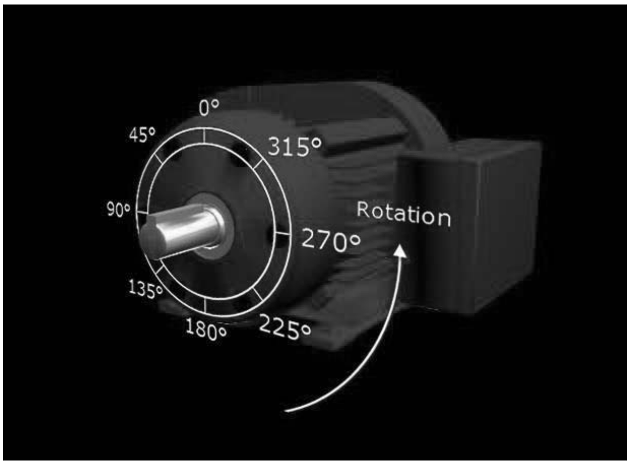


Figure 4-28

Representing phase

Thus far we have always discussed phase readings as simple numbers: 0°, 45°, 180°, 360°, etc. We have also seen how we can represent the amplitude and phase in a vector plot. In many cases we are not concerned about the absolute phase number, we are interested in relative phase – how the phase at one point compares to the phase at another. In a moment we will learn more about the applications of phase; however no matter how we use it, we have the option of the numerical value itself, or a visual representation.

In this illustration you can see how convenient it is to represent the phase angle visually. By drawing a circle and a tail at the desired angle, it is easy to quickly determine the angle, and the relative movement, with a quick glance.

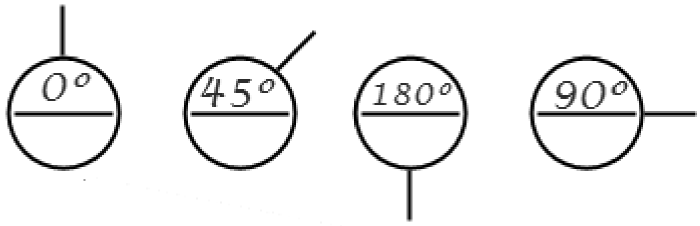


Figure 4-29

You don't even need to write down the phase angle - you can just draw the tails; either inside or outside the circle. You can easily see that these two readings are 180° out of phase.

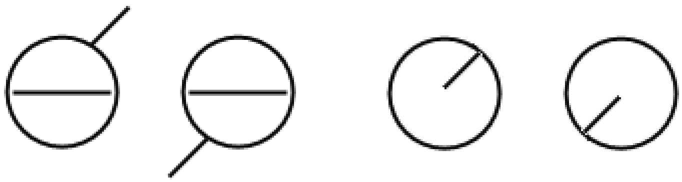


Figure 4-30

This data can be used in a number of ways, but one common method is called the *bubble diagram* (developed by Ralph T. Buscarello). You can take readings around the machine and enter them into the diagram, adding the tails according to the angle. We will discuss the use of this diagram in greater detail in the phase analysis section.

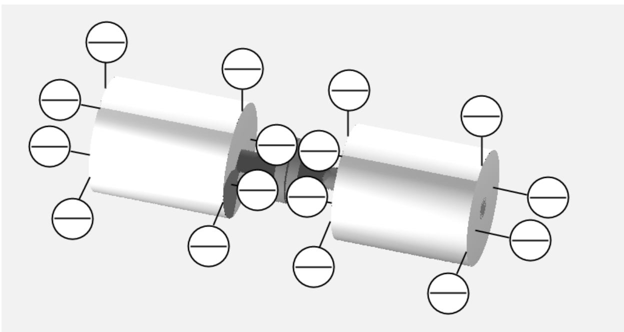


Figure 4-31

Or for just a single component:

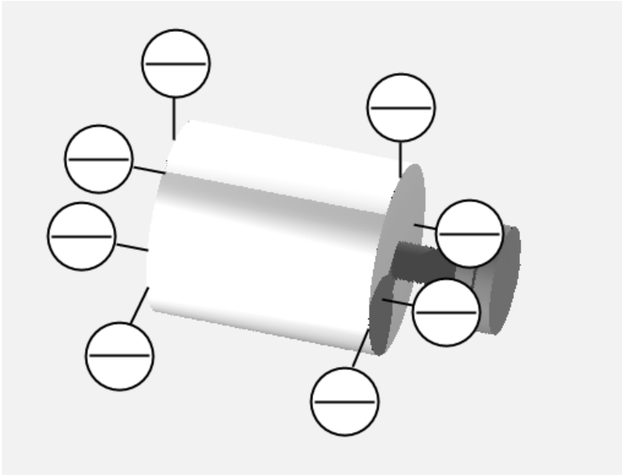


Figure 4-32

You can then visualize how different points on the machine are moving relative to one another.

Measuring phase

Let's take a closer look at how we measure phase. In the previous section we described two basic methods: using a tachometer reference, and using the vibration from another sensor. There is also a variation on those methods that utilize a strobe, but we'll get to that later.

Using a tachometer

There are a number of ways to obtain a once-per-revolution tachometer signal. The most common involves the use of reflective tape and an optical photo-tach.

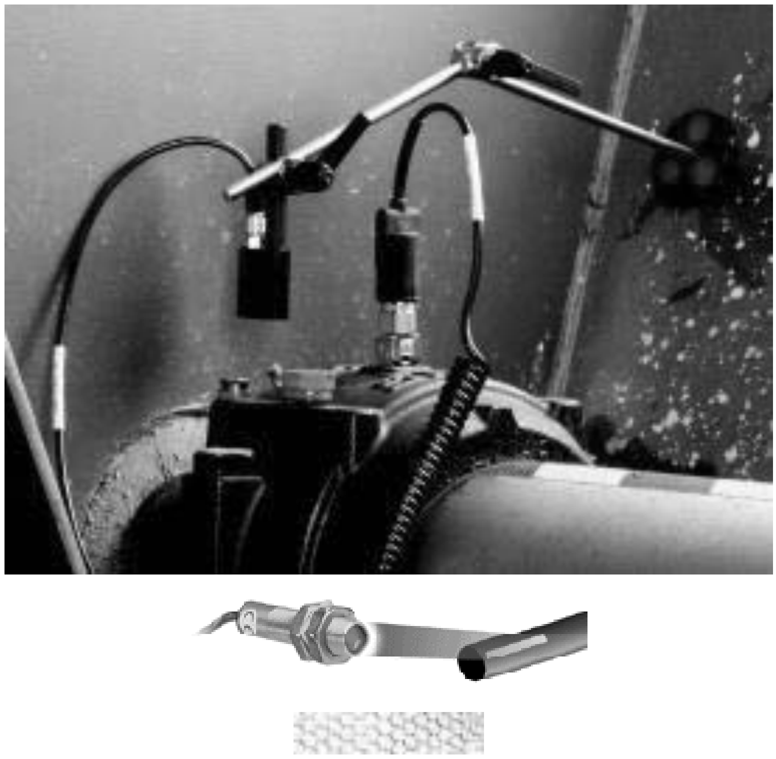


Figure 4-33

There are a number of products available that can use reflected light, including laser light, to generate the tach signals. Many will work without reflective tape, as long as there is an area of high contrast – for example, a paint spot.

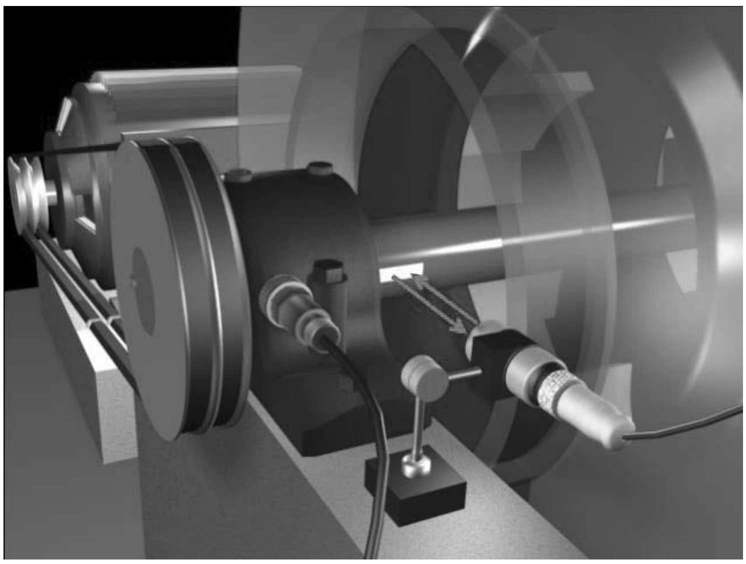


Figure 4-34

The photocell shines a light (visible or laser) on to the shaft. Due to the surface texture and color, the light does not reflect. When the tape passes underneath, the light reflects. The tachometer generates a TTL signal that is fed into the data collector.

As mentioned, another way is to use a displacement probe which is aimed at a keyway or setscrew. The change in displacement provides the step in voltage used as the reference. This is often called a keyphasor.

The output from the tachometer is fed into the tachometer input of the data collector. You will need to refer to your data collector to understand where to connect the tachometer signal and how to use it to collect phase readings.

Figure 4-35

The data collector is then able to use the tachometer signal to determine the speed of the machine, and to compare the vibration at the running speed from a vibration sensor to the tachometer signal. It will then provide a phase angle of between 0° and 360° (in some cases the data collector may provide a reading of -180° to +180°).

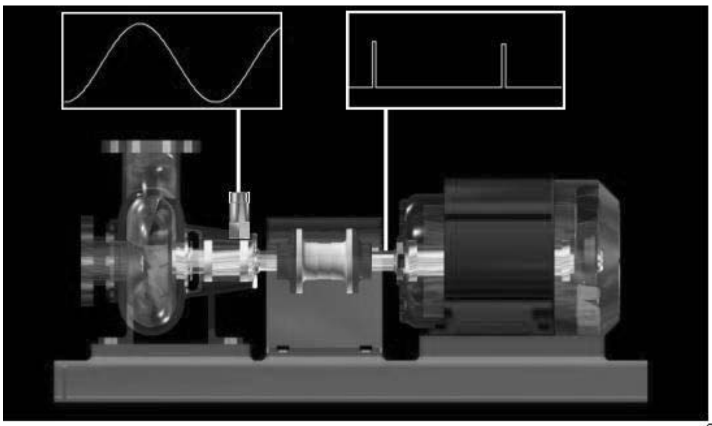


Figure 4-36

The data collector can determine the phase angle in a number of different ways. It can apply the two-channel method that will be discussed next, or it can use the tachometer to trigger the data acquisition process and acquire the phase angle from the FFT process. Triggering and data acquisition are described elsewhere.

Two channel phase

Did you know that when your data collector takes a measurement on a machine and creates the FFT (spectrum), it actually computes the magnitude (amplitude) spectrum and phase spectrum? But because you do not have a reference signal (the data collector starts sampling when you press the button, not according to any pre-defined reference on the shaft) the phase data does not have a lot of value. So it is discarded and we only keep the magnitude spectrum.

However, there are two possibilities available to us. If the data collection was synchronized to the tach reference, the phase data would be relevant. We could look at the phase at the running speed and use that information. This is one of the ways that data collectors measure phase when using the tachometer. But there is another way.

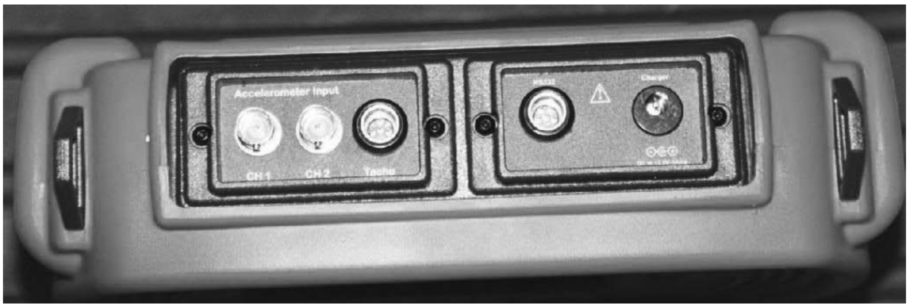


Figure 4-37

If we connected one accelerometer to one channel of a two channel data collector, and we connect another sensor to the second channel, the data collector can sample them simultaneously (this is essential) and compare the phase spectra. We would place one sensor at a reference location, and the second sensor at the point of interest. We can also move that sensor around to see how the phase angle changes (while leaving the reference sensor in the same location the whole time).

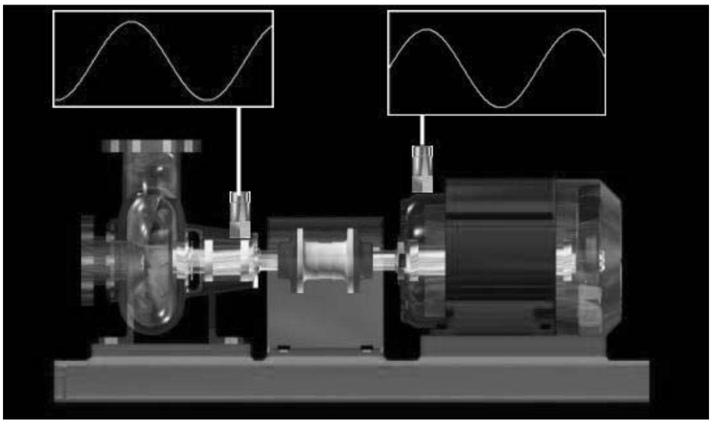


Figure 4-38

We might see that the channel one phase value at the running speed was 33°, but what would that tell us? Nothing. But if the phase reading at running speed from the second channel was 113°, we can then compare the two values to see that there is a 180° phase difference between them. We can see that the vibration from the two points is 180° out of phase. That means one point is moving upwards and the other point is moving downwards. And that opens all kinds of possibilities to us.

Using a strobe

Strobe as a tachometer

There are two ways that we can use a stroboscope. If we tune the strobe to the running speed so that the shaft or coupling appears to have stopped rotating, the output of the strobe can be fed into the tachometer input of the data collector. The data collector would treat the signal from as if it were a normal tachometer input.

However, if the machine speed varies slightly, the signal from the strobe will no longer represent the exact speed of the machine – the phase reading will be inaccurate. If you set up the strobe so as to freeze a keyway, setscrew or some other point on the shaft or coupling, then you should use that as your reference before you record the amplitude and phase reading. If the speed varies then you will see the keyway/setscrew begin to rotate forward or backwards.

Data collector driving the strobe or vice versa

There is another way to use a strobe that is very effective, however not all strobes or data collectors have this capability.

The vibration sensor is connected to the strobe and it is placed in "EXT" mode. You control the flash rate of the strobe until you freeze the motion of the shaft. Switch to "LOCK/TRACK" and the strobe will now use its internal circuitry to filter the vibration signal and extract the vibration at the running speed. The strobe can now track any slight changes in speed. The strobe will typically have a TTL output signal that can be used to drive a data collector for balancing purposes or triggered data collection.

Figure 4-39

Alternatively, some data collectors can be used to track the dominant running vibration and drive the strobe. The data collector can either track the 1X speed, or the user can move the cursor on a spectrum to set the speed. A cable is connected from the output of the data collector to the input of the strobe. The flash rate of the strobe is now under the control of the data collector.

When the strobe or data collector is set to track the running speed, you can then perform visual phase measurements. The strobe will flash at the running speed of the machine, thus the shaft (or coupling) will appear to freeze. (Of course, you must be very careful – the shaft has not stopped and you must be careful not to touch it.) You should then set a visual reference, like a keyway or setscrew, and use the "Relative Phase" knob on the strobe to adjust the keyway/setscrew so that it is at the 12:00 position.

If you watch the shaft/coupling while you move the accelerometer, it will appear as if the shaft/coupling rotates. The amount of rotation is dictated by the phase difference between the original sensor position and the new position. For example, if the machine was out of balance and you move the accelerometer 90°, the shaft/coupling will appear to rotate 90° (a quarter turn).

This is a very effective phase analysis method. As you move the sensor around you can see how the phase changes without even looking at actual phase values. It is best if you can use a setscrew, keyway or reflective tape as your visual reference. You should start by adjusting the strobe so that the reference is at the top of the shaft. As you move the sensor it is very easy to note the change in phase.

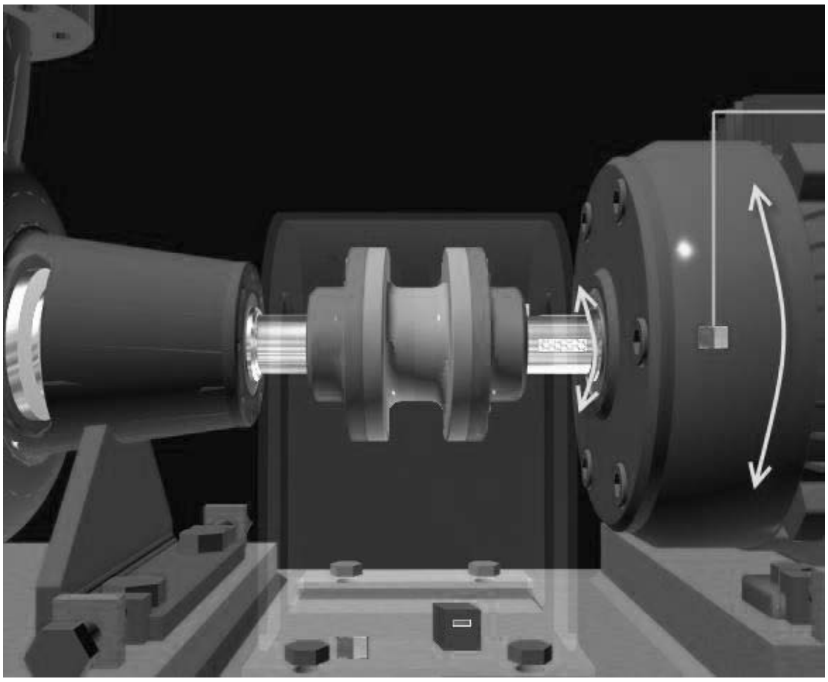


Figure 4-40

Applications of phase analysis

There are a number of ways that you can put phase data to use. There are basically three main vibration applications: balancing, structural resonance analysis, and fault diagnosis. We will discuss balancing and resonance analysis separately. For now we will focus on fault diagnosis.

Machine fault diagnosis

Phase can be used to diagnose fault conditions. There are a number of fault conditions that develop similar patterns, making it difficult to accurately diagnose a fault. While spectra and time waveforms collected in multiple axes can help you to accurately diagnose faults, if you understand the underlying forces involved, phase analysis can provide conclusive evidence as to the exact nature of the fault: unbalance, misalignment, bent shaft, eccentricity, foundation flexibility, cocked bearing, and even looseness.

We can take a number of measurements in order to understand the motion of the machine. We can take readings vertically and horizontally at each end of the machine. We can compare the amplitude and phase of vertical versus horizontal; we can compare the vertical readings at both ends of the machine, and we can compare the horizontal readings at both ends of the machine. We can also take phase readings on either side of the coupling and compare the readings in the radial directions.

Axial readings are also very important. Rather than a single reading, we can take readings on either side of the shaft; to compare the left side to the right side, and compare the top to the bottom reading. And again we can compare axial readings taken on either side of the coupling (i.e. on the motor and pump).

You might routinely collect a single axial vibration reading, but when you are collecting phase readings it is important to collect two axial readings, and in certain cases that we will discuss later, you may even collect four readings. Due to restricted access, safety issues, and machine construction, you may only take axial measurements at one end of the machine.

Precautions when collecting phase data

It is also worth revising the fact that you must be careful when comparing phase readings taken at opposite ends of a machine, or when comparing phase readings taken across a coupling. Phase readings are sensitive to direction. For example, if you were to measure phase on a solid block of steel then the phase reading on one side of the block will be 180 different to the reading taken on the opposite side of the block.

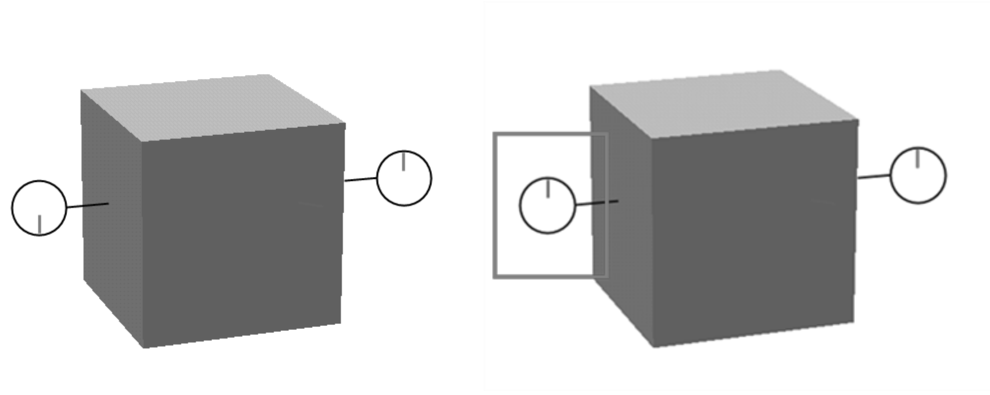


Figure 4-41

You must therefore decide upon a reference direction and stick to it. If you decide that your reference direction is toward the driven end of the machine, you must add 180 to any readings taken with the sensor facing the driver of the machine.

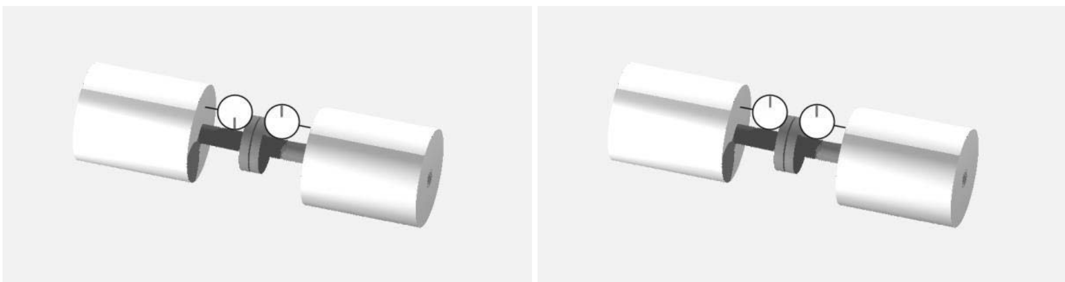


Figure 4-42

This is especially important if you are using a strobe to measure phase and you are observing how the keyway, setscrew or some other physical object tied to the shaft moves when the sensor is moved. In one situation you may be facing east when making that observation and west when you are making a different measurement. Again, you must decide on a reference direction and adjust the other readings by 180°.

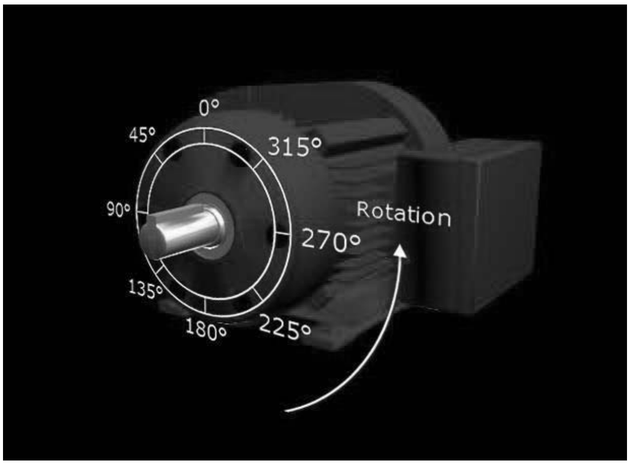


Figure 4-43

In the sections that follow we will talk about the phase relationships between certain points and axes on the machine. You may be required to observe whether the readings are in-phase (0° difference), 180° out-of-phase, or 90° out-of-phase. However, it is not expected that the difference is exactly 0°, 180° or 90°. It is generally considered that if the readings vary by 30° or less, then the rule holds. For example, if the difference between two readings was between 150° and 210°, then you can consider the readings to be 180° out-of-phase.

Diagnosing unbalance

Although considered by some to be the most common and simplest fault to diagnose, it is actually quite easy to confuse unbalance with other fault conditions. If you find a high 1X peak and assume it needs to be balanced, you may be quite wrong – and generate a lot of unnecessary work - and still not correct the fault.

We need to go back and study the motion of a rotor when it is not balanced correctly. If you understand the underlying motion you will be able to use phase data to prove that the rotor is in fact out of balance, and rule out other possibilities.

We will now quickly review the different forms of unbalance, and then look at how we can look at the end-to-end phase readings, and the vertical-to-horizontal phase readings and relative amplitude levels in order to diagnose unbalance.

Static unbalance

The simplest type of imbalance is equivalent to a heavy spot at a single point in the rotor. This is called a static imbalance because it will show up even if the rotor is not turning - if placed in frictionless bearings the rotor will turn so the heavy spot is at the lowest position.

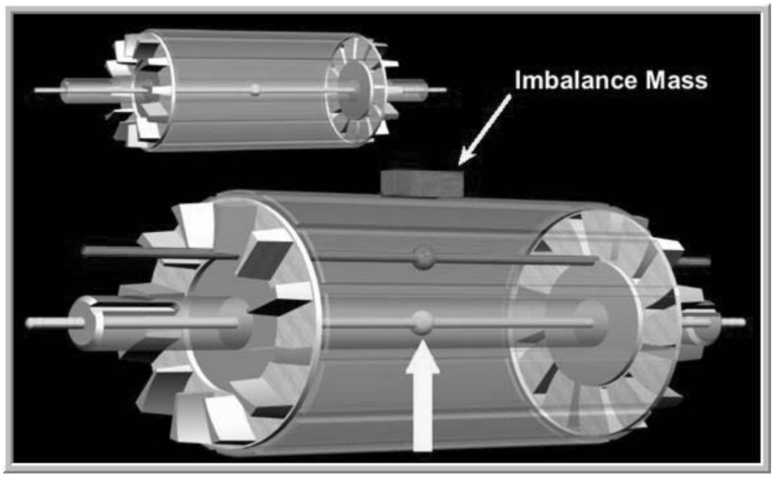


Figure 4-44

Static imbalance results in running speed (1X) rotational forces on both bearings of the rotor, and the forces on both bearings are always in the same direction. The vibration signals from them are "in phase" with each other.

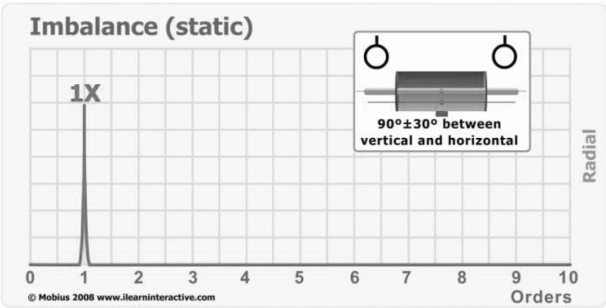


Figure 4-45

A pure static imbalance will produce a strong 1X peak in the vibration spectrum, the amplitude of which is proportional to the severity of the imbalance and the square of the RPM. The relative levels of the 1X vibration at the bearings depend on the location of the heavy spot along the rotor.

We can look at the motion of the shaft and learn five important facts:

1. Both ends of the machine should be in-phase (in the case of pure static unbalance), both in the horizontal and vertical directions. The phase difference between the two vertical readings should be very similar to the phase difference between the two horizontal readings – that is, they should be in-phase vertically and in-phase horizontally.

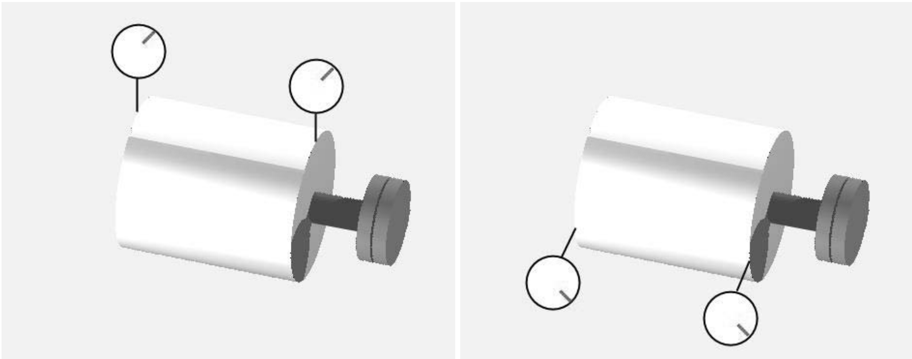


Figure 4-46

2. There should be approximately 90° between the vibration readings taken in the vertical and horizontal directions. Depending upon the direction of rotation, the horizontal measurement will either lead or lag by 90° (therefore you will measure approximately 90° or 270°). This will be true at both ends of the machine.

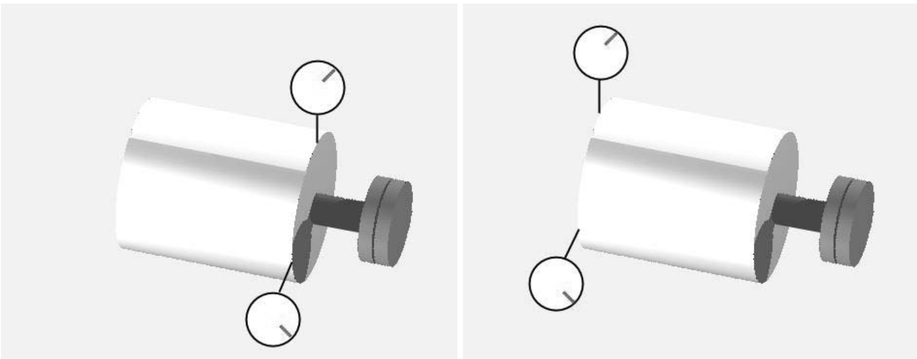


Figure 4-47

3. In theory the vertical and horizontal vibration levels should be equal. However because the machine will be less stiff in the horizontal axis, it is more likely that the measurement taken in the horizontal direction will be higher than the vertical measurement. However, if the horizontal reading was greater than twice the amplitude of the vertical then you might suspect that the machine had a looseness/foundation problem.

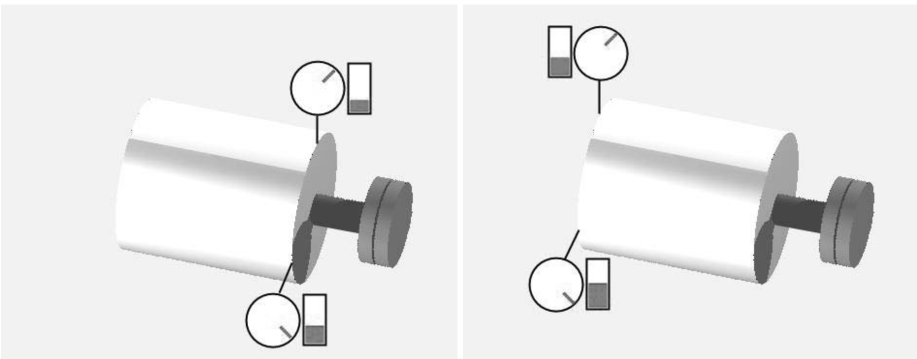


Figure 4-48

4. As stated earlier, the vibration level at one end of the machine will be higher than the vibration measured at the other end. However the ratio between the drive end and driven end should be roughly the same in the vertical and horizontal directions. For example, if the drive end was twice the amplitude in the vertical direction, you would expect the vibration at the drive end to be approximately twice as high in the horizontal direction as well.

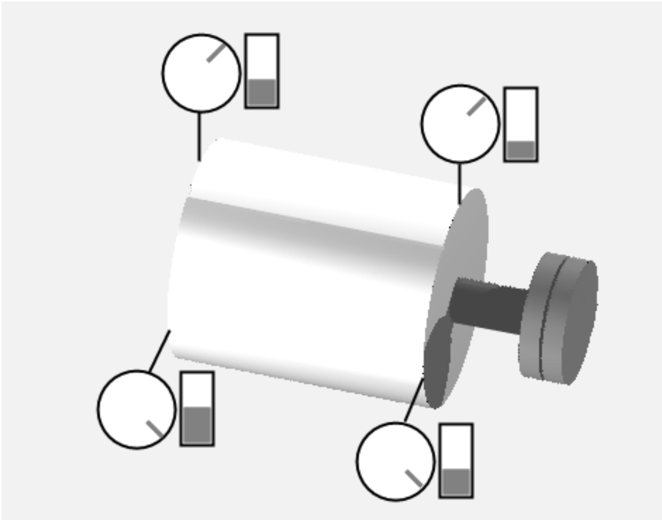


Figure 4-49

5. The axial amplitude and phase readings are also very important. If you consider the movement of the rotor/machine with pure static unbalance, the ends of the machine are moving up and down in phase. There is no rocking, so the phase readings across the shaft will be in-phase, and the amplitude levels should be very low; possibly less than one-third of the horizontal readings.

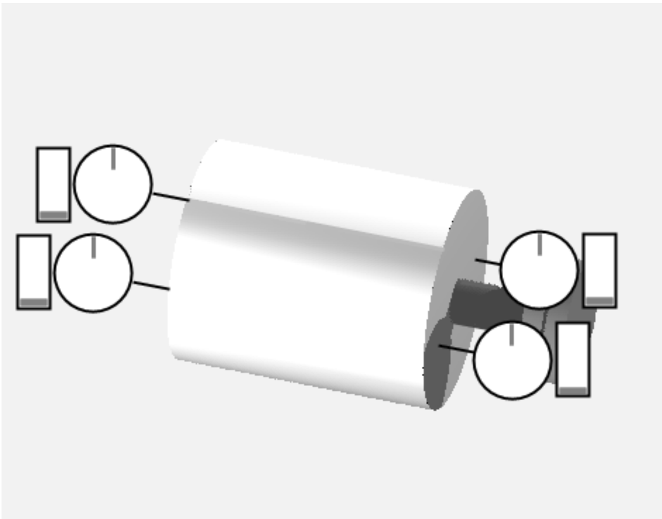


Figure 4-50

When you put all of that together you have a complete picture of the motion of the machine component.

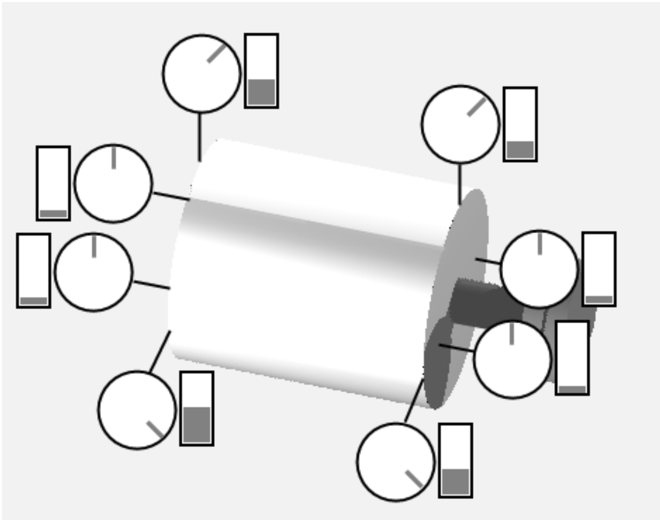


Figure 4-51

We could use our animation program to demonstrate this action. We have entered data for one component and one end of the machine.

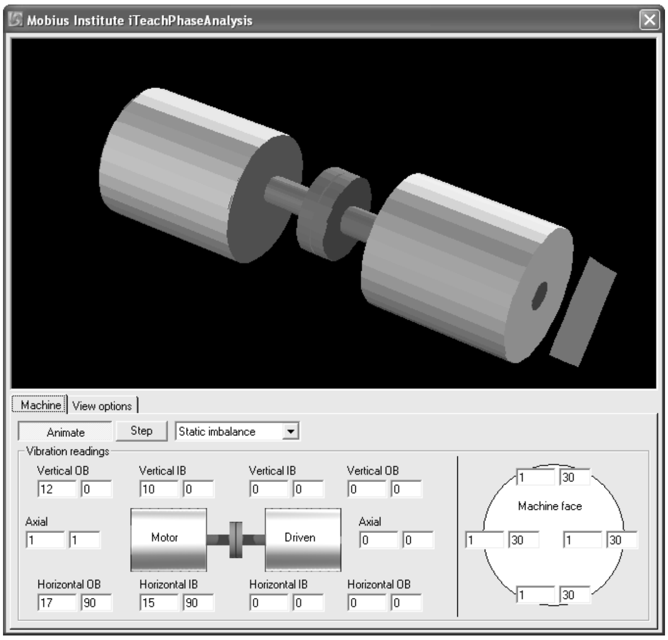


Figure 4-52

Couple unbalance

A rotor with couple imbalance may be statically balanced (it may seem to be perfectly balanced if placed in frictionless bearings). But when rotated, it will produce centrifugal forces on the bearings, and they will be of opposite phase.

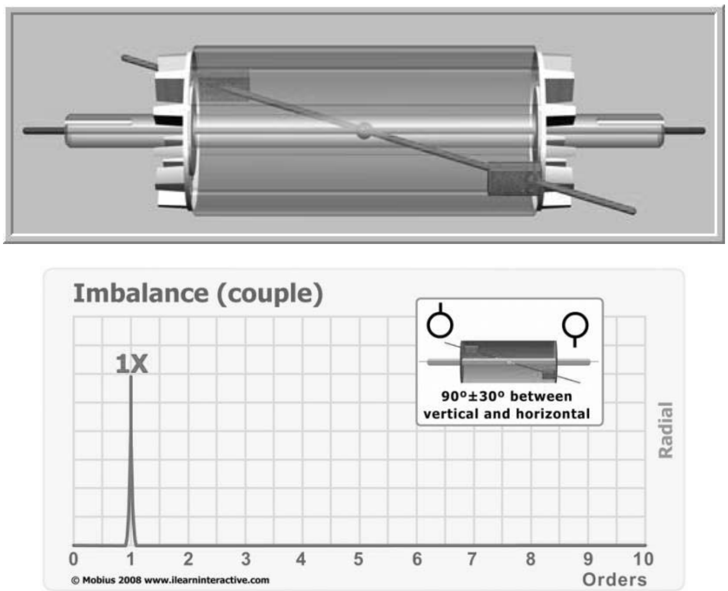


Figure 4-53

Let's revisit the four issues we considered earlier:

1. The phase relationship has changed. Both ends of the machine should be 180° out-of-phase (in the case of pure couple unbalance). You should therefore measure the phase between both ends of the machine in the vertical and horizontal direction. The phase difference between the two vertical readings should be very similar to the phase difference between the two horizontal readings.

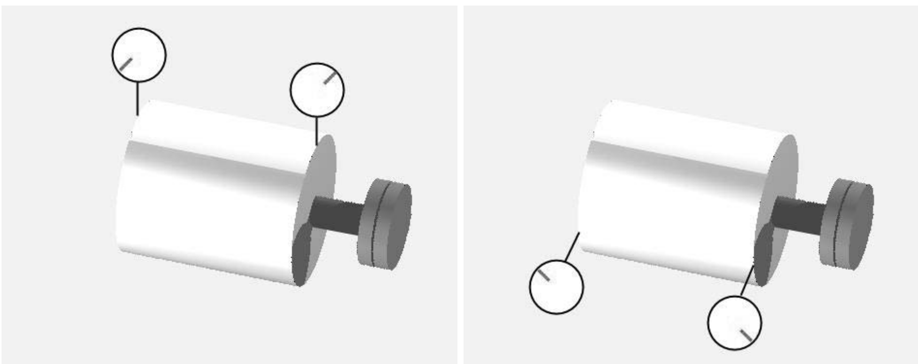


Figure 4-54

2. As we found earlier, there should be approximately 90° between the vibration readings taken in the vertical and horizontal directions. Depending upon the direction of rotation,

the horizontal measurement will either lead or lag by 90° (therefore you will measure approximately 90° or 270°). This will be true at both ends of the machine.

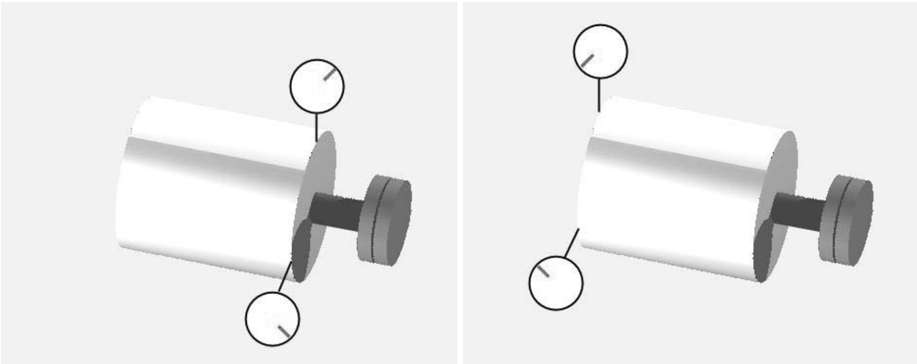


Figure 4-55

3. Once again, there will be a difference in vibration amplitude in the vertical and horizontal directions, with horizontal normally being higher due to lower stiffness. If the horizontal reading was greater than twice the vertical reading you might consider checking the foundations/base.

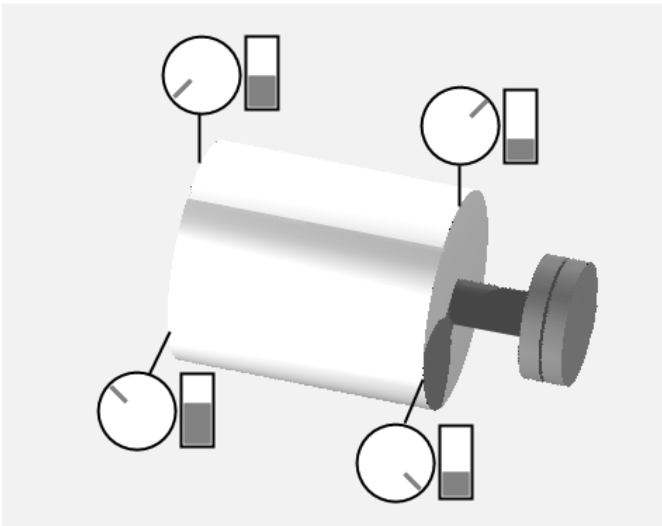


Figure 4-56

4. Once again, the vibration level at one end of the machine will be higher than the vibration measured at the other end. However the ratio between the drive end and driven end should be roughly the same in the vertical and horizontal directions. For example, if the drive end was twice the amplitude in the vertical direction, you would expect the vibration at the drive end to be approximately twice as high in the horizontal direction as well.

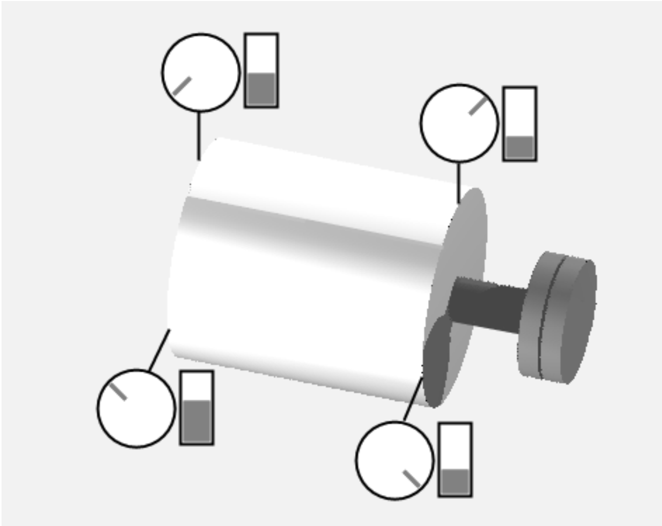


Figure 4-57

5. The axial amplitude readings will again be low compared to the radial readings – the forces are predominantly in the radial direction. However, because we now have a rocking motion, with 180 phase difference from one end of the machine to the other, the phase readings taken across the shaft will also be out-of-phase.

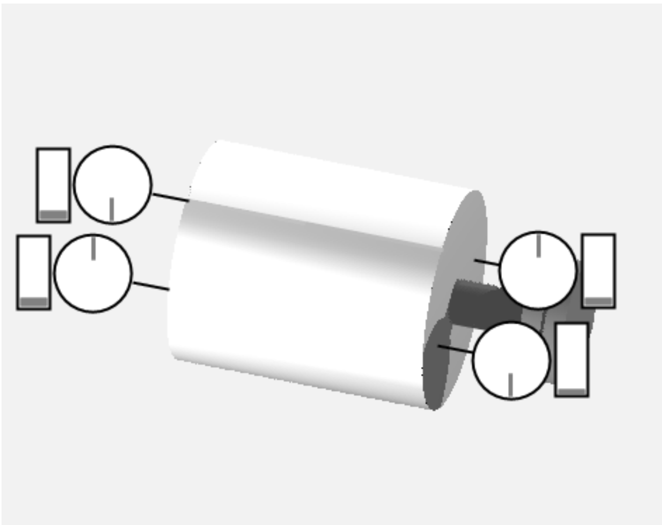


Figure 4-58

We could use our animation program to demonstrate this action. We have entered data for one component and one end of the machine.

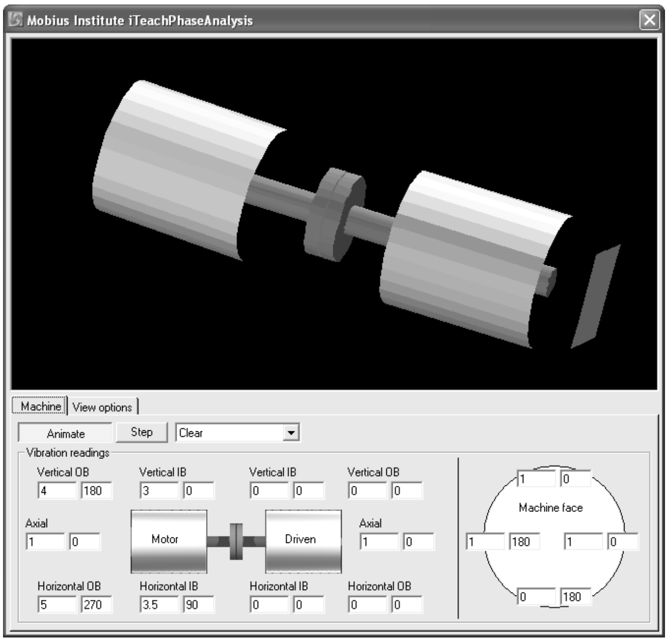
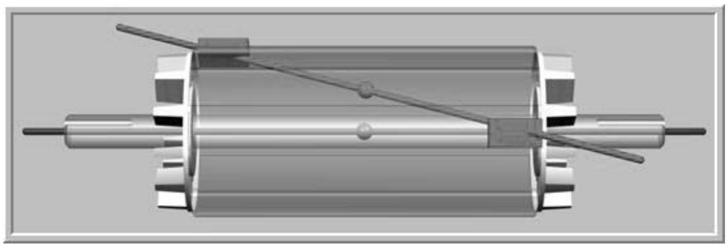


Figure 4-59

Dynamic unbalance

In reality the amount of unbalance will not be evenly distributed along the rotor (unless it is a very narrow rotor or axial fan, in which case it will approximate static unbalance). We are likely to have a combination of static and couple unbalance. The combination is called dynamic unbalance.



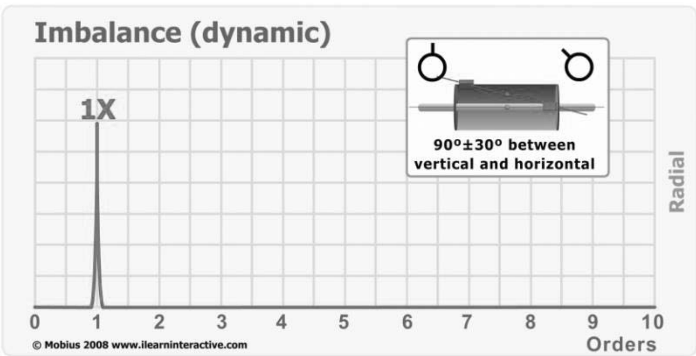


Figure 4-60

This is the most common situation found with rotating machinery (except with fans, grinders and pulleys where the ratio of length to radius (L/R) is less than 1) and you must perform a "two-plane" balance to correct the situation.

Let's revisit the five issues we considered earlier:

1. The phase relationship between the ends of the machine will be neither in-phase or 180° out-of-phase. The readings should be somewhere between 0° and 180°. The phase difference between the two vertical readings should be similar to the phase difference between the two horizontal readings. For example, if the phase difference between the vertical readings was 163°, you would expect the phase reading between the horizontal readings to be 163° give or take 20°.

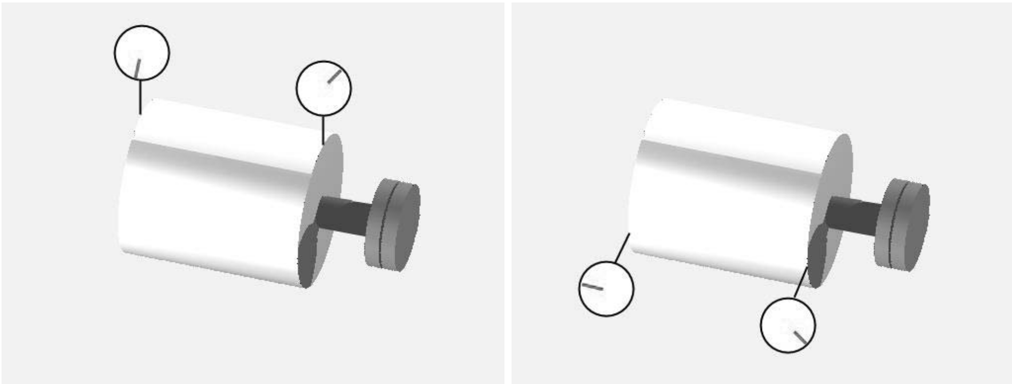


Figure 4-61

2. As we found earlier, there should be approximately 90° between the vibration readings taken in the vertical and horizontal directions. Depending upon the direction of rotation, the horizontal measurement will either lead or lag by 90° (therefore you will measure approximately 90° or 270°). This will be true at both ends of the machine.

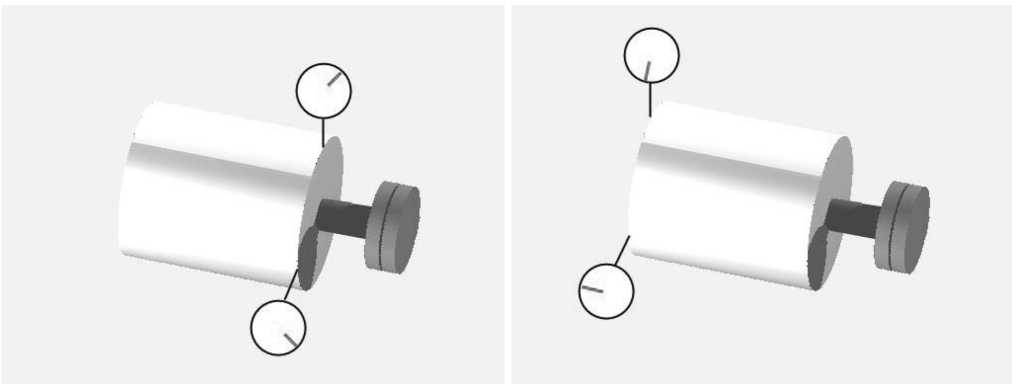


Figure 4-62

3. Once again, there will be a difference in vibration amplitude in the vertical and horizontal directions, with horizontal normally being higher due to lower stiffness. If the horizontal

reading was greater than twice the vertical reading you might consider checking the foundations/base.

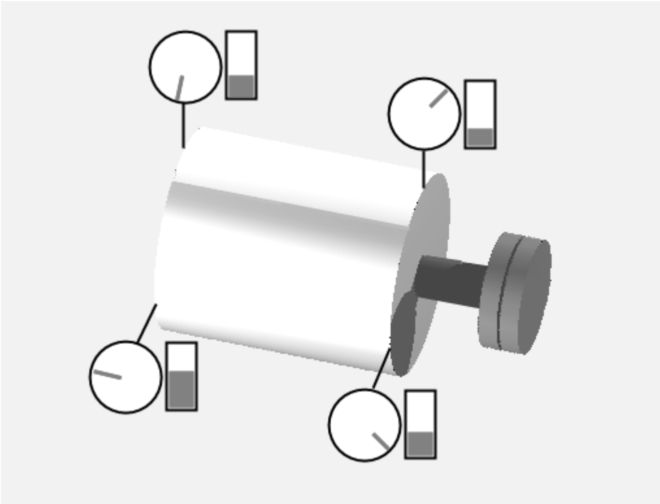


Figure 4-63

4. Once again, the vibration level at one end of the machine will be higher than the vibration measured at the other end. However the ratio between the drive end and driven end should be roughly the same in the vertical and horizontal directions. For example, if the drive end was twice the amplitude in the vertical direction, you would expect the vibration at the drive end to be approximately twice as high in the horizontal direction as well.

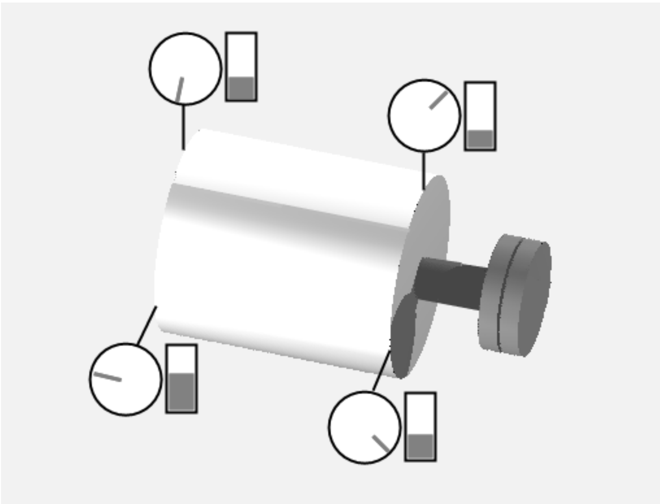


Figure 4-64

5. Once again the axial amplitude and phase readings are very revealing. The amplitude readings will be significantly lower than the radial readings. The phase readings taken across the shaft tend to follow one of two patterns. If there is a significant rocking motion (in which case the vertical phase readings will approximate 180° out-of-phase), then the axial readings will also approximate 180° out-of-phase. However, in the absence of a

rocking motion, the axial phase readings will have a smaller phase difference – they will be close to in-phase.

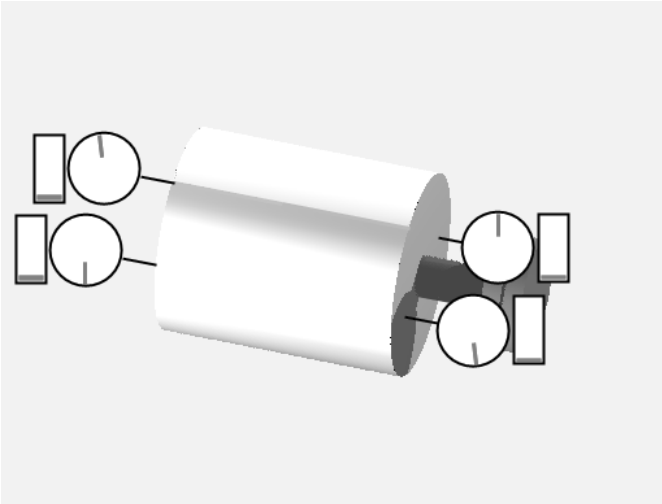


Figure 4-65

Taking all of the readings into account, we can build up an image of how the machine is moving – what forces are at play.

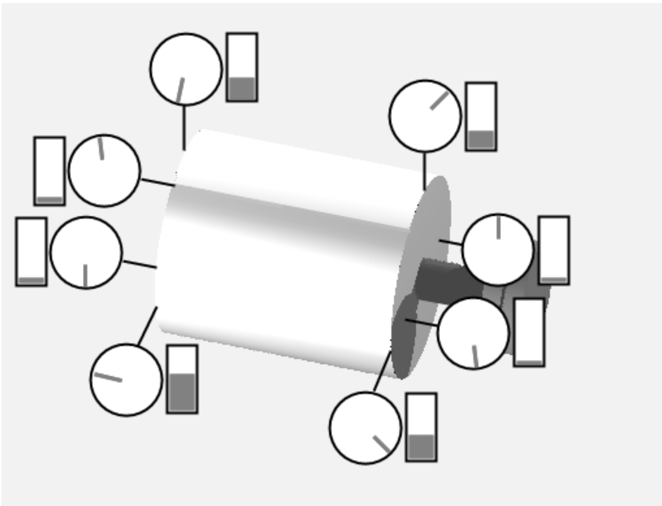


Figure 4-66

We can use the animation software to help us to visualize the movement.

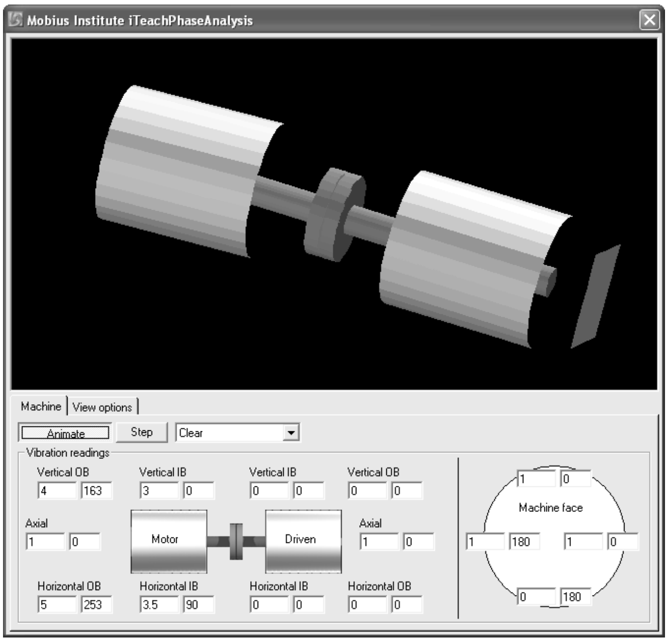


Figure 4-67

Vertical machine unbalance

Vertical machines, such as vertical pumps, are usually cantilevered from their foundation, and they usually show maximum 1X levels at the free end of the motor regardless of which component is actually out of balance.

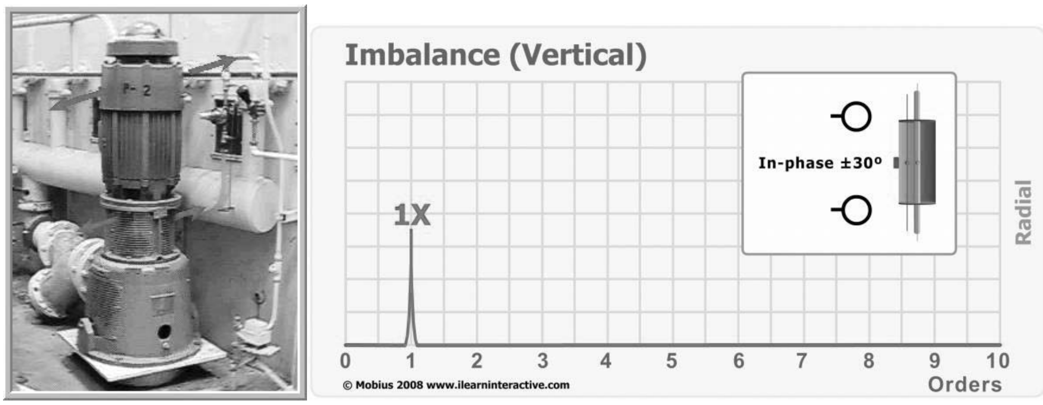


Figure 4-68

The spectrum will again show a strong 1X peak when measured in the radial direction (horizontal or tangential), and phase readings collected along the machine should be basically in-phase. Because of the circular motion that results from unbalance, the phase readings taken 90° around from the reference measurements should be 90° greater or lower; depending upon the direction of rotation.

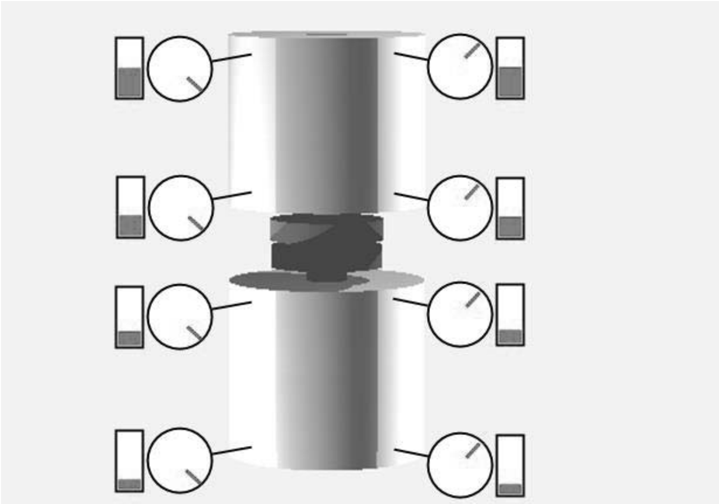


Figure 4-69

The amplitude readings should be higher toward the top of the machine. The machine will normally have greater stiffness in one direction (typically in-line with the discharge pipe), thus the amplitude readings will be lower in that axis.

Unbalance in overhung machines

The dynamics of an overhung machine are quite different; therefore our study of relative vibration levels and phase readings is quite different. Overhung pumps and fans are common in industry so you must examine the machine closely to ensure that you know whether a component is in fact overhung or supported on both sides by bearings.

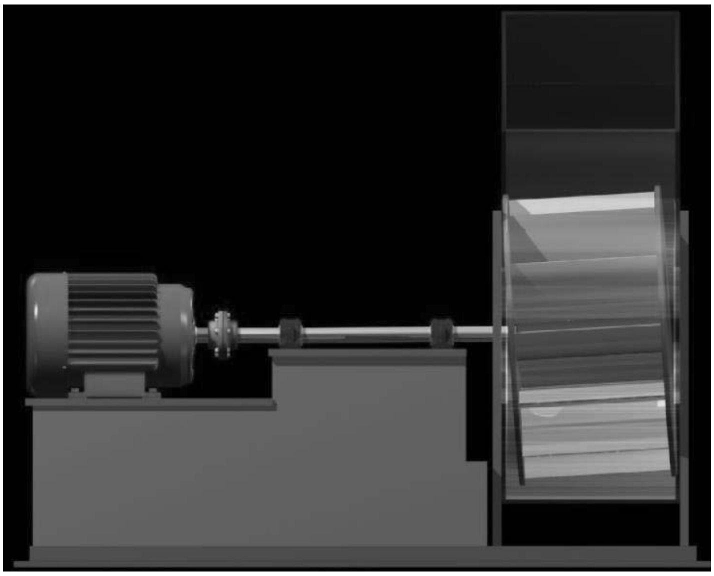


Figure 4-70

In an overhung or cantilevered machine, you will again see a high 1X vibration level, however this time it will be observed in the axial direction as well as in vertical and horizontal.

Measurements should be taken from the bearing closest to the overhung impeller or fan blades.

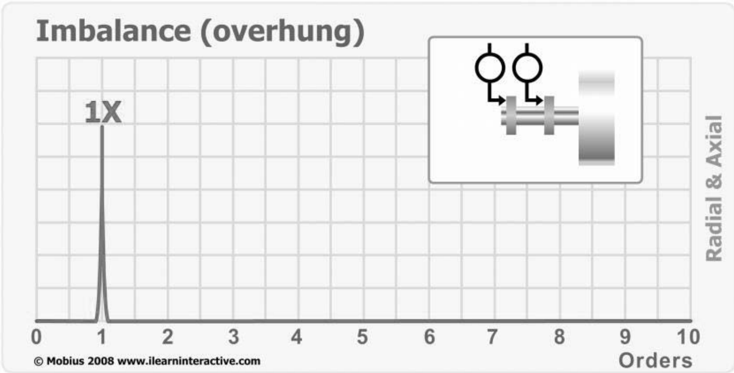


Figure 4-71

We see the high 1X in axial because the imbalance creates a bending moment on the shaft, causing the bearing housing to move axially. The readings will be in-phase in the axial direction.

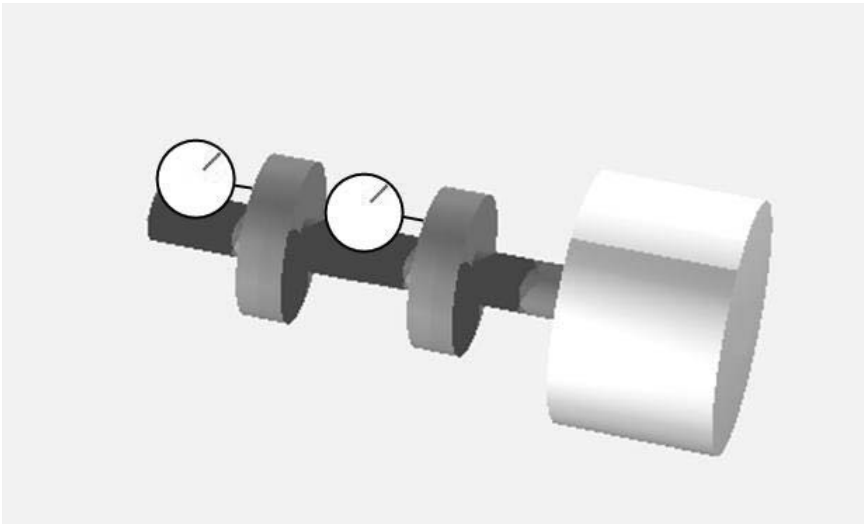


Figure 4-72

Because of the bending motion there will be somewhere between 0° and 180° between the two horizontal readings, likewise between the vertical readings. The phase difference between the vertical readings will be similar to the phase difference between the two horizontal readings.

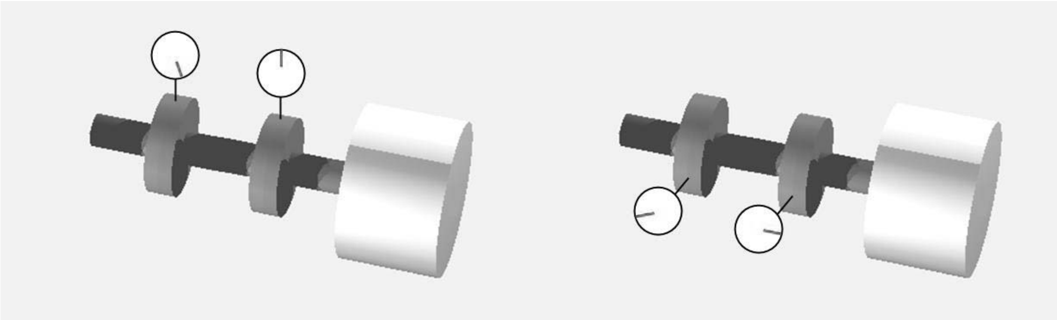


Figure 4-73

And because of the circular motion, there will be approximately 90° between the vertical and horizontal readings.

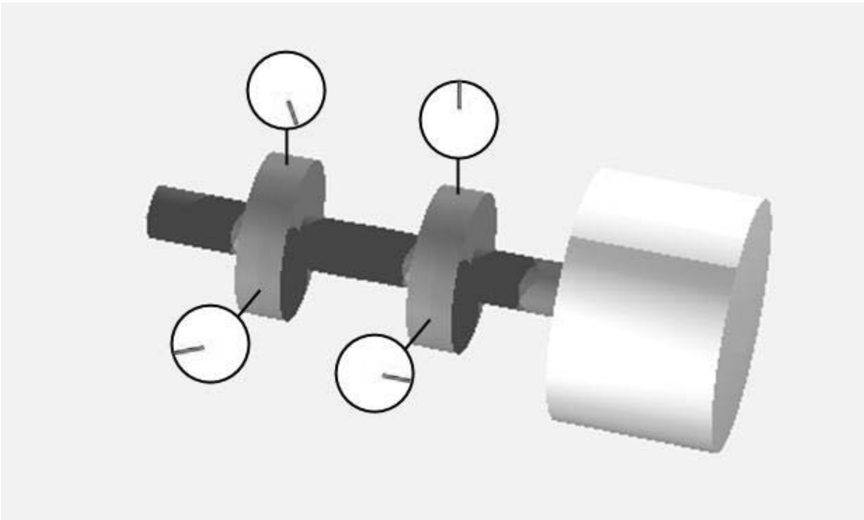


Figure 4-74

Misalignment

So many people have simple rules regarding misalignment. "If the 1X peak is high in axial then it may be misalignment. If the 2X is high in radial then it may be misalignment." Some people acknowledge that a 3X or 4X peak may also be present in the radial spectrum, however if 1X is high in radial the assumption is that the machine is out-of-balance. That can be a dangerous assumption to make. Misalignment also generates 1X in the radial direction, and phase should be used to distinguish between unbalance and misalignment.

We have previously discussed the circular motion that develops as a result of unbalance. As such we will see certain phase relationships when we compare vertical to horizontal readings, and when we compare the phase relationships from one end of the machine to the other. If we perform those comparisons and the rules do not follow those that would confirm an unbalance condition, you may suspect a misalignment condition.

There are also phase relationships that help to confirm the presence of misalignment, and to confirm conditions such as eccentricity, bent shaft and cocked bearing that could also be confused with either unbalance or misalignment. We will discuss these conditions in the following sections.

Given that the misalignment condition relates to the state of alignment between two components, we are now interested in phase relationships on more than one component. And we are interested in how the phase angle changes across the coupling. In contrast, in the balance section we focused on a single component that was thought to be out-of-balance.

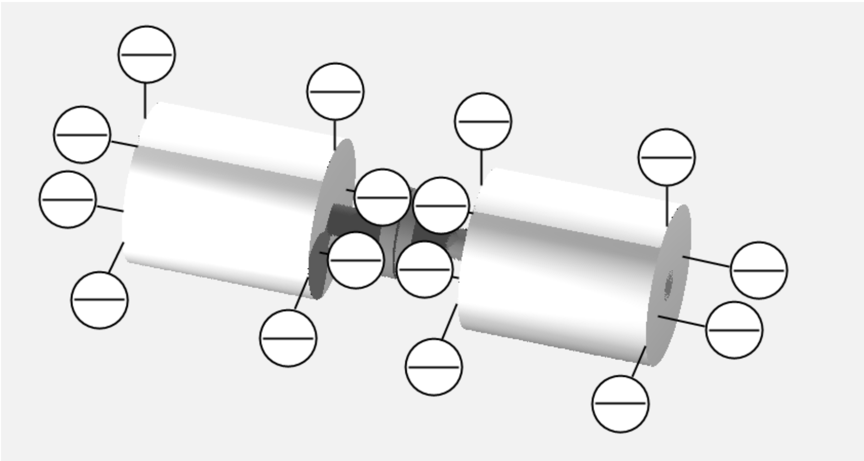


Figure 4-75

Quick review:

If the misaligned shaft centerlines are parallel but not coincident, then the misalignment is said to be parallel (or offset) misalignment.

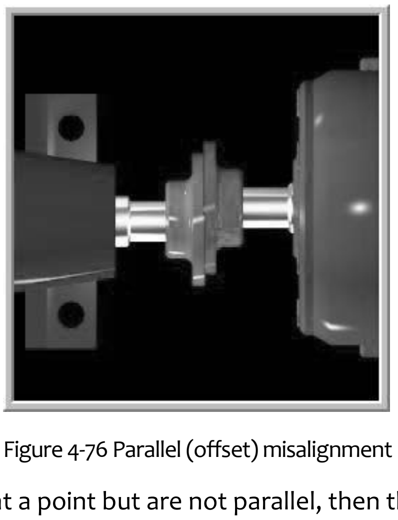


Figure 4-76 Parallel (offset) misalignment

If the misaligned shafts meet at a point but are not parallel, then the misalignment is called angular misalignment.

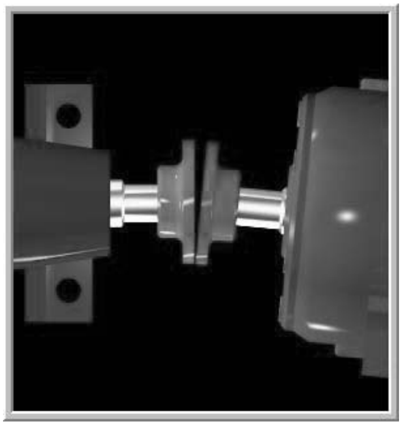


Figure 4-77 Angular misalignment

Almost all misalignment conditions seen in practice are a combination of these two basic types.

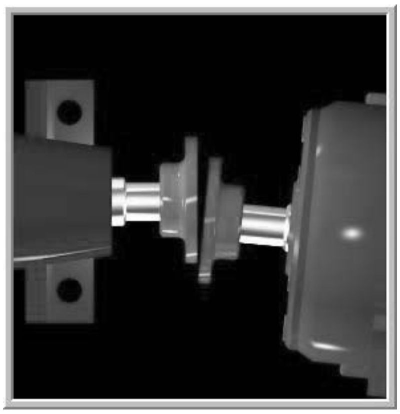


Figure 4-78 Combined parallel and angular misalignment

Angular misalignment produces a bending moment on each shaft, and this generates a strong vibration at 1X and some vibration at 2X in the axial direction at both bearings. There will also be fairly strong radial (vertical and horizontal) 1X and 2X levels, however these components will be in phase.

Misaligned components will usually produce fairly high axial 1X levels at the bearings on the other end of the shafts as well. This means that you can collect the axial reading on the outboard bearings of the motor or pump, for example, and still detect misalignment.

Parallel misalignment produces both a shear force and bending moment on the coupled end of each shaft. High vibration levels at 2X as well as 1X are produced in the radial (vertical and horizontal) directions on the bearings on each side of the coupling.

Most misalignment cases are a combination of parallel and angular misalignment. Depending upon the coupling type and other factors, you may see 3X and 4X peaks, and in some cases it may look like looseness with peaks out to 8X. One difference with looseness is that the noise floor will not be raised.

Phase relationships:

Let’s revisit the five issues we considered in the unbalance section and learn how they relate to misalignment:

1. The phase relationship between the vertical and horizontal readings taken at the ends of the machine will not follow the rules that we saw earlier. Because of the motion created with angular and offset misalignment, and the affect that different coupling types will have on that motion, the phase angle between the ends of the machine will not be consistent in the vertical and horizontal directions. If we saw that the difference between the vertical readings was 180°, then we might expect that the difference between the horizontal readings might be 0°.

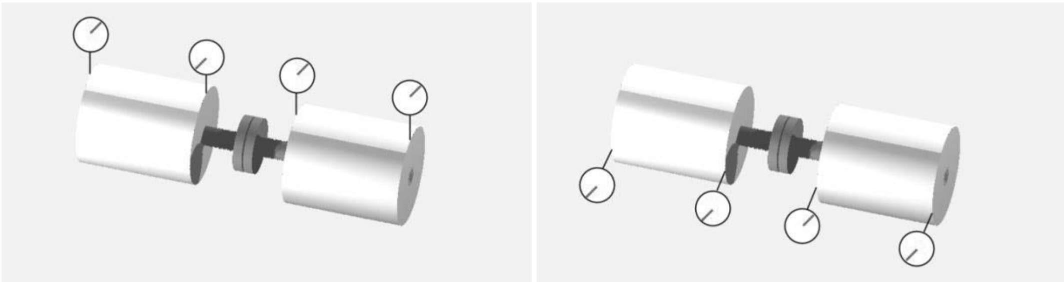


Figure 4-79

2. If a machine is misaligned, we would not expect to see 90° or 270° degrees between the vertical and horizontal readings taken at the same bearing. Instead they are likely to be closer to 0° or 180°.

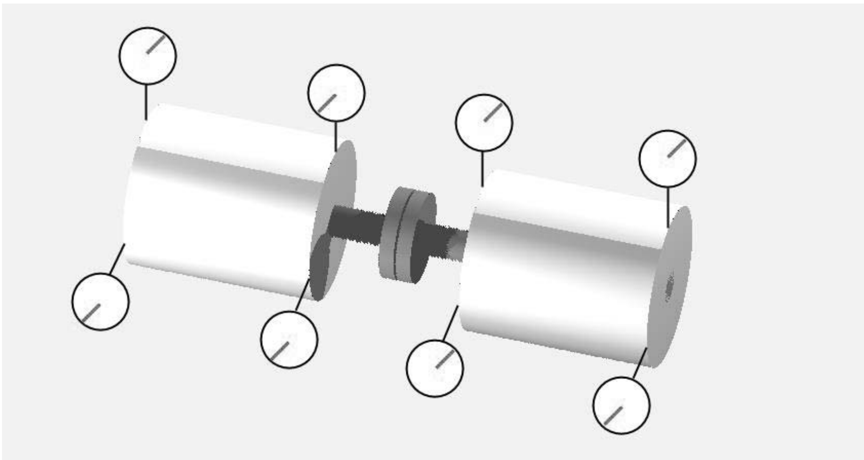


Figure 4-80

3. The amplitude readings in the vertical and horizontal directions follow quite different rules to those we learned in the unbalance section. First, the vertical readings may be higher than the horizontal readings – it depends upon the nature of the offset. It has been noted that if there is substantial lateral offset, the vibration readings will be quite high in the vertical direction.

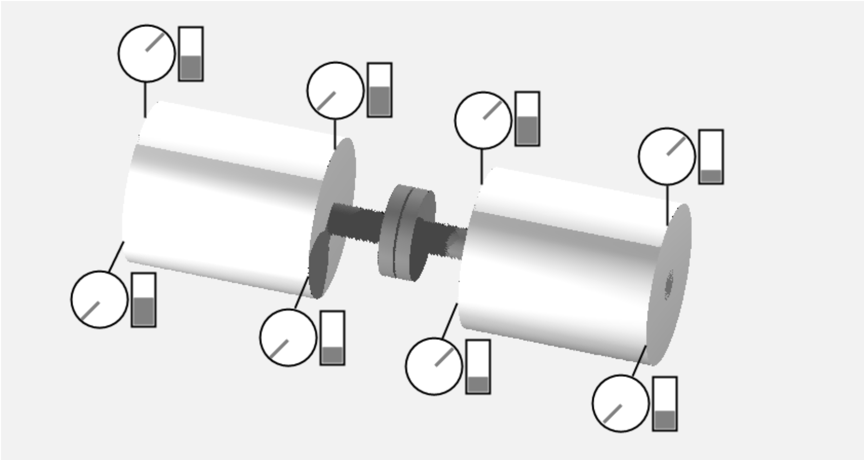


Figure 4-81

4. If you compare the ratio between the vertical and horizontal 1X amplitudes at each bearing, the ratios will be different if the machine is misaligned. The vertical may be higher in one case and lower in another.

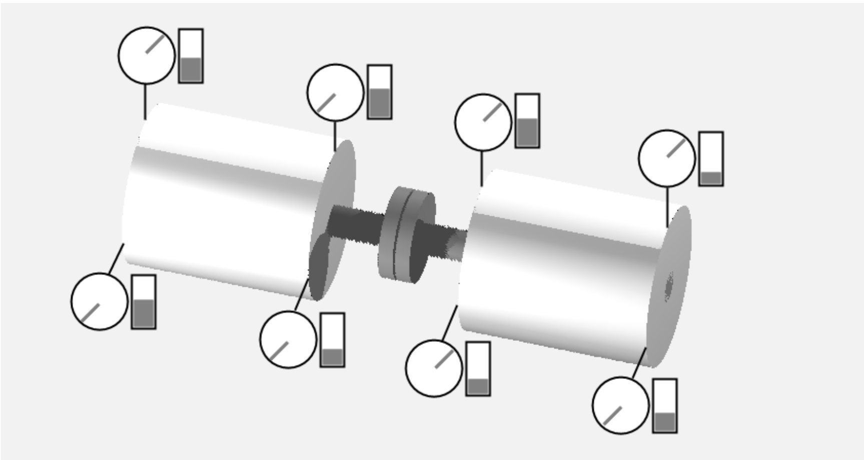


Figure 4-82

5. The relative amplitudes of the axial readings could vary significantly. If there is substantial angular misalignment, the axial readings will be high. If there is mostly offset misalignment, axial could be low in comparison to the radial readings.

   When we consider the phase readings, we have to consider two types of relationships: the phase difference from one side of the shaft to the other; and the phase difference across the coupling.

When there is strong angular misalignment you would expect the phase reading to be 180 out-of-phase across the coupling.

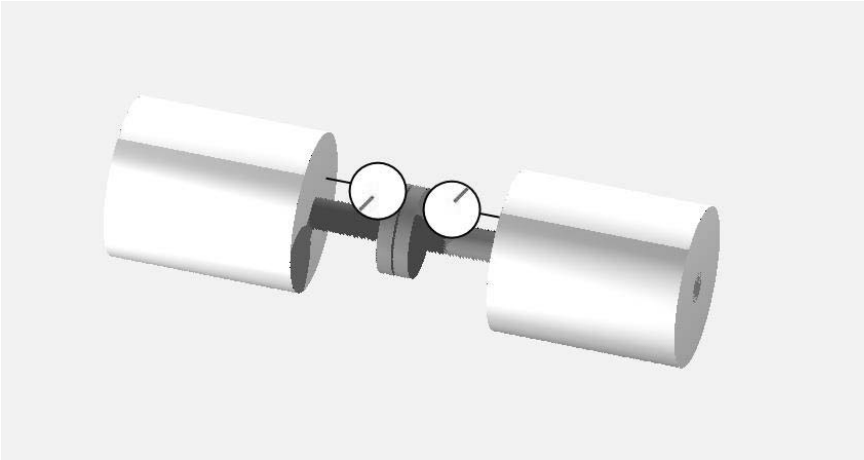


Figure 4-83

When you compare the phase readings from one side of the shaft, things can get a little interesting. If the readings across the coupling (from one machine component to the other) should be out-of-phase, then you would expect a consistent phase relationship around the shaft – in order to maintain the cross-coupling relationship. However, because of the type of coupling, the design of the two machine components, and the actual balance of angular misalignment and angular misalignment, the phase readings may surprise you. On one component the readings may be in-phase, but on the other they may be out-of-phase.

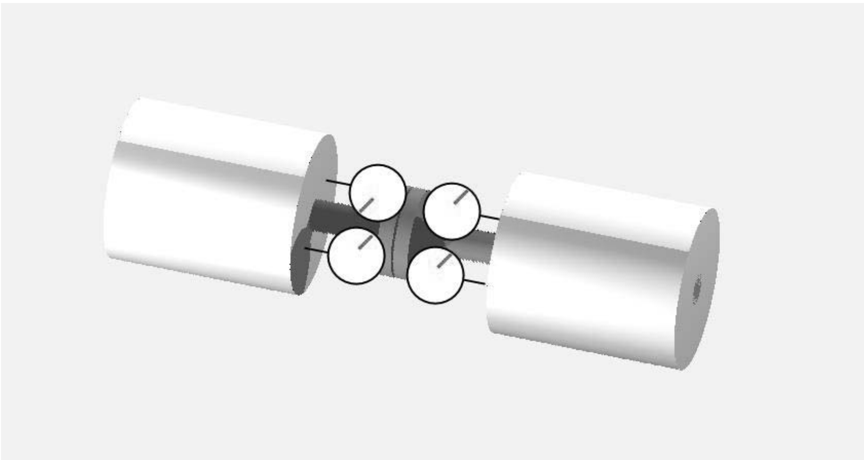


Figure 4-84

Eccentricity

Eccentricity occurs when the center of rotation is offset from the geometric centerline of a sheave (pulley), gear, bearing, or rotor.

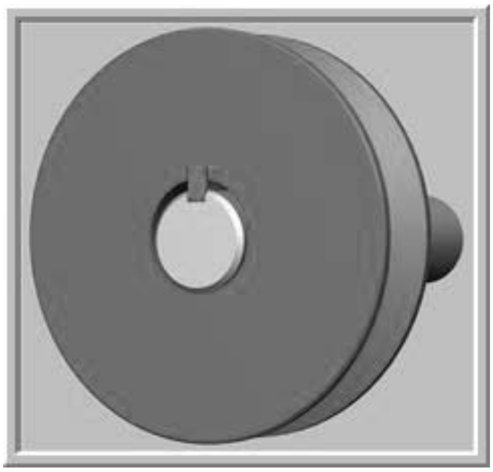


Figure 4-85

Eccentric sheaves/gears will generate strong 1X radial components, especially in the direction parallel to the belts. This condition is very common, and mimics imbalance. The phase relationship, however, is quite different, because the motion is quite different.

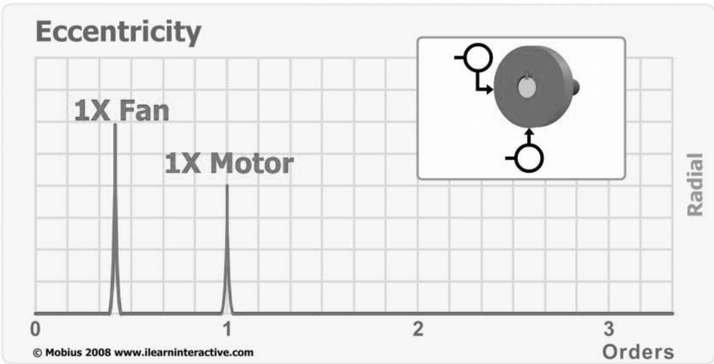


Figure 4-86

In belt driven machines, there will be a high 1X vibration level on both components (motor and fan for example), however due to the change in speed, these will be at two different frequencies.

Figure 4-87

The highest vibration will be in the direction of orange arrow (Figure 4-87), so measurements should be taken in this direction. There will be a phase difference between the measurement taken in the direction of the arrows and **at right angles** to that direction of 0° **or** 180°. Note that we are not taking phase measurements in the true vertical and horizontal directions. We are taking one measurement in line with the belts, and the other at right-angle to this direction.

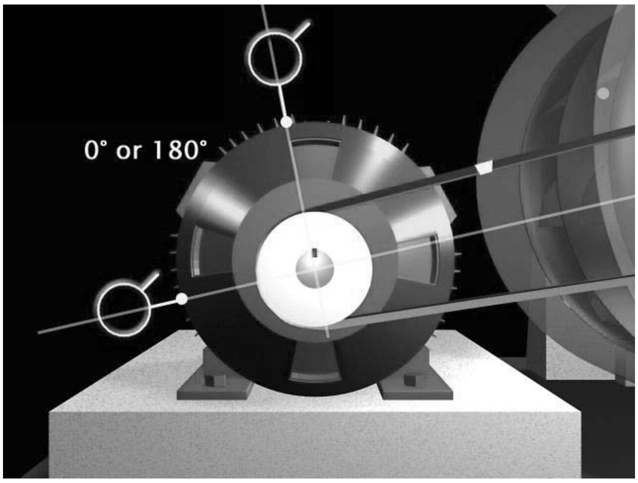


Figure 4-88

Bent shaft

A bent shaft predominantly causes high 1X axial vibration. The dominant vibration is normally at 1X if the bend is near the center of the shaft, however you will see 2X vibration if the bend is closer to the coupling. Vertical and horizontal measurements will also often reveal peaks at 1X and 2X, however the key is the axial measurement.

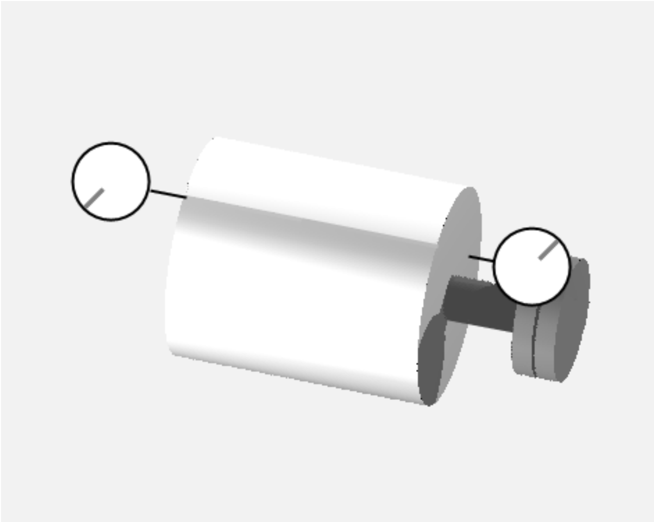


Figure 4-89

Phase is also a good test used to diagnose a bent shaft. The phase at 1X measured in the axial directions at opposite ends of the component will be 180 degrees out of phase.

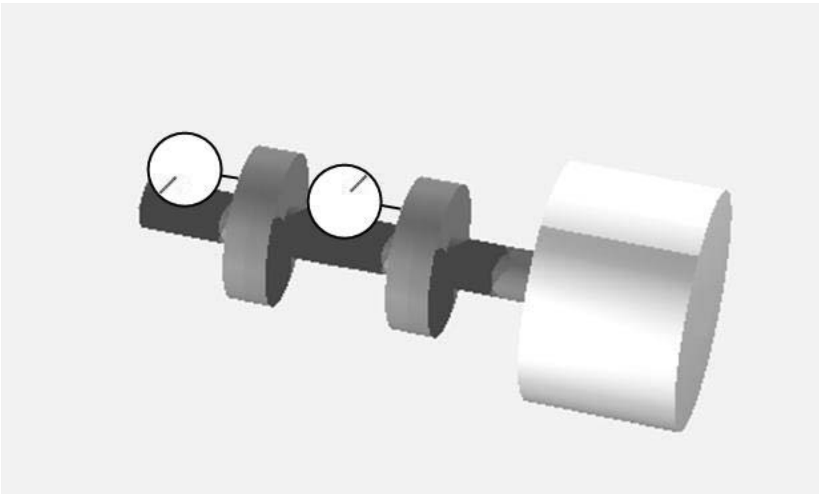


Figure 4-90

It is also possible to take phase readings around the shaft – on both sides of the shaft, and above and below. We expect all of the readings to be in-phase.

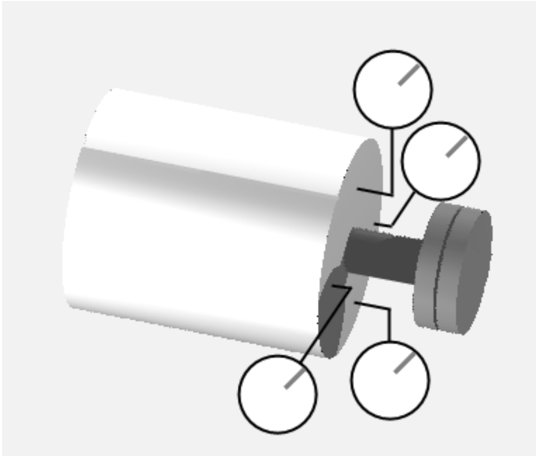


Figure 4-91

Cocked bearing

A cocked bearing, which is a form of misalignment, will generate considerable axial vibration. Peaks will often be seen at 1X and 2X in the axial direction, and even 3X and other harmonics can be seen. Given that there is such a strong axial vibration, it too can be confused with misalignment, and with imbalance in an overhung pump or fan. The presence of peaks at 2X (and 3X) in the axial direction would indicate a cocked bearing condition, distinguishing itself from imbalance.

There are actually two possible forms of cocked bearing. If the outer race of the bearing is cocked, the axial phase readings will indicate a 180° difference from one side of the shaft to the other. However, it all depends how it is cocked. The 180° difference may be seen from the left side to the right or it may be seen from the top to the bottom – but not both.

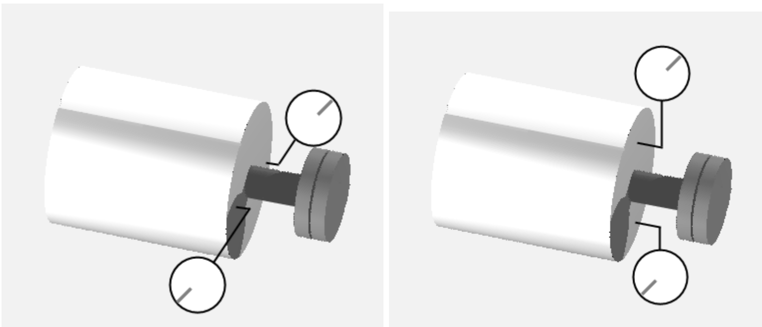


Figure 4-92

If the inner race is cocked on the shaft, then the bearing will appear to “wobble” as it rotates, generating a rotating 180° phase difference. There will be 90° difference as you move from top to right to bottom, to left (or 12:00 to 3:00 to 6:00 to 9:00).

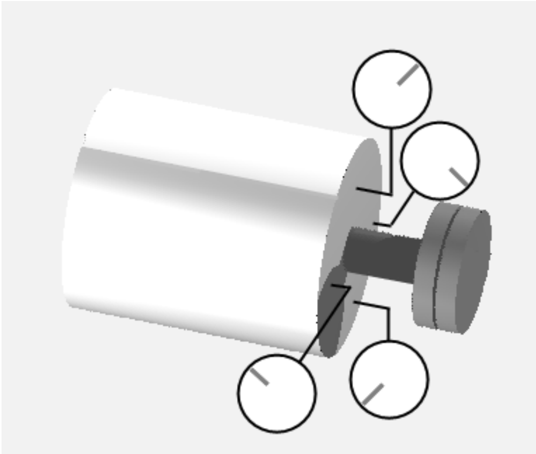


Figure 4-93

Looseness

Phase analysis can also be used to help to identify looseness and foundation problems – but in a slightly different way than we have discussed thus far.

First, because rotating looseness involves a 1X peak and harmonics, it can, in some cases, be confused with misalignment and even bent shaft and cocked bearing. However the phase readings will not follow the rules we have discussed thus far, and will be random in nature (erratic). So this can help you distinguish between the two fault conditions.

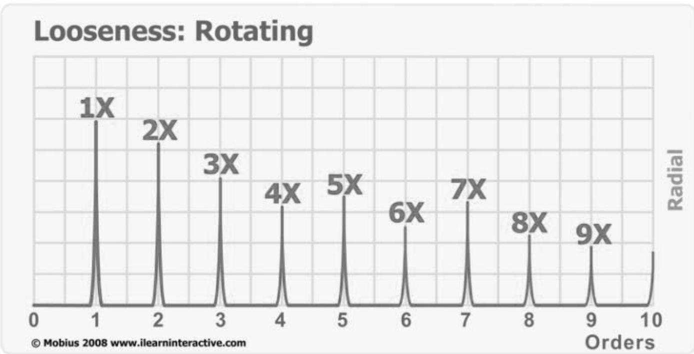


Figure 4-94

In the case of structural looseness, where there is a problem with the foundations, phase can be used in two ways.

First, if the vibration levels are high enough, the machine may rock back and forth. Phase readings taken in the horizontal direction could be in phase, but unlike unbalance, there will not be a 90° phase difference between vertical and horizontal.

If there is a crack in the foundation or a loose hold-down bolt, you can monitor the phase while you move the accelerometer from point to point. When the accelerometer moves across the crack or loose boundary, the phase angle will change by 180°.

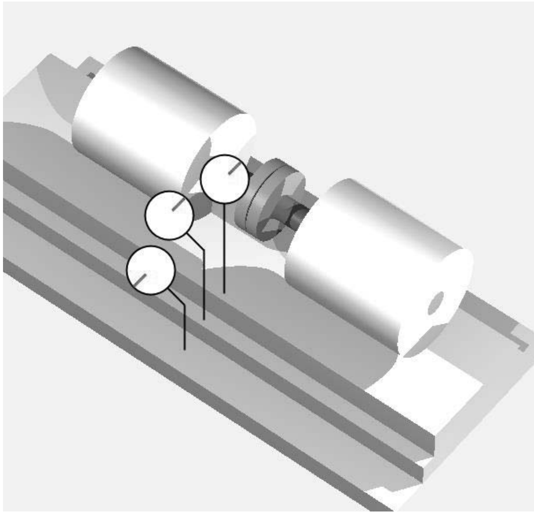
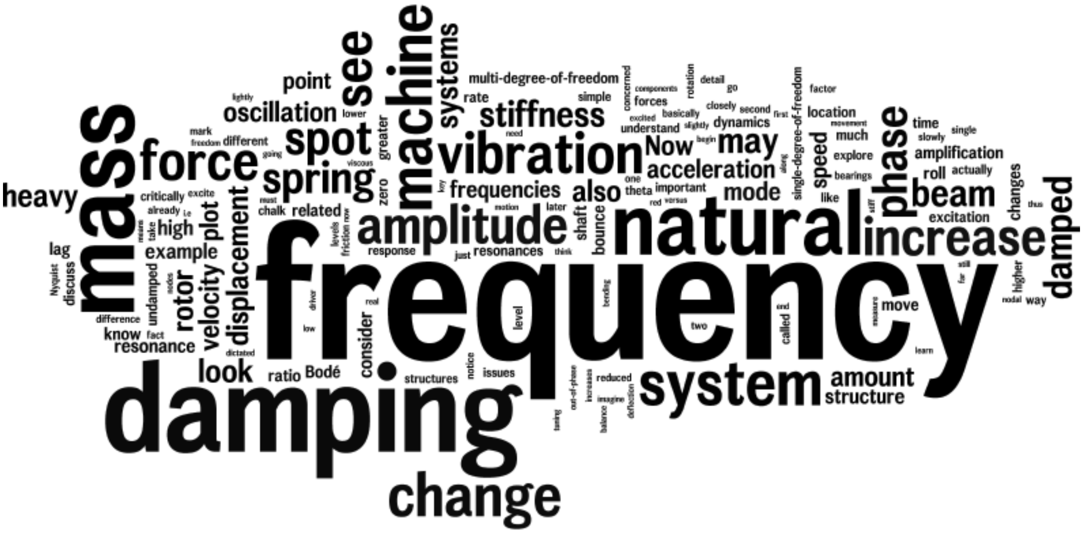


Figure 4-95

Conclusion

I hope you now have renewed interest in phase analysis. Phase can be used to help you to positively diagnose a wide range of fault conditions, and to visualize resonance conditions. If you have a two channel data collector, phase readings are not difficult to collect and should be performed frequently.



Chapter 5

System Dynamics

Objectives:

- Describe single degree of freedom systems in terms of mass, stiffness and damping
- Describe the concept of damping and the amplification factor
- Describe the concept of a natural frequency
- Explain Bode plots and Nyquist diagrams
- Describe multi degree of freedom systems

This chapter discusses system dynamics and natural frequencies in terms of mass, stiffness and damping. Various types and characteristics of damping and the amplification factor are covered in detail. Bode plots and Nyquist diagrams are explained as are multi degree of freedom systems.

System Dynamics

We are now going to take a closer look at mechanical systems. This information will provide us with some fundamental information that forms the basis for many of the issues we face in machinery monitoring. We are going to begin by looking at the three basic components: mass, stiffness and damping. We need to understand these components before we can really understand issues such as resonance, balancing, sensor use and operation, measurement location selection, and machinery vibration itself.

If we did not consider these issues, we may be compelled to think of every mechanical system as a simple linear device. If a vibratory force is generated within a machine, that same force will be transmitted through the machine and structure without any alteration, regardless of the frequencies generated. We all know that this is not the case.

Figure 5-1

We know that a machine may internally generate a dynamic force with varying amplitude levels at different frequencies, but that what we actually measure can be quite different. Amplitudes at some frequencies will be lower; much lower at certain frequencies. While at other frequencies the amplitudes will be higher than those generated.

If we did not consider machine dynamics we might also believe that a machine will generate the same pattern of vibration regardless of running speed. Of course, centrifugal forces vary with speed, but are there other changes related to speed? Sure there are. Depending on the machine, the vibration amplitudes and patterns can change dramatically. We'll learn why that is.

And if we did not consider machine and structural dynamics, we might not understand why certain structures failed (welds, brackets, etc.), or why intolerable levels of vibration were generated.

For example, you may be in a plant where bearings on a machine seem to fail regularly, even though the machine has been aligned and lubrication practices are good. Or you may witness pipes bouncing up and down or hear reports of brackets that have failed at the welds.

Figure 5-2 - (From Simon Hurricks, Genesis Energy)

Alternatively, you may be aware of very high vibration levels just at certain times during the process (or when a fault condition exists on a machine). People may complain about vibration levels when it seems that a floor in a building vibrates excessively. All of these issues are related to resonances and we will now look into this important area in more detail.

Mass, stiffness, damping – the basics

Let's boil every machine component down to its simplest elements. Systems can be described in terms of their mass, stiffness and damping. They are traditionally depicted as shown:

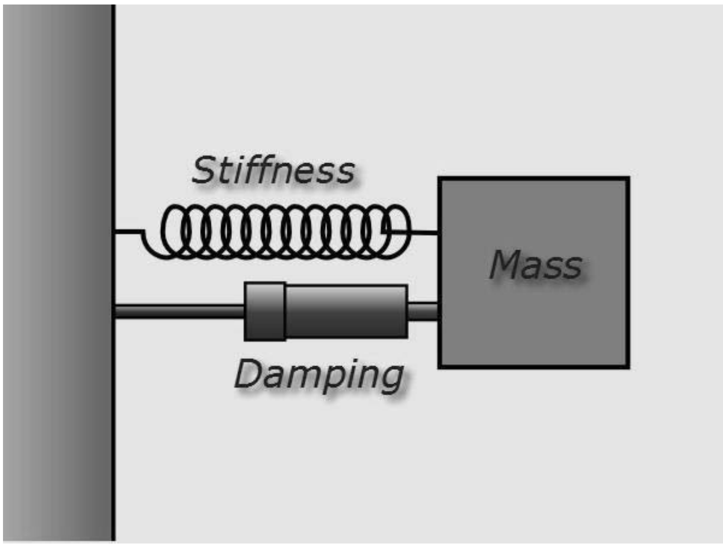


Figure 5-3

Before we look more closely at each of these elements, let's make a simple observation. If we were to pull on the mass and let it go, we would expect it to bounce up and down. We intuitively know that the amount that it bounces, the frequency of the bounce, and the length of time that it continues to bounce will be the same each time (with the same applied force), but we also know that these parameters depend upon the mass, stiffness and damping.

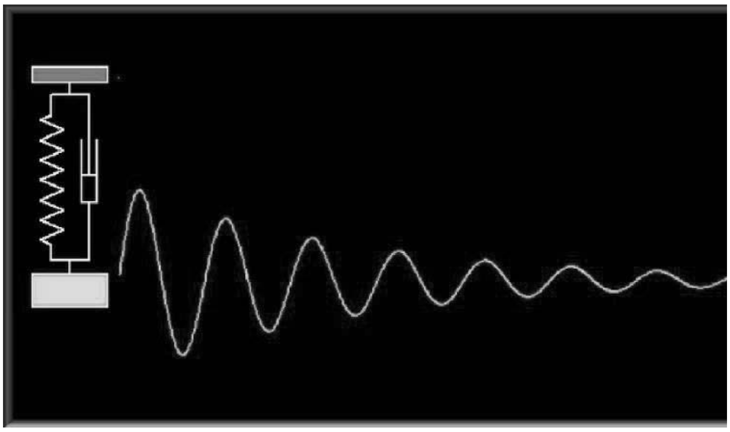


Figure 5-4

Produced by Sumico Technologies

Mass

If we increase the mass, it will bounce more slowly. Conversely, a reduced mass will cause the mass to bounce at a higher frequency.

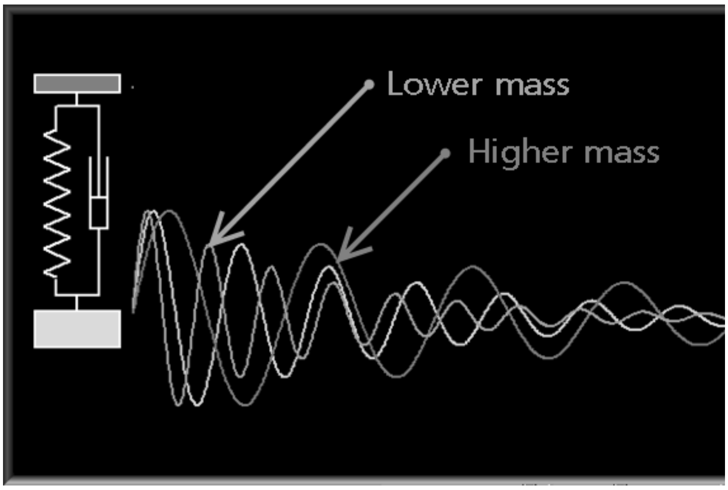


Figure 5-5

When we consider the dynamics of a real machine, the mass can basically mean one of two things. If we are considering a turbine, a paper machine, or any machine where we are concerned with the dynamics of the rotor itself, then we are concerned with the mass of the rotor and any components attached to it (turbine blades, for example).

When we are concerned with the entire structure, then of course we are concerned with the mass of the entire machine, and possibly the foundations.

In fact, in the study of the dynamics of smaller structures, the mass of the vibration sensor becomes an issue because it changes the dynamic properties of the structure we are trying to measure.

Stiffness

If we increase the stiffness of the spring, it will bounce at a higher frequency. Conversely, if the spring is less stiff, the frequency will be reduced.

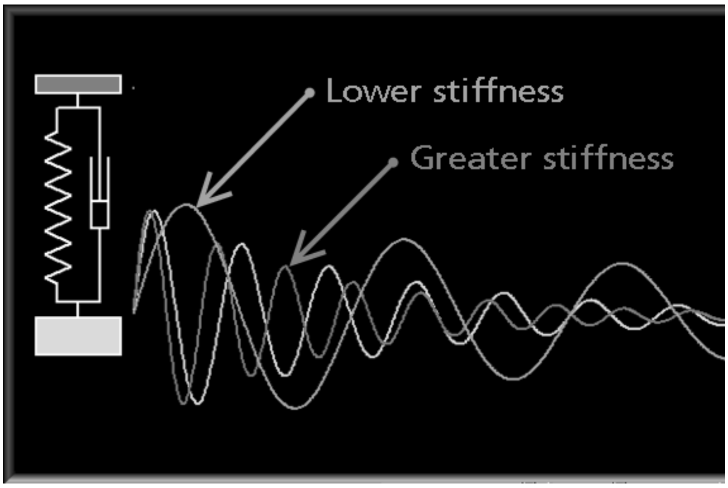


Figure 5-6

The spring represents the stiffness in the system. We can think of the foundations as the spring, or the bearings as the spring. Depending upon how closely we wish to examine the various components within the machine, we can find many examples of springs.

Damping

If we alter the damping, the frequency is not changed (very much). The damping controls the amplitude of each bounce, and thus the time it takes for the mass to stop bouncing.

Damping causes the energy in the system to be dissipated. If we had zero damping, the system would continue to vibrate endlessly.

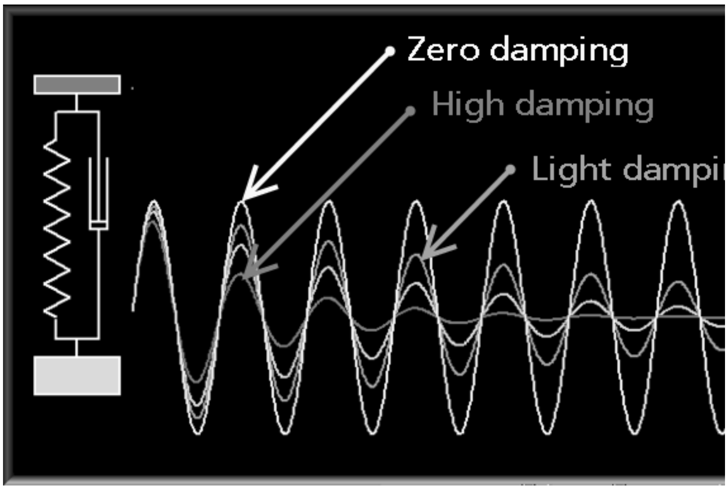


Figure 5-7

There are three main forms of damping: friction damping, viscous damping, and hysteresis damping. Frictional damping occurs when two parts slide against each other. The friction causes heat which is a loss of energy. Bearings are designed to minimize friction, so this is not the form of damping that we are most concerned with.

Hysteresis damping involves internal molecular friction. For example, as the spring is compressed and released, the friction between the molecules generates heat and eventually causes the spring to stop oscillating and come to rest. Naturally, this form of damping results throughout the machine as dynamic forces in the machine cause blades to bend, bearings to vibration, housings to flex, and so on.

We are most interested in viscous damping. Viscous damping is basically the resistance to fluid flow. An excellent example is the shock absorbers on your car. As you can imagine, the amount of damping depends on a number of factors. When you look at the shock absorber, the damping characteristics will be dictated by the viscosity of the fluid, and the size of the cylinder and plunger.

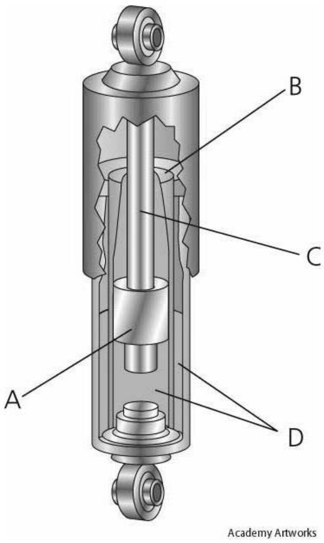


Figure 5-8

A. piston

B. cylinder

C. piston rod

D. oil

In rotating machinery, the lubrication in the bearings provides viscous damping. While the lubricant is used to facilitate the motion of the balls/rolls in a rolling element bearing, or the shaft to turn in a journal bearing, it does provide a resistive force to that motion. It also dampens vibration – the transmission of the vibration from the rotor to the machine housing.

The process itself also provides viscous damping. Whether it is a fan moving air or a pump pushing fluid, the process provides damping which is attempting to resist the rotation of the shaft.

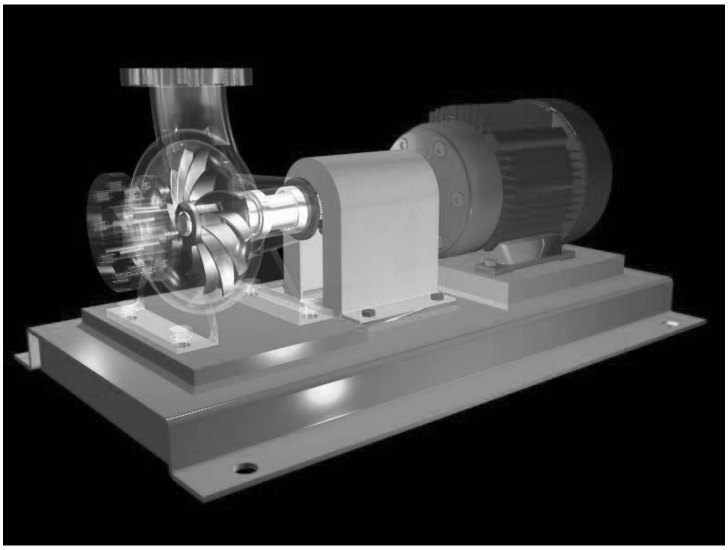


Figure 5-9

Degrees of Freedom

In this demonstration we are considering one axis of movement only – up and down. It is only free to move in a single direction. For this reason we call this system a single degree of freedom (SDOF) system.

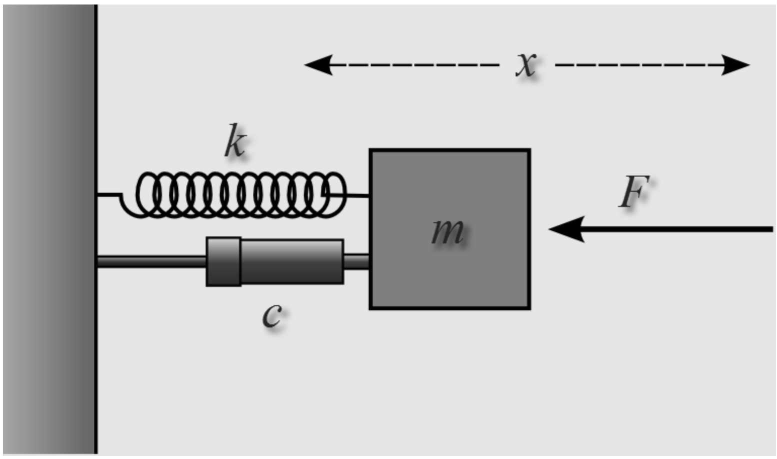


Figure 5-10

We will focus our attention on single degree of freedom systems, however it must be understood that real systems are multi-degree of freedom systems. A component may be able to move in three axes: left/right (x-axis), up/down (y-axis), and toward/away (z-axis) and rotate around each of those axes. And a component may actually have the equivalent of multiple springs, masses and dampers.

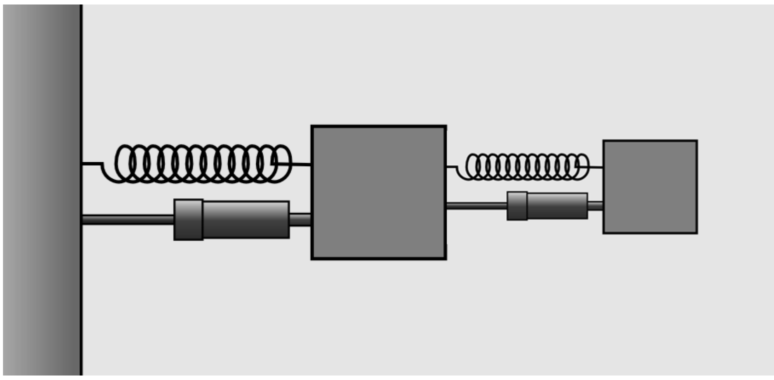


Figure 5-11

When we discuss modal analysis, and to a lesser extent ODS, we will look more closely at multi-degree-of-freedom systems. We can basically say, however, that we can model the system as a series of single-degree-of-freedom systems, and take an interest in the natural frequency of each mode.

So, let's look more closely at the single-degree-of-freedom system. We already know that if we apply a force the spring will be compressed or stretched, and then it will begin to oscillate. The damping affects the duration of oscillation, but not the frequency of oscillation (later we will learn that it does actually does change the frequency slightly, but we'll explore that later on). The frequency is dictated by the stiffness of the spring and the mass. The change in force over time can be described as:

$$f(t) = kx + c_v\dot{x} + m\ddot{x}$$

$f = Force$

$k = Spring\ constant - Stiffness$

$c_v = Viscous\ damping$

$m = Mass$

$x = Displacement$

$\dot{x} = Velocity\ (rate\ of\ change\ of\ displacement)$

$\ddot{x} = Acceleration\ (rate\ of\ change\ of\ velocity)$

This equation may seem very complicated, but it is not really. It simply means that the force is a combination of the displaced stiffness (the compression and stretching of the spring), the product of the damping and the velocity (rate of change of displacement), and the product of the mass and the acceleration (rate of change of velocity).

We already know that if we integrate acceleration we get velocity, and if we integrate velocity we get displacement. These equations are the opposite – they are derivatives, which basically means "rate of change". The derivative (or differentiation) of displacement is velocity, and the derivative of velocity is acceleration – therefore the double derivative of displacement is acceleration.

In fact, we can think of this in other more intuitive ways. We know that the force is related to the compression and stretching of the spring. As we push down on a spring it tries to push back with a force. If it is not very stiff, it is easy to compress and stretch. A very stiff spring may be very difficult to compress and stretch. The reaction force is:

$$f(t) = kx$$

Now let's look at the damping. Visualize a shock absorber, or even a simple bicycle pump. If we slowly push down on the pump handle, it will slowly move. There is resistance, but we can move it. But if we push hard and try to make it move quickly (i.e. we increase the velocity) then the resistance is much greater. The viscosity of the air is great enough to resist the velocity. We will also notice that the pump gets hot – that is the loss of energy. The reaction force is:

$$f(t) = c_v \dot{x} \quad f(t) = c_v \frac{dx(t)}{dt}$$

We should also be familiar with the equation F = mA, or Force is equal to the product of mass and acceleration. Our accelerometers work on this principle – the force on the mass on a crystal with an output proportional to acceleration. So the last part of the equation is basically related to the acceleration of the mass. The reaction force is:

$$f(t) = m\ddot{x} \quad f(t) = m\frac{d^2x(t)}{dt^2}$$

There are two key things we can learn from this. The first key point is that at low frequency, acceleration is very low, so the mass has very little effect on the force, but the stiffness of the spring has much greater impact as displacement is higher – we need a greater force to compress (or stretch) the spring.

$$f(t) = kx + c_v\dot{x} + m\ddot{x}$$

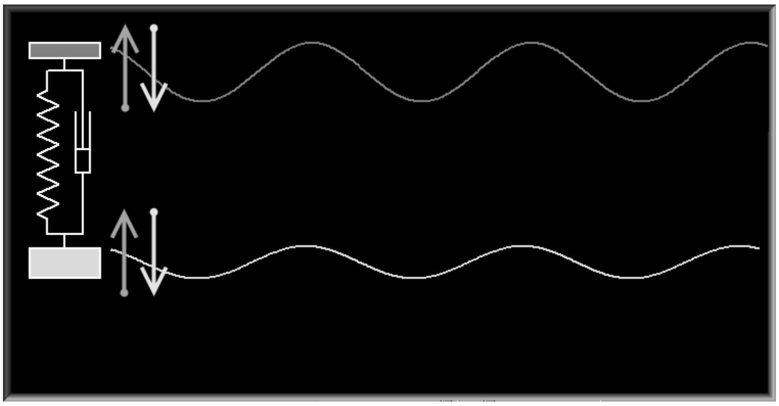


Figure 5-12

At high frequency the mass has far greater impact on the forces, and spring stiffness has a much reduced affect as the displacement is much lower. We will explore this in more detail shortly.

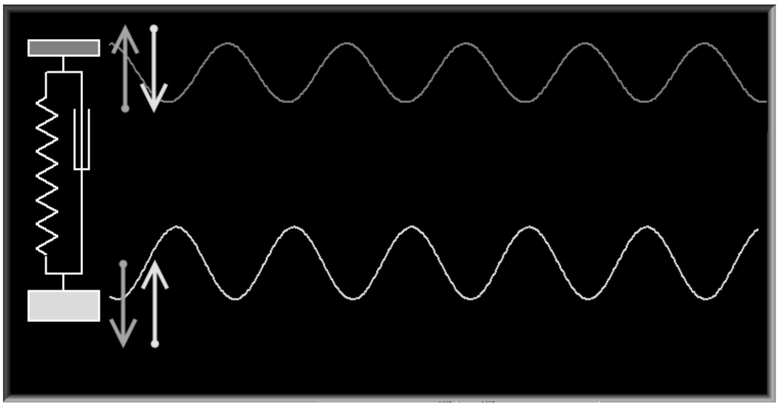


Figure 5-13

Frequency of oscillation

What we have not yet discussed is the frequency of oscillation. We can calculate the damped or undamped natural frequency. We will first consider the 'undamped' case, followed by the 'damped' cases: critically damped and over-damped.

Undamped natural frequency

The undamped frequency of oscillation as follows:

$$\omega_n = \sqrt{\frac{k}{m}} \quad \text{or} \quad f_n = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$$

If the stiffness 'k' is higher, the frequency will be higher. If the mass 'm' is greater, the frequency will be lower. (Note: omega 'ω' is in units of rad/sec and 'f' is in Hz)

If the system is undamped the vibration will not decay.

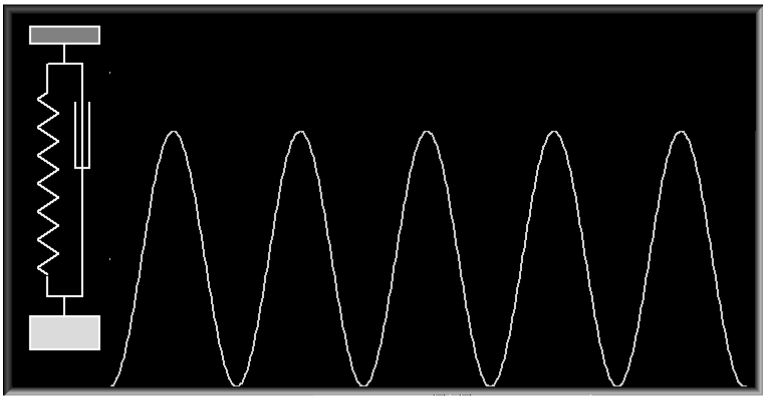


Figure 5-14

Critically damped system

The critical damping factor c_c can be interpreted as the *minimum damping* that results in non-periodic motion (i.e. simple decay). The damped natural frequency is zero.

$$c_c = 2m\sqrt{\frac{k}{m}} = 2m\omega_n$$

We can consider the damping as a ratio between zero damping (undamped) and critically damped. It is denoted by zeta:

$$\zeta = \frac{c_v}{c_c}$$

Based on the equations we have already developed, we can exchange the damping terms with zeta to determine the damped natural frequency.

$$\omega_d = \omega_n \sqrt{1-\zeta^2} \quad f_d = f_n \sqrt{1-\zeta^2}$$

We have already seen that:

$$f_d = f_n \sqrt{1-\zeta^2}$$

Where:

$$\zeta = \frac{c_v}{c_c}$$

When the system is critically damped:

Therefore zeta is equal to 1. Therefore the natural frequency is zero – i.e. there is no frequency of oscillation.

$$\zeta = \frac{c_v}{c_c} = 1$$

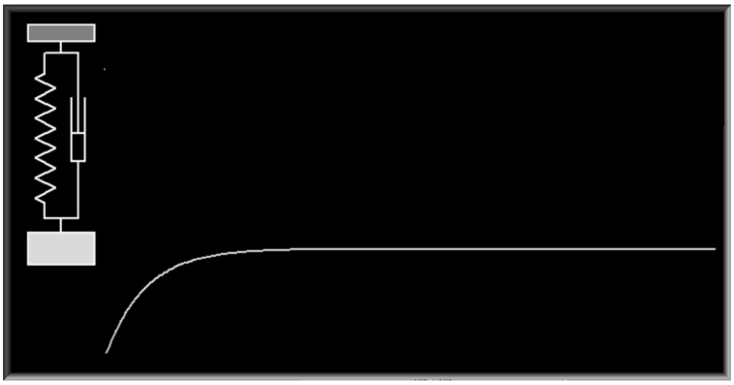


Figure 5-15

Over-damped systems

In an over-damped system there is no oscillation. The damped natural frequency is zero. It takes longer to decay than if it were damped.

$$c_v > c_c$$

Therefore:

$$\zeta = \frac{c_v}{c_c} > 1$$

Here are a few examples:

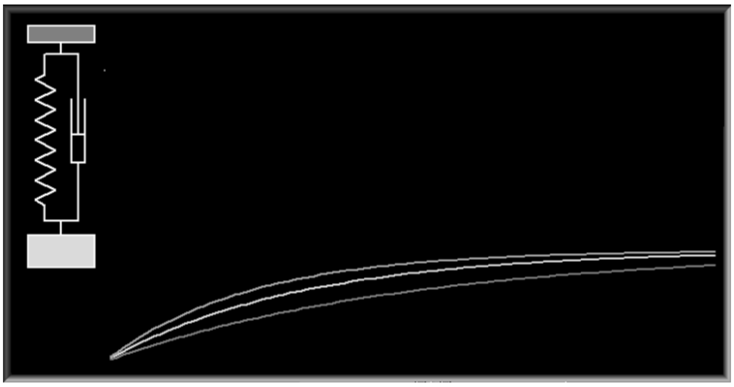


Figure 5-16

Under-damped systems

Under-damped systems are a special case where, theoretically, the oscillation never ends. The frequency of oscillation is the same as the undamped natural frequency.

If we again consider damping as a ratio of the undamped and critically damped values, then the ratio must be less that '1' when the system is under-damped.

$$c_v < c_c$$

$$\zeta = \frac{c_v}{c_c} < 1$$

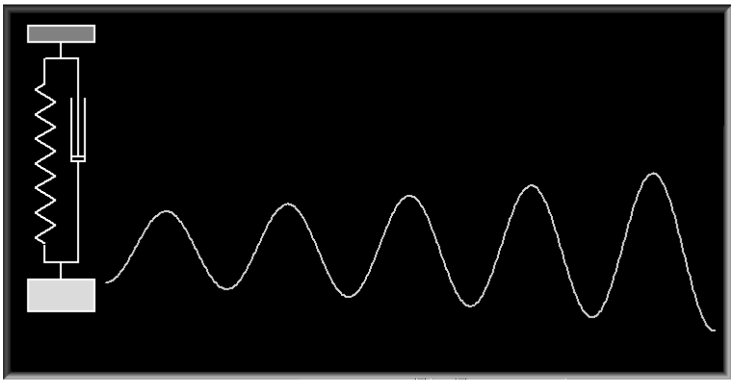


Figure 5-17

Damping and amplitude ratio

The amount of damping determines the amount of amplification (or attenuation) we experience in vibration level at the natural frequency. Here you can see the amount of amplification at frequencies above, below and equal to the natural frequency for different values of zeta (the damping ratio).

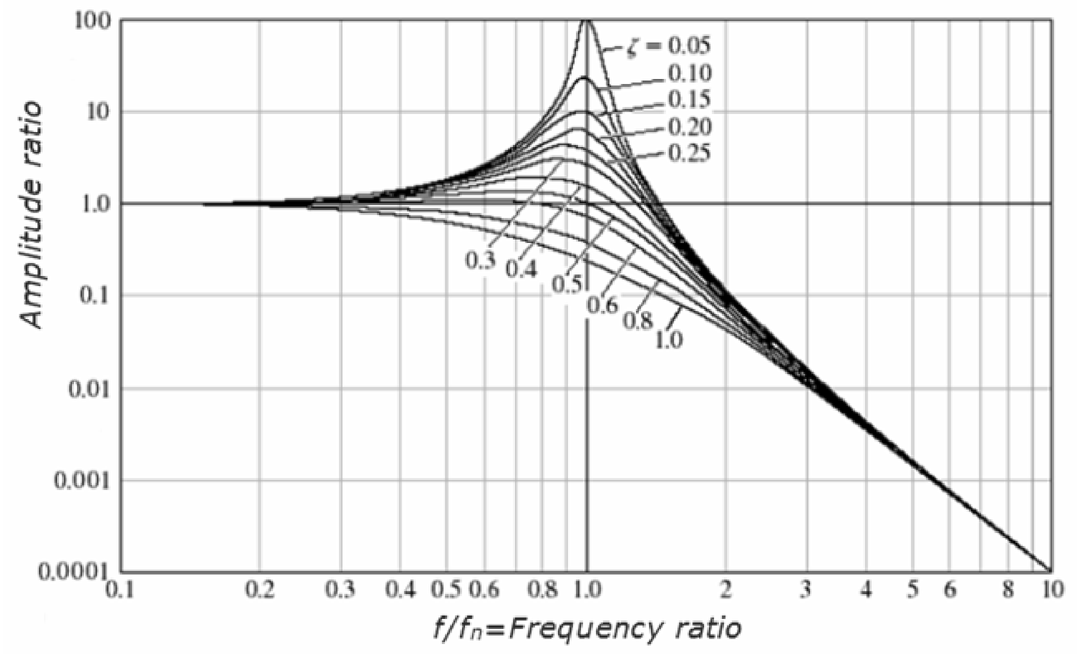


Figure 5-18

Remember: When zeta is '1' the system is critically damped. When zeta is very small (for example, 0.05) the system is very lightly damped and there will be significant amplification at the natural frequency (and above and below that frequency).

Note that the x-axis of the chart is normalized to the natural frequency, and the y-axis is a logarithmic ratio of the input vibration to the output vibration (i.e. the response compared to the input excitation force).

Damping and phase change

We will explore the relationship between the natural frequency and phase in greater detail later in this section, however this chart not only illustrates how the phase will change as a function of frequency, but also how it is affected by damping. The phase will always change by 180°, but the rate at which it changes as we approach the natural frequency is determined by the damping.

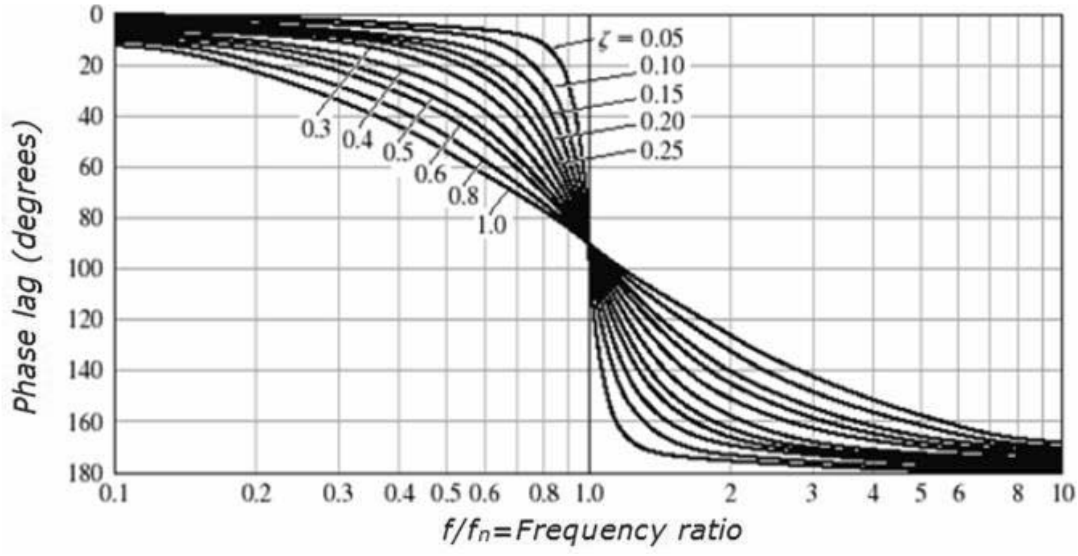


Figure 5-19

Amplification factor: 'Q'

The amplification factor is the amount of mechanical gain of a structure when excited at a resonant frequency. Q is the ratio of the amplitude of the steady state solution (amplitude at resonance) to the static deflection for the same force F.

The amplification factor is a function of the system damping.

For a damping ratio ζ =0 (no damping) the amplification factor is infinite,

For ζ =1 (critically damped) there is no amplification.

$$Q = \frac{X_o}{X_{ST}} = \frac{1}{\sqrt{\left[1-\left(\frac{f}{f_n}\right)^2\right]^2 + \left[2\zeta\frac{f}{f_n}\right]^2}}$$

X_{ST} Static deflection

X_o Dynamic deflection

Excited natural frequencies

Now we will look at the natural frequency, and the effect of mass, stiffness and damping in practice.

A single-degree-of-freedom system has a single natural frequency. As stated earlier, most structures we deal with are multi-degree-of-freedom systems, and will have multiple natural frequencies – but we'll discuss that later. So, all structures have natural frequencies. In many cases they lie dormant within a machine and we are not aware of them.

A tuning fork is a classic example of a structure with a known natural frequency. If the tuning fork is for the note A, the frequency of the tuning fork is 440 Hz. But unless someone strikes the tuning fork, the natural frequency lies dormant.

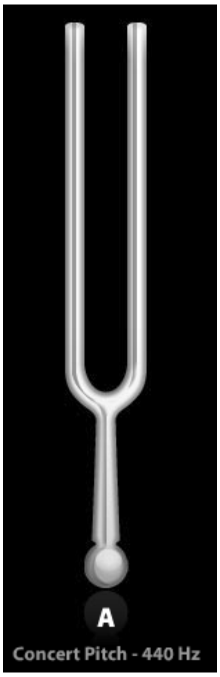


Figure 5-20

When we pull our mass and let go, the system oscillates at its natural frequency. When a system is excited at its natural frequency, it is said to "resonate". We will discuss resonance in greater detail shortly.

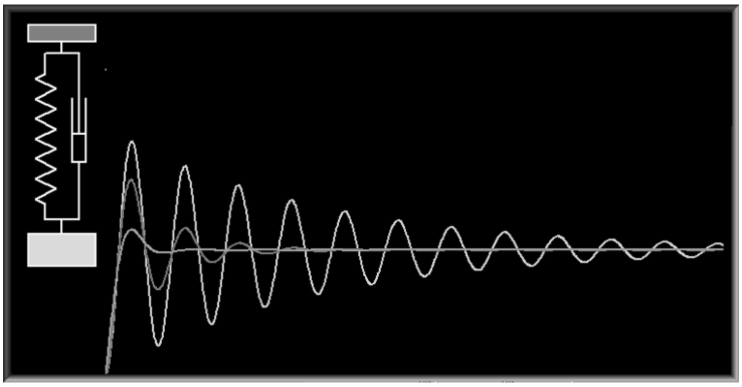


Figure 5-21

As we learned earlier, the stiffness and mass dictate the natural frequency, and the damping dictates the duration of the vibration.

Exciting the system

Now we need to take this to the next level. Rather than plucking the mass and watching how it responds, we will now excite the structure and see what happens. Now the frequency of oscillation is dictated by the frequency of excitation (the red block at the top of the spring, damper, mass system).

If we start with the frequency down low we will see that the red driver and the green mass seem to be in phase. The sine waves rise and fall together. It looks like the spring is incredible stiff. You could almost imagine that it is a steel rod between the driver and mass.

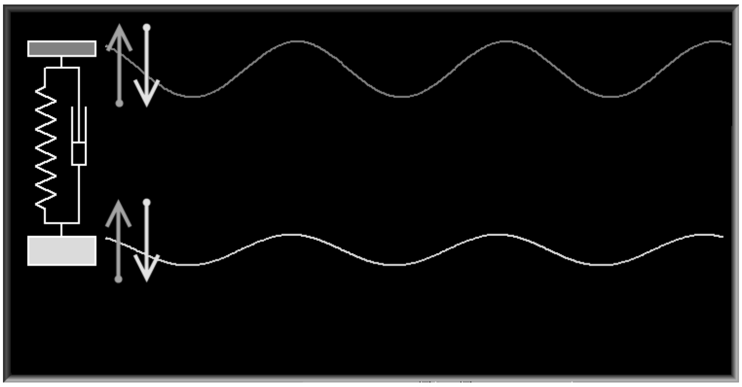


Figure 5-22

As we increase the frequency it may look like they are still in phase. You just might notice that the amplitude of the oscillation increases slightly. In fact, if you look very closely, you may see that there is a slight lag between the oscillation of the red driver and the green mass.

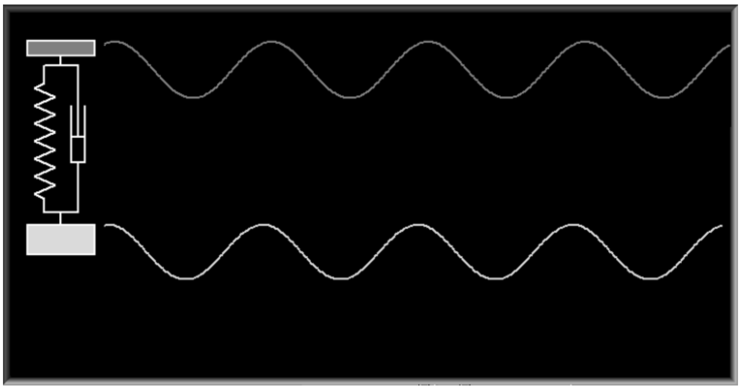


Figure 5-23

But a curious thing happens as we increase the frequency of excitation. You will notice that the amplitude of the oscillation begins to increase more dramatically. You may also notice that the *lag* between the driver and the mass is increasing.

If we tune the frequency so that the amplitude is the greatest we can see two things. First, we can see that there is a 90° phase difference between the driving force and the response of the mass. This is the definition of a resonance.

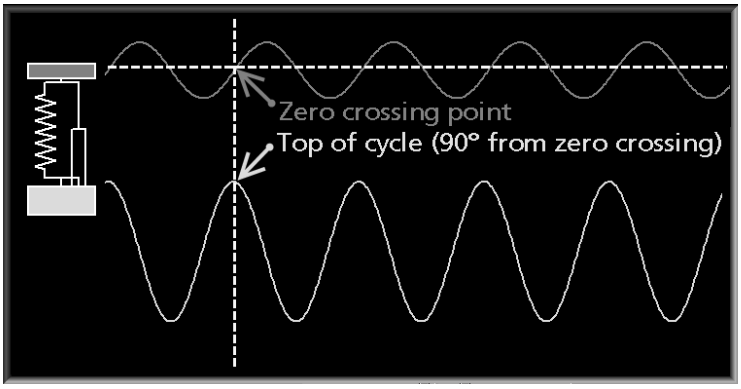


Figure 5-24

The second key point is that at resonance, the amplitude of vibration is dictated entirely by the damping. At the natural frequency, the forces due to stiffness and the mass cancel each other out (we'll explore this a little more in a moment). If we increase the damping, the amplitude will decrease, but the frequency will not change. This is called being "heavily damped".

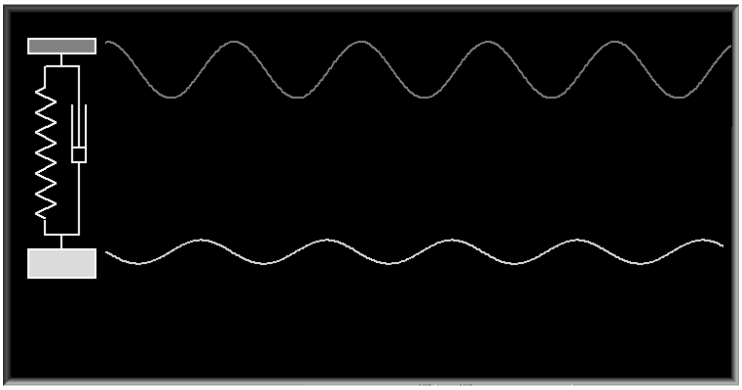


Figure 5-25

If we reduce the damping, the amplitude will increase. If we have a low amount of damping we call it lightly damped. If damping was zero, the amplitude would be infinite!

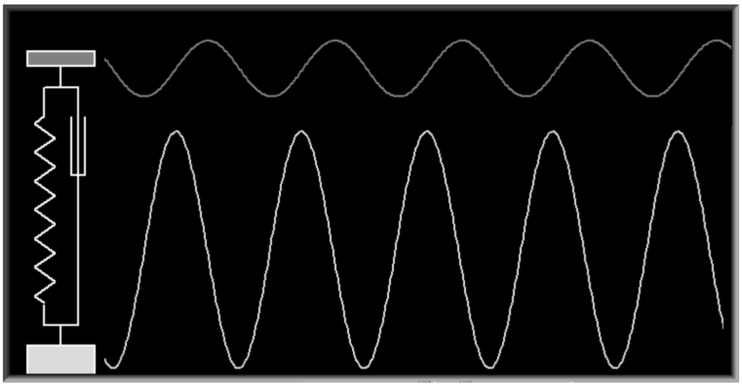


Figure 5-26

If we continue to increase the frequency, an unusual thing happens. The amplitude drops off (which you would have expected), but the phase lag also increases. In fact, you don't have to increase the frequency by very much (especially when it is lightly damped) to see a 180° lag between the excitation and the response. Now you can see that as the red block moves down, the green mass is moving up.

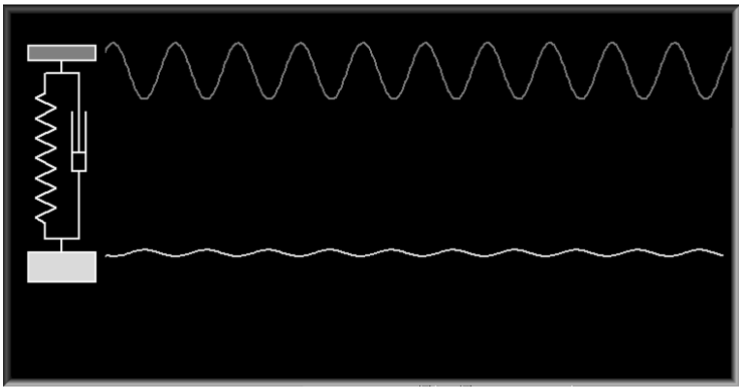


Figure 5-27

Perhaps this is not so unusual. If you visualize holding a long spring with a mass on the end and slowly moving it up and down, the mass would move up and down slowly. If you control your hand movement, you will find the resonant frequency. If you go faster, the mass will bounce up as you are pushing down, just as you can see in this simulation.

Phase relationships and vibration signals

Let's try to understand this in a different way. Do you remember the relationship between displacement, velocity and acceleration? You may remember that velocity leads displacement by 90°. Acceleration leads velocity by 90°, which means that acceleration leads displacement by 180°. For this reason, as you may remember, when you integrate acceleration to velocity there is also a phase shift.

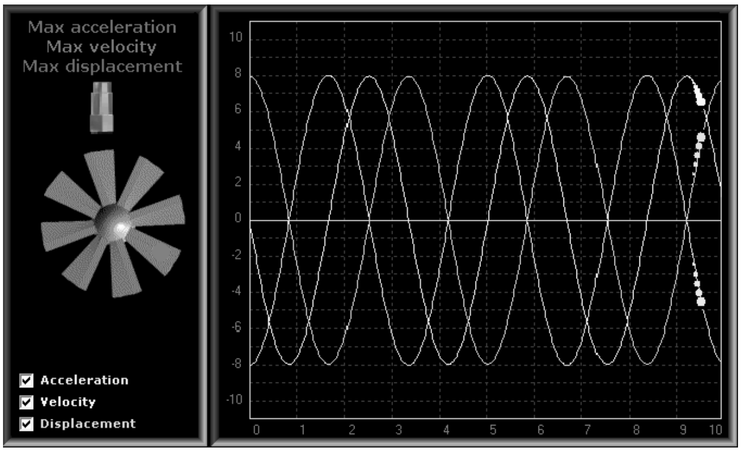


Figure 5-28

So, if we look at the equation again, we can see that the force related to stiffness is 180° out-of-phase with the force related to the acceleration of mass. And both are 90° out-of-phase with the force related to damping.

$$f(t) = kx + c_v\dot{x} + m\ddot{x}$$

That's why the force related to stiffness is 180° out-of-phase with the force due to the mass, and when they cancel (the natural frequency) the force is proportional to damping, which is 90° out-of-phase.

The Bodé plot

There is another way we can visualize this characteristic. It is called a Bodé plot. The Bodé plot includes a graph of magnitude versus frequency, and a second graph of phase versus frequency. We can see how the amplitude and phase change as the frequency is changed.

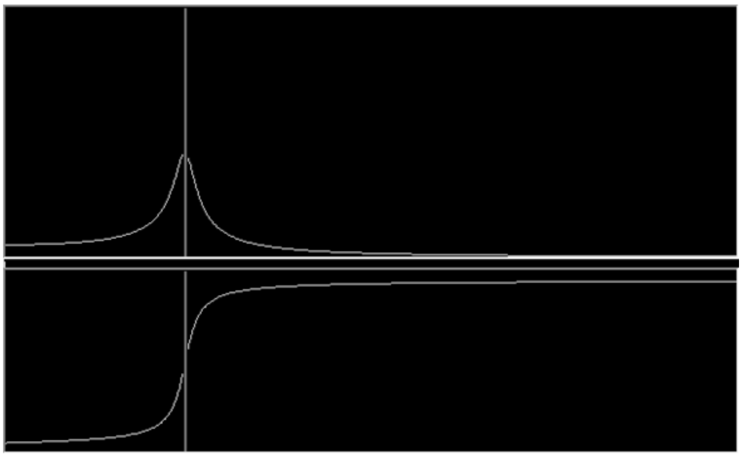


Figure 5-29

The Bodé plot is very helpful, as it shows us how the amplitude will change; the location of the natural frequency, and the effect of damping. For example, in Figure 5-30 damping is set to 0.025.

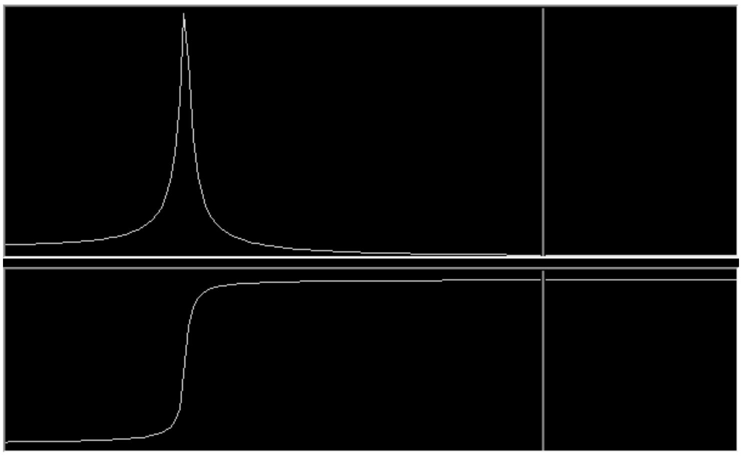


Figure 5-30

In Figure 5-31 it is set to 0.25.

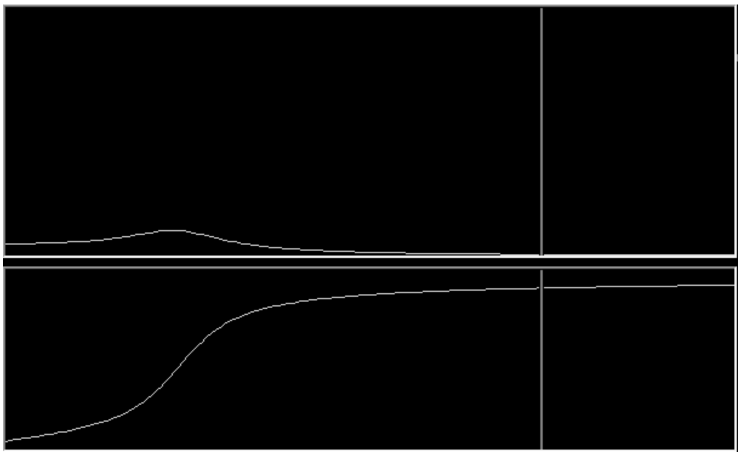


Figure 5-31

You can see the effect that damping has on the amount of amplification, and the rate of change of phase. If the system is lightly damped, the phase will not change until the frequency is close to the natural frequency.

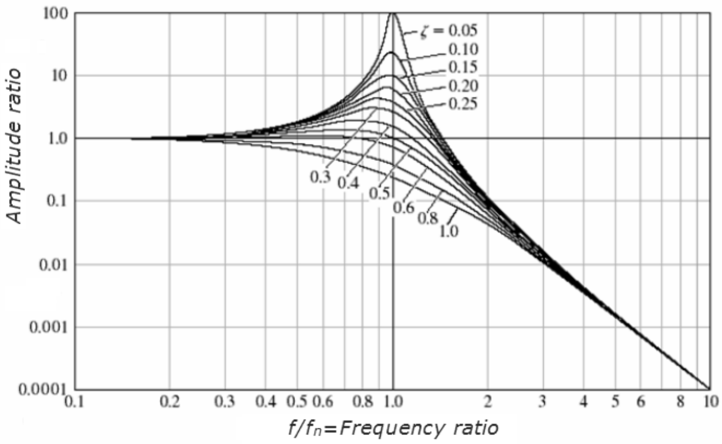


Figure 5-32

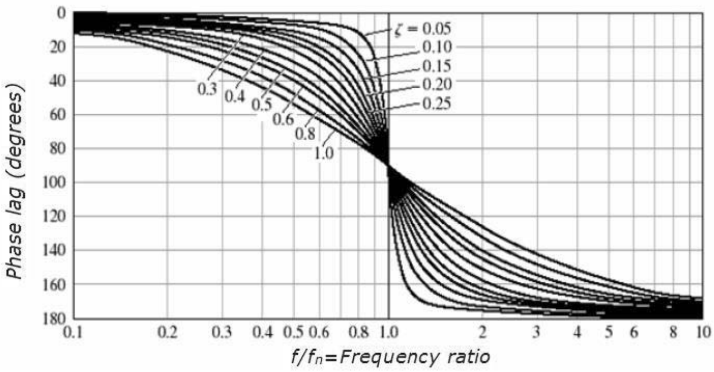


Figure 5-33

The Bodé plot provides a very useful way to display amplitude and phase data versus frequency. Later we will discuss run-up and coast-down tests where we excite a machine through a range of frequencies. The Bodé plot is key because not only do we see how the amplitude changes as the machine speed passes through a resonance, but we also see how the phase changes. The phase must change by 180° for it to be a resonance.

Dynamic response of a rotor

Let's take a look at how the system response affects a rotor in a machine. Now we are looking at the shaft from an axial view point. We have added an amount of mass to create unbalance centrifugal forces. And we placed the reflective tape at the heavy spot – the point at which the out-of-balance mass exists. We will discuss phase measurements in more detail at another time, but we have placed the displacement probe in line with an optical tach. What would we measure? You would expect that the phase reading to be 0°.

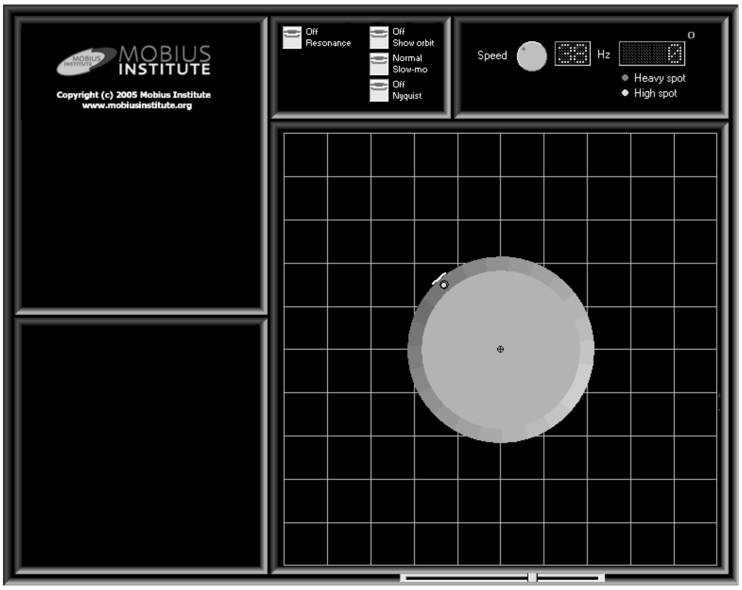


Figure 5-34

In actual fact, the reading might be slightly different for a whole variety of reasons – there will be a slight lag between the heavy spot (the unbalance mass) and the high spot that we register with the displacement probe.

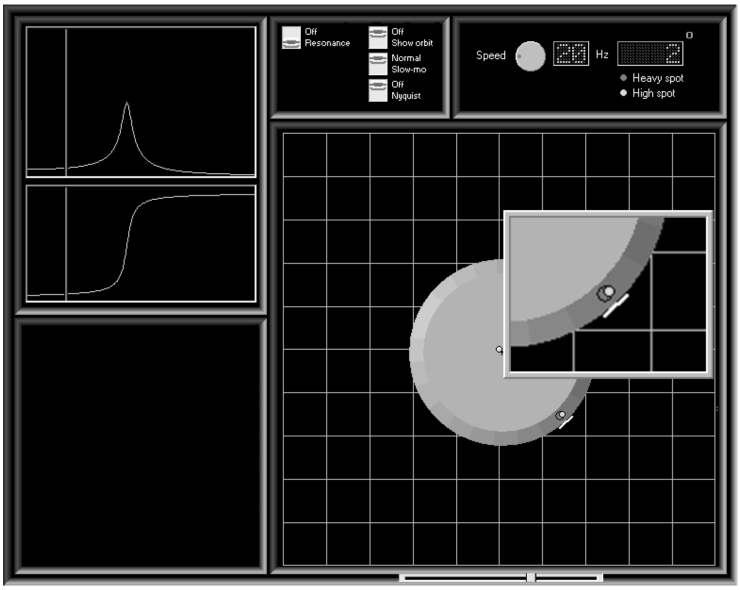


Figure 5-35

Let's just step back a moment and make sure we understand the "heavy spot" and "high spot". The high spot has been defined as "the angular location on the shaft directly under the vibration transducer at the point of closest proximity." The displacement probe is sensitive to the high spot. The heavy spot has been defined as "the angular location of the imbalance vector at a specific lateral location on a shaft." As we will learn, they are not the same angular location on the rotor.

As we increase the speed of the machine we naturally see the shaft turn more quickly, but we will also notice a separation between the heavy spot and high spot. Of course, the actually heavy spot does not move, but to the displacement probe it will appear as if it does.

Up to a certain frequency, which is approximately 60% of the natural frequency, the rotor acts like a rigid rotor. There is only a small difference between the heavy spot and high spot, and the force is centrifugal. The heavy spot is at the outer radius of rotation. If we add the orbit display you can see that the heavy spot is traveling at the outer radius.

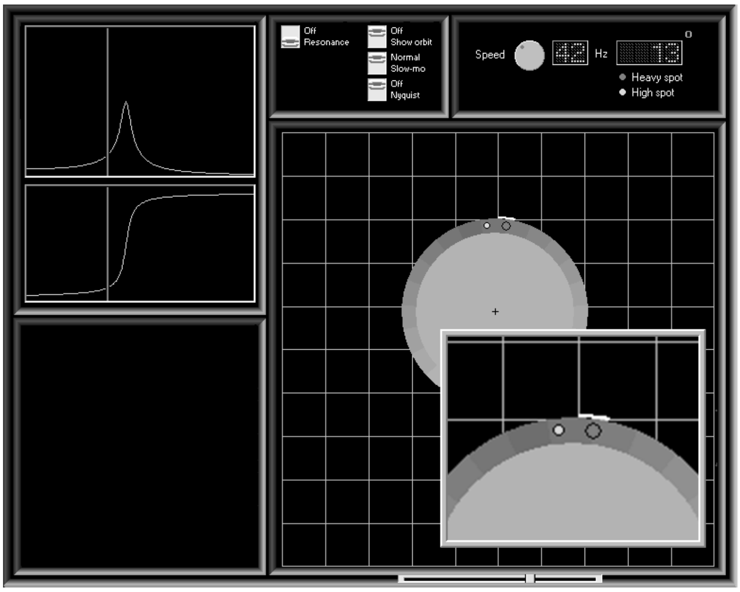


Figure 5-36

Another way to visualize what is happening is to imagine a roll on a balancing machine. Imagine that we have deliberately added weight to the inside of the roll such that we know the location of the heavy spot. As the roll increases in speed the rotation will follow the circular path shown. If we were to take a piece of chalk and move it just close enough to the roll so it barely touches, and then stop the roll, we would see a chalk mark at the heavy spot. At this stage it is acting as a rigid rotor.

Now let's increase the speed. As we increase the speed, the amplitude will increase quickly (depending upon the damping), but the high spot will begin to lag behind the heavy spot more dramatically. At the natural frequency, the high spot will lag the heavy spot by 90°.

If we went back to our example with the roll on the balancing machine, the chalk mark would be a quarter rotation *behind* the actual heavy spot (90° against the direction of rotation).

Now we can increase the speed of the rotor even further, as we pass the critical speed the lag will increase still further until it is 180°. Now the motion looks unusual. Now the shaft is rotating around the center of mass of the rotor. The center of mass is the point about which the mass is evenly distributed.

If we imagine the roll on the balance machine again, the chalk mark would be *opposite* the heavy spot. If we were actually attempting to balance the roll, we certainly need to be aware of where we are in relation to the natural frequency. If we are less that 60% of the speed then we could place a weight opposite the chalk mark in order to balance the rotor. If we are well above the natural frequency, we would place the weight at the chalk mark.

Nyquist Diagram

Thus far we have been looking at the Bodé plot. Another common plot is a Nyquist diagram; also called a polar plot. The Nyquist plot is a plot of amplitude versus phase angle as frequency changes.

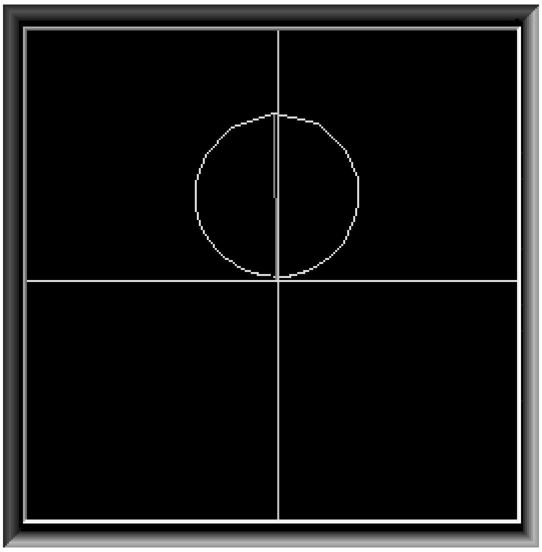


Figure 5-37

In this example you can see it looks like a simple circle. That is because we are studying a single-degree-of-freedom system and there is only one resonance. Remember, as we pass through a resonance the phase angle swings by 180°. Although you may think that a circle requires 360°, you can see that the amplitude increases to a maximum level after the vector has swept through 90° and then it sweeps through a further 90° as it passes to the other side of the natural frequency, thus creating a circle.

If we were dealing with a multi-degree-of-freedom system, the Nyquist diagram would have multiple circles; one for each mode. But we'll discuss this again later.

Multi-degree of freedom (MDOF) systems

Now we are going to explore multi-degree-of-freedom systems, and discuss modes and nodes. Although we have thus far looked at simple single-degree-of-freedom systems, in the real world we will face multi-degree-of-freedom systems. As stated previously, we can look at a multi-degree-of-freedom system as a series of single-degree-of-freedom systems, but there are a number of important issues to consider.

Let's start with a beam that is pinned at each end. We see structures around all plants that look like this, and even a shaft pinned at either end by bearings can be modeled this way. We will excite the structure at a single frequency. As we increase the frequency and get closer to the first natural frequency we will see the beam respond.

Figure 5-38

Well before we excite the beam at its natural frequency we will see the vibration levels increase. At the natural frequency the level of deflection is greatest. And as we increase the frequency further the amplitude level (amount of deflection in the center of the beam) will be reduced.

Earlier in this section we experimented with a mass, a spring, and a damper. Now we are dealing with a real structure. The beam also acts like a spring, mass, damper system. If we made the beam thicker, it would become stiffer and its mass would increase. We can change the thickness of the beam, the width of the beam, and the length of the beam and we will see the natural frequency change. If the system is stiffer, the frequency will go up. If it is heavier, the natural frequency will go down. But no matter how we change these physical characteristics, we will still be able to find an excitation frequency that makes it vibration in the same characteristic way.

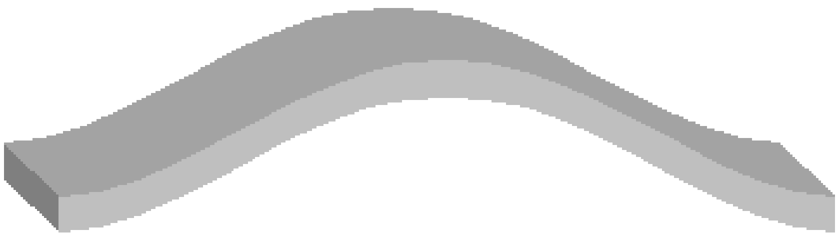


Figure 5-39

As we change the length, width and thickness, we change the mass and stiffness, but we do not change the damping, so although the natural frequency will change, the amount of vibration will not.

If we artificially alter the amount of damping, the frequency will change slightly, but the amplitude will change in proportion to the increase or decrease in damping.

$$f_d = f_n \sqrt{1 - \zeta^2} \quad \zeta = \frac{c_v}{c_c}$$

Modes and nodes

The beam is actually a multi-degree-of-freedom system. That means that it has more than one natural frequency. What we have seen thus far is the "first bending mode".

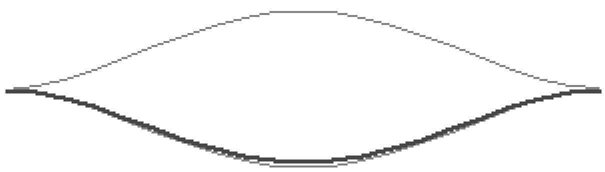


Figure 5-40

If we increase the excitation frequency, the vibration we observed will be reduced. But as we continue to increase the frequency the vibration will begin to increase again – but this time the pattern is different.

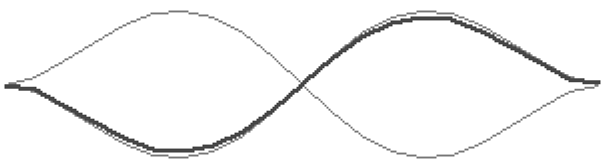


Figure 5-41

The amplitude level will probably be lower, but you can clearly see a difference in the "shape" of the beam. This is called the "second bending mode".

We can introduce another term at this time: node (and anti-node). The point mid-way along the beam is called a node (and there are nodes at each end). You can see that there is no movement at the mid-point or at the ends. 25% and 75% along the beam you can see the movement is the greatest. This point is called the "anti-node".

If we increase the excitation frequency still further, the second-bending-mode vibration will die away and then a new mode shape will appear.

Figure 5-42

Now we have four nodes and three anti-nodes.

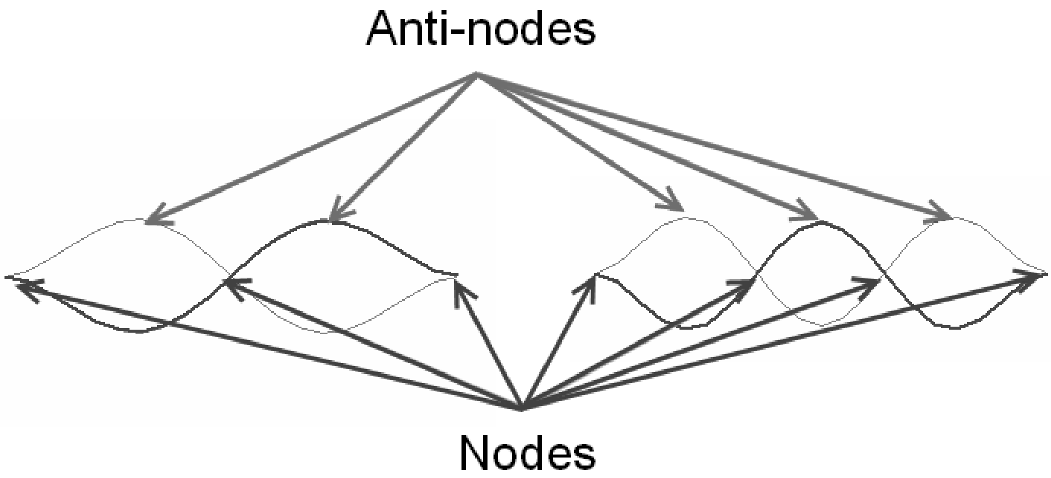


Figure 5-43

The difference between the node and anti-node is obvious to look at, but if you ever intend to test for resonances, or correct resonances, then the difference is very important. If the structure was being excited at the frequency that corresponds to the second bending mode and you measured the vibration at the nodal point in the center of the beam, what would you measure? A spectrum may be quite complicated (depending upon how it is being excited), but the amplitude at the frequency of the second bending mode would be zero. If you were to insert a new column under the beam at that point (midway along the beam), the vibration along the beam would not change because it was never vibrating at that point (assuming the addition of the column did not add to the mass of the beam or change the damping).

So the message is that measurements taken at nodal points will not reveal information about the mode; forces input at the nodal points will not excite the mode; and modifications made at the nodal points will not affect the mode. We will explore this further in the section on testing and correcting resonances.

There are a variety of ways that we can test structures to better understand the frequency of each mode, and the amount of damping of each mode. The two graphical methods we have already used will reveal the presence of the modes via changes in amplitude *and* the changes in phase.

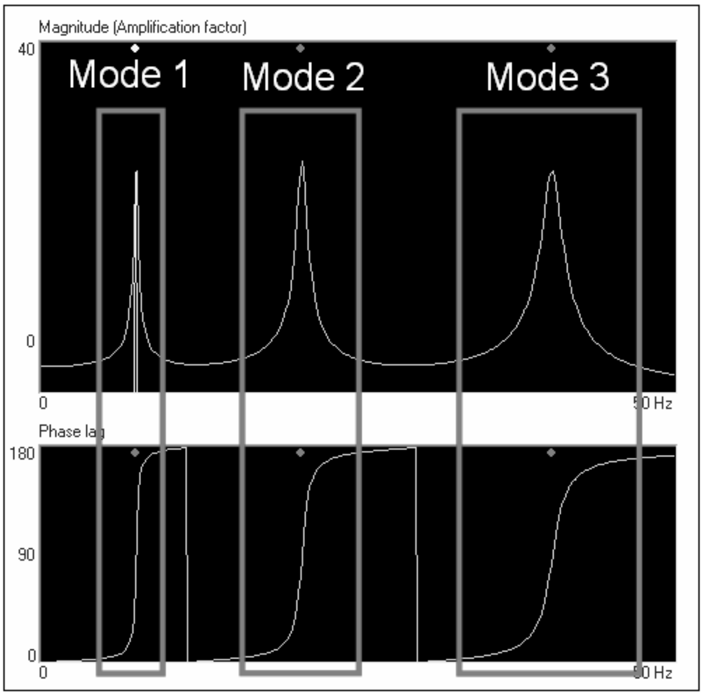


Figure 5-44 - Bodé plot

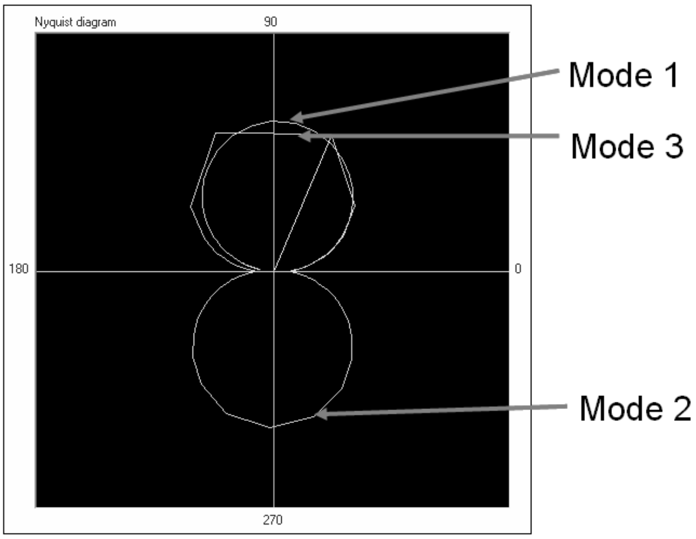


Figure 5-45 - Nyquist diagram

Why is this important?

If you are dealing with turbines then you will have to consider the dynamics of the rotor. Balancing the rotor, and managing the run up and coast down of the turbine passes through the natural frequencies are very important. But that is the subject of the Category IV course.

Dynamics, system response, transmissibility, and resonances are issues that every vibration analyst must face. Whether you are considering how to mount an accelerometer; how to balance a machine; how to collect good phase readings; how to diagnose bearing faults, or how to solve structural resonances; the underlying dynamics of the machine and supporting structure are very important. For now we are going to focus on resonances.

Chapter 6

Resonance and Natural Frequencies

Objectives:

- Describe common tests for resonance including run-up / coast-down, bump tests and calibrated hammer tests
- Describe common data types used in resonance testing including the FRF, coherence and transmissibility, waterfall and Campbell diagram

This chapter builds on the prior chapter and discusses practical test techniques for diagnosing and detecting resonance conditions. Tests include run-up / coast-down, bump tests and calibrated hammer tests. When conducting these tests, certain data types are typically analyzed. These are also covered in this section and then expanded on in the next two sections on modal analysis and ODS.

How Can You Tell If You Have A Resonance Problem?

Without doing any special tests, there are two basic ways to tell if you have resonance problems: unusual mechanical failures, and tell-tale signs in the spectrum.

Unusual Failures

If you have machinery failures or structural failures that seem to be as a result of fatiguing and you do not have any other explanations, then you should consider resonance as a possible cause. Structures should last a very long time, and fatigue failures should only occur after MANY years of service. Examples of failure modes include:

- Broken welds
- Cracked and leaking pipes
- Premature machine failure
- Broken or cracked shafts
- Foundation cracks

In the following example, high vibration due to resonance caused the hand rail so slowly wear away.

Figure 6-1 (Images from Simon Hurricks, Genesis Power, NZ)

Tell-Tale Signs in the Spectrum:

There are four tell-tale signs:

Unusually high peaks in the spectrum.
- Peaks are amplified

Directional vibration.
- High vibration levels in one direction/axis but not in another.

"Haystacks" and "humps"
- Areas in the spectrum where the noise floor, and any peaks in the vicinity, seem to have been raised.

Peaks that change amplitude when machine speed changes.

Because resonances amplify vibration, peaks in the spectrum that may normally be quite low (at the blade pass frequency, or gear mesh frequency, etc.) can increase dramatically. When you see a high peak, remember to consider the possibility that it is being amplified by a resonance.

If you have a variable speed machine, you may observe that when the machine is operating at its highest speed, the blade pass frequency may coincide with a natural frequency, so the machine resonates, and the blade pass frequency peaks will increase in amplitude. However, you may also observe that when it is operating at a lower speed, the blade pass frequency would be below the natural frequency, so the amplitude would drop to "normal" levels, and the machine will not be in resonance.

Also, when you study the structure more closely, you will notice that the mass, stiffness and damping can be very different vertically, horizontally and axially. So, there may be a resonance in the horizontal direction (i.e. it is vibrating from side-to-side), but not so vertically (there will be natural frequencies, but they may not be excited). In the case of vertical pumps, however, resonances can quite often exist radially, but not at the same frequency vertically (axially along the machine).

"Humps" and "Hay Stacks"

As we have just seen, any vibration that coincides with a natural frequency will be amplified. And, as we have just seen, it does not have to be a perfect match. In Figure 6-2 you can see the "hump" in the spectrum from 30 to 65 Hz. It looks like the 3X peak is being amplified.

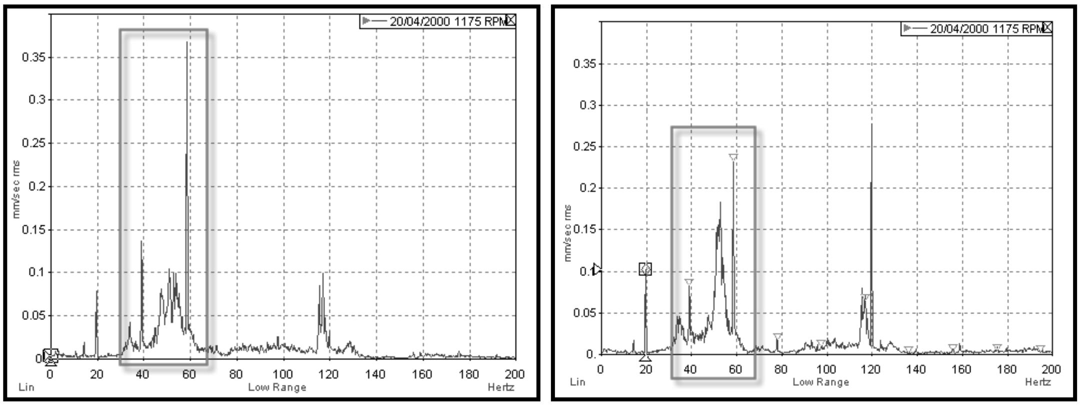


Figure 6-2

The data above came from the vertical pump pictured below. You can see the photo of the installed machine (where the data was taken), as well as an old machine that had previously failed (Figure 6-3).

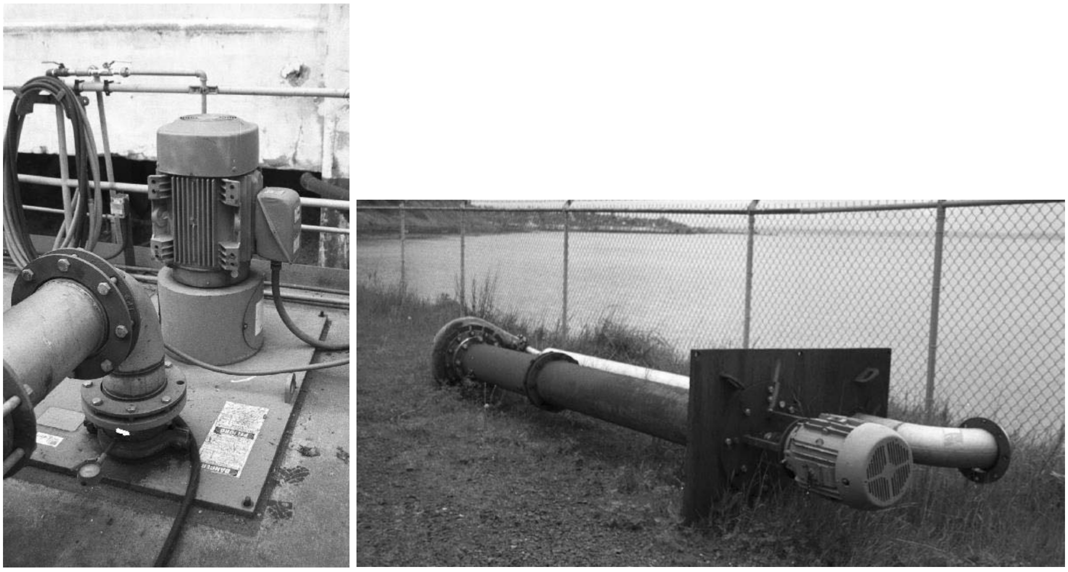


Figure 6-3

Special Tests to Identify Natural Frequencies

If you suspect that a machine is resonating, but you need to better understand the situation before you attempt a correction, then there are a few tests that you can perform that will show where the natural frequencies are located, and give you an indication of how the machine/structure is moving.

In order to understand the resonances, we need to "excite" them – we must inject energy at either all frequencies or just the frequencies of interest and see how the structure responds. Multiple natural frequencies exist in all structures. If we excite them we can see where they are. We will discuss a number of ways to perform this function.

Note: Remember that machines do move in three axes. Unless you already suspect that the problem is predominantly in one direction (for example the motor/pump is rocking from side to side), then you should perform these tests while collecting data in two or three axes.

Note: When attempting to excite resonances, be aware of the possible location of nodes. When you perform these tests, you must be careful not to inject energy at a nodal point (e.g. during bump or impact tests), or measure vibration at a nodal point, as there will be no vibration to measure. If you feel along the structure, you can often tell where the nodes are located. You can do the same with an accelerometer while watching the collector in real-time mode.

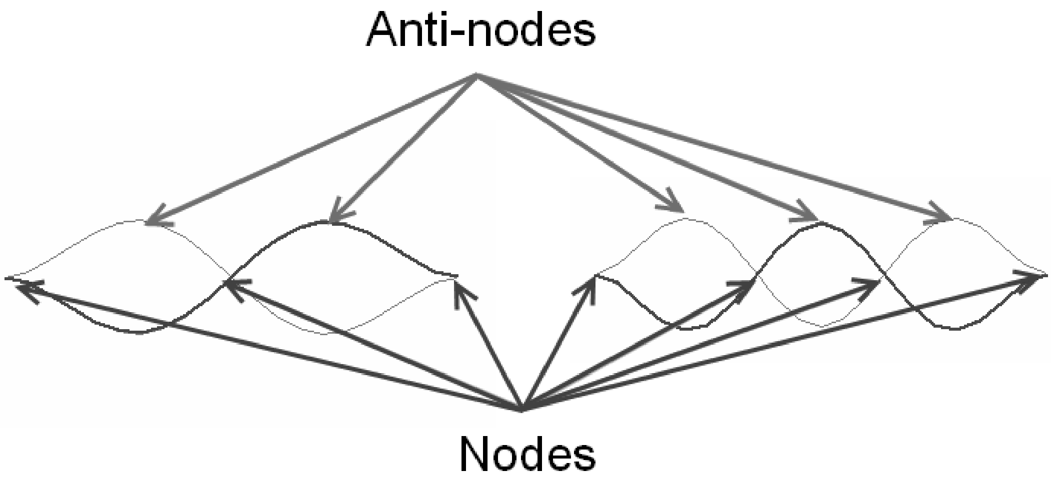


Figure 6-4

Changing Running Speed

We can use the machine itself to inject vibration energy into the machine and structure. If it is possible to change the speed of a machine, set it at its highest speed and slowly reduce the speed while watching the vibration levels in the spectrum. If possible, use the peak-hold averaging function so that only maximum amplitudes are maintained.

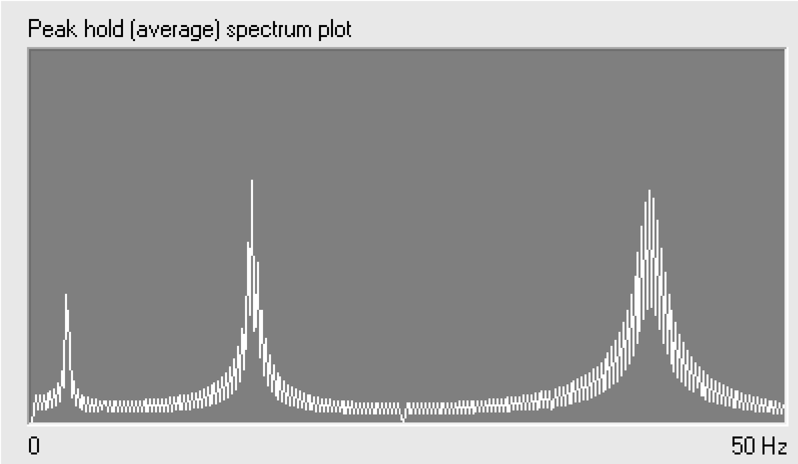


Figure 6-5

In the example above (Figure 6-5) we were able to sweep across a broad range of speeds. In the example below, the speed could not be changed over such a wide frequency range. While we have revealed the presence of a natural frequency, more information would be useful.

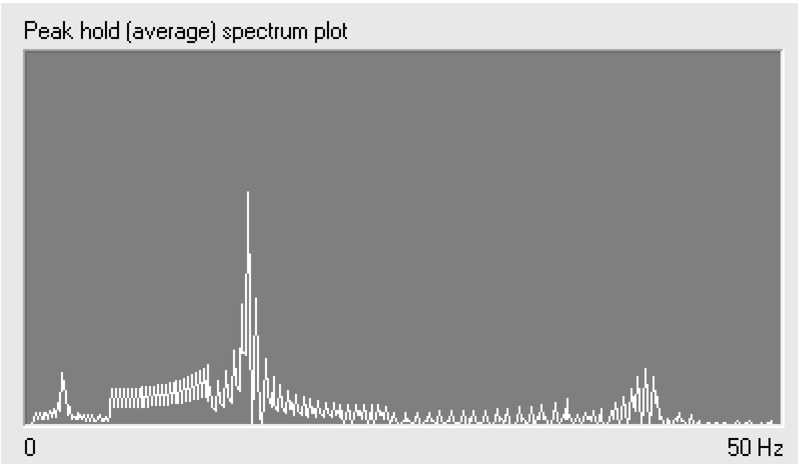


Figure 6-6

Bump Test

If you were to strike a machine with a large piece of timber, energy would be injected into the structure at all frequencies. Why is that so, I hear you ask? If you were to look at a spectrum of an impulse, instead of seeing the traditional peaks, you would see that the amplitude across the entire frequency range is raised. That means that all of the natural frequencies will resonate, amplifying the energy at those frequencies.

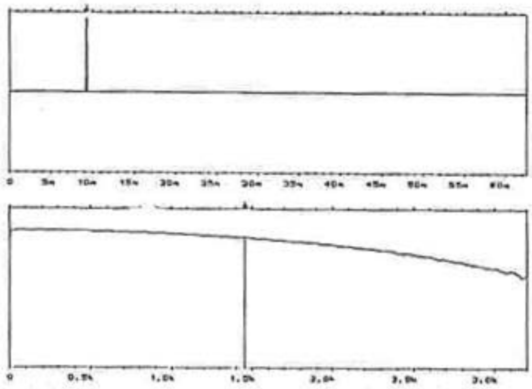


Figure 6-7

So, if you stop the machine, set the data collector to peak hold mode (20 averages), start data collection and immediately strike the machine again, the natural frequencies will be excited, and therefore the data collector will record vibration levels at all of the frequencies.

Note that when you strike the machine, it will vibrate for a very short time (depending upon the damping). Thus if you use 800 lines, auto-ranging and a low Fmax then the measurement will take too long. Resolution is not important, so select a higher frequency range (Fmax), a lower resolution measurement (400 lines or even lower), and if possible, set an appropriate gain setting and turn off auto-ranging.

You can also do this with the machine running, but you do need to be extra careful not to damage yourself or the machine, and you need to impact the machine with enough force that the increase in vibration amplitude can be measured.

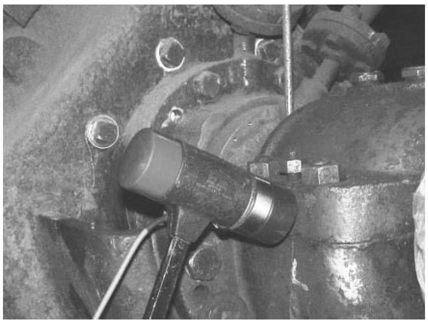


Figure 6-8

Note that a bump test will not always excite all of the frequencies, and the make-up of the machine and structure will affect the results considerably.

The act of striking the machine, and the tool used to strike the machine (a piece of timber suggested above), is a science unto itself. There are special hammers you can buy for this purchase. You see, if you were to strike the machine with a piece of metal, more energy will be input at higher frequencies, less so at lower frequencies (and you may damage the machine!).

The opposite is true for soft items. It you used a rubber mallet; you may not inject enough energy at the higher frequencies. This is discussed further in the modal analysis chapter.

The aim is to inject enough energy so that you excite the natural frequencies in the frequency range of interest (the range where your machine generates frequencies, or the range where you suspect the resonance). If we were to strike our cantilevered bar, we might get a graph like the one shown below. You can clearly see where the natural frequencies are located. You would then compare this graph to a normal spectrum to see if any of the dominant peaks match.

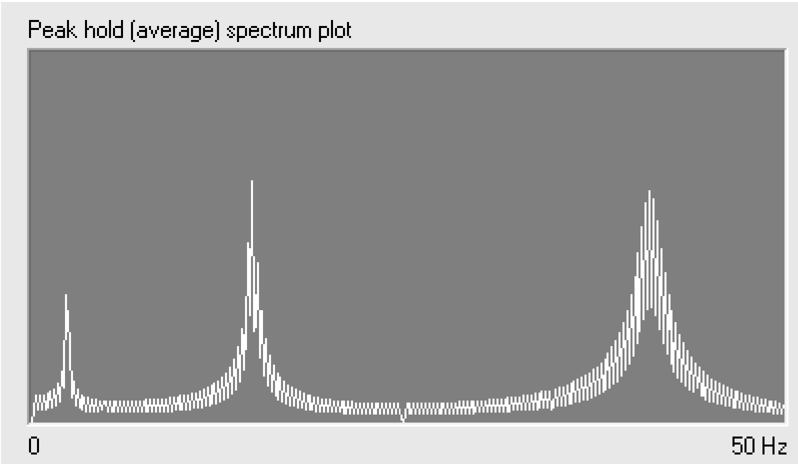


Figure 6-9

Using Negative (subtraction) averaging

Instead of using peak hold averaging when performing the bump test, you can use a method known as "negative (or subtraction) averaging". If supported by your analyzer, this can be a very effective tool. There are two stages to this testing method.

Phase one

When you are ready to begin the test, start the "negative averaging mode" on the analyzer.

Now bump the machine a number of times while the machine is running. The analyzer will be recording the normal machine vibration *and* the response to being bumped.

Phase two

In phase two of the test you continue to record the vibration, however now you do not bump the machine. The analyzer will be recording the normal machine vibration *without* the vibration related to the machine's response to the bump.

As the analyzer records the "normal" vibration it is able to subtract the new vibration from the old vibration that included the response to being bumped. The difference between the two sets of data is the response to being struck; all of the 1X, 2X, and other sources of "normal vibration" for the machine are removed.

If performed correctly the measurement can be quite effective at revealing the natural frequencies in the machine. Of course, it does not prove the existence of the resonances; we need to see the phase shift to be sure, but it is a very good indicator.

Run Up and Coast Down Tests

If you can control the speed of the machine as it runs up to speed or as it coasts down from operating speed, then as the machine slows down/speeds up, you can look for the tell-tale resonances.

For example, if you attempt a coast-down test, as the speed of the machine drops, the running speed vibration, or vibration at some other frequency (e.g. 2X, vane pass rate, etc.), may coincide with one of the natural frequencies, and it will resonate. You will see the amplitude increase.

If the machine comes up to speed too quickly (or stops too quickly in a coast-down), then you will not have time to collect sufficient data, nor will there be enough time to excite the resonances. As described above, you need to set up the data collector to take measurements quickly – opt for lower resolution spectra.

You can use peak hold averaging, but better yet you should use a waterfall plot (often called a cascade plot or spectral map). Whereas we are familiar with waterfall plots showing individual spectra from different dates, in this case the spectra are collected a few seconds apart (or longer, depending upon the rate of change of speed).

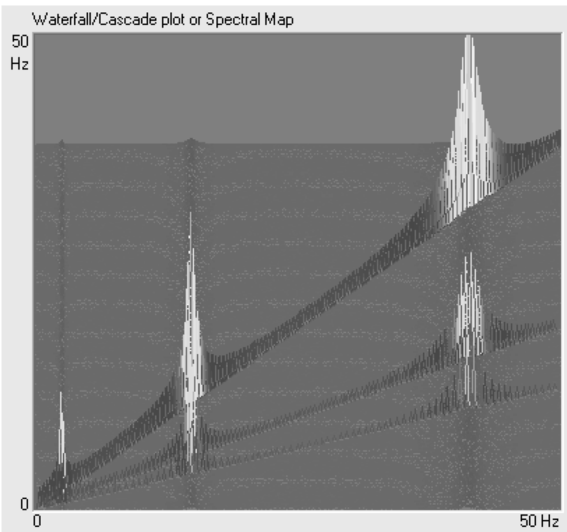


Figure 6-10

There are two interesting characteristics. The three lines of peaks radiating out result from a strong 1X, 2X, and 3X peak in the source vibration. As the speed of the machine increases, the 3X peaks excites the first mode (and is amplified), then as it moves out the 2X moves in, and finally the 1X frequency excites the first mode.

You can look at the three regions up the graph to see clearly where the resonances exist. Even when the dominant frequencies from the machine do not coincide with a natural frequency, other sources of vibration, including noise, will excite them.

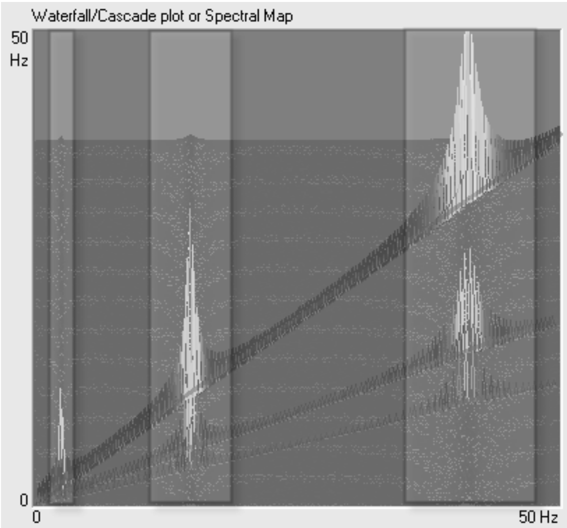


Figure 6-11

Order Tracking

If your data collector supports order tracking, it is a test you can perform while the machine is running up to speed or coasting down. The data collector will take a once-per-rev tachometer signal so that it knows the exact running speed. It extracts the amplitude and phase data as the speed is changed.

If you consider our waterfall data, it is the same as extracting the amplitude and phase from the 1X frequency.

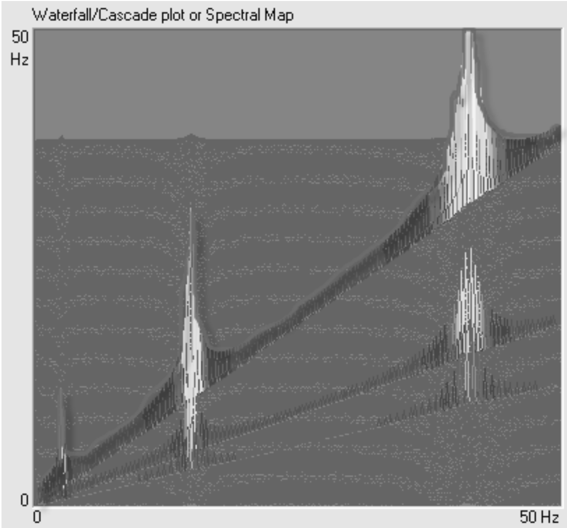


Figure 6-12

The data should look like the Bodé plot we described earlier.

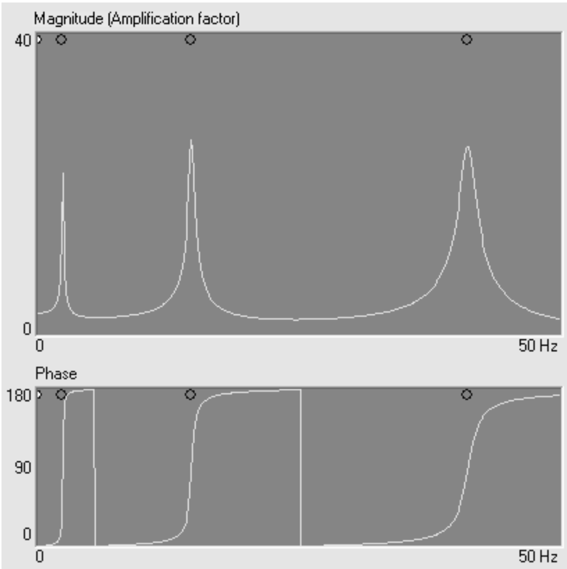


Figure 6-13

Case study: DC Motor at a Printing Press

A DC motor in a printing press was mounted on an I-Beam. Isolators were used to suppress the vibration amplitudes, however high vibration readings of >30 mm/sec or 1.7 in/sec were recorded, and when the machine was shut down the vibration would grow higher.

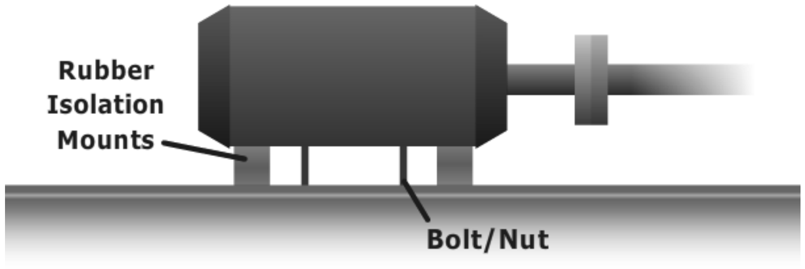


Figure 6-14

It was suspected that there was a resonance; in fact it was suspected that the isolators themselves were responsible.

A run-down test was performed. As the speed was reduced, the amplitude climbed as high as 45 mm/sec or 2.5 in/sec. As can be seen in Figure 6-15, the amplitude rose and fell away quickly, suggesting the possibility of a resonance.

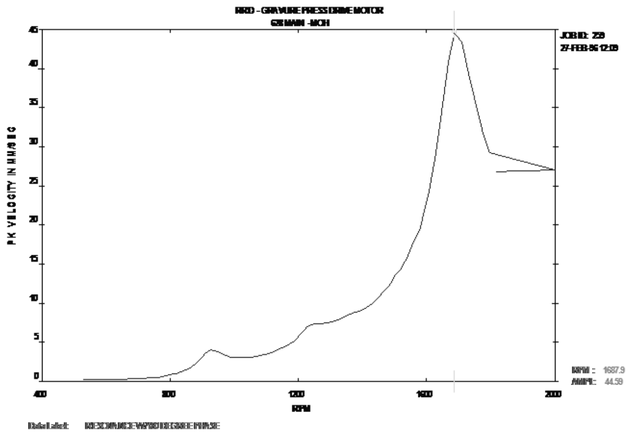


Figure 6-15

When the phase data was studied it was clear that we were dealing with a resonance – you can see the 180° phase shift.

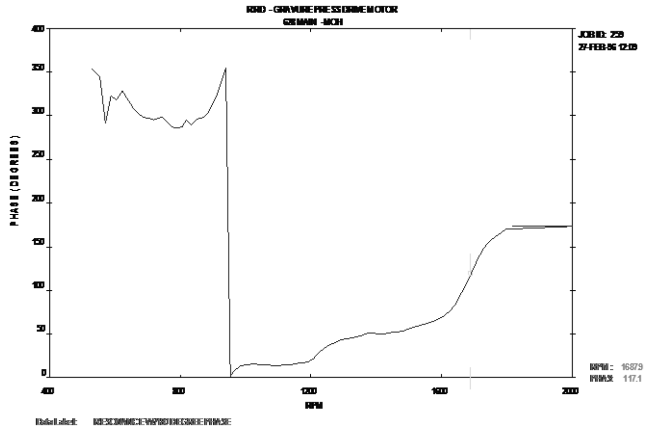


Figure 6-16

Isolators are designed as single-degree-of-freedom systems. If designed and installed correctly, they will cause the vibration levels to be greatly reduced, as long as the operating frequency of the machine is substantially above the natural frequency of the isolator. It appeared that the natural frequency of these isolators was too close to the operating speed of the machine – thus amplifying the vibration!

A bump test was performed to double-check the results. They also indicate the likelihood of a resonance at around the running speed (note the change in frequency scale).

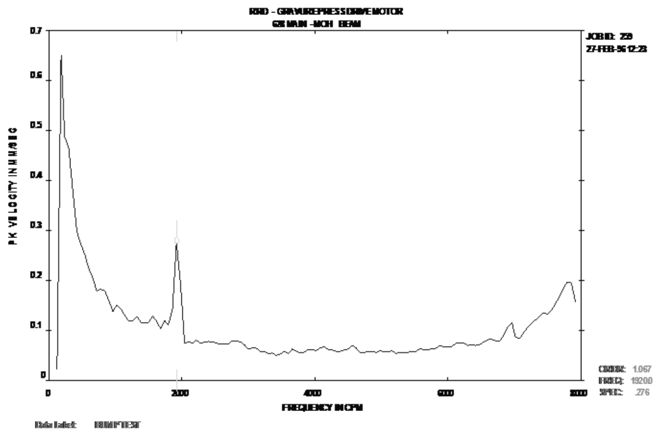


Figure 6-17

To negate the amplifying effect of the isolators, shunts were added to the machine; jacking bolts between the underside of the motor and the supporting I-beam.

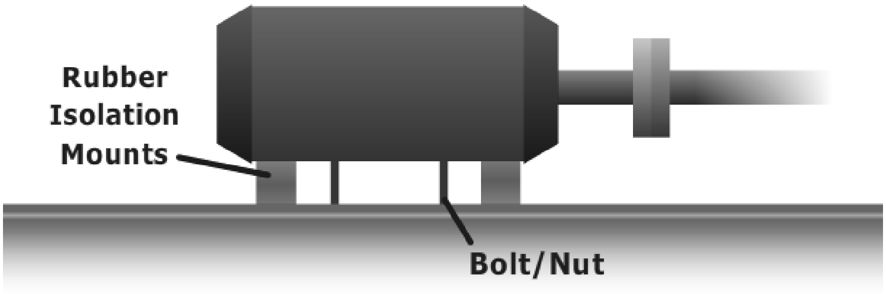


Figure 6-18

The run-down test was repeated and dramatically lower vibration readings were recorded.

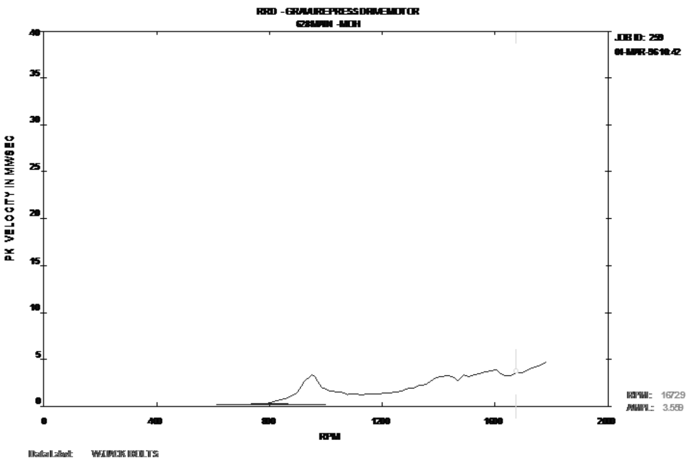


Figure 6-19

The peak amplitude is now just 3.5 mm/sec, or 0.2 in/sec.

Cross-channel measurements

Thus far we have used amplitude and phase in the Bode plot only. The phase readings were based on a tachometer reference on the machine. A single-channel analyzer can be used to acquire the data quite easily (although reflective tape must typically be added to the shaft), however the one major limitation is that phase data is only available at the running speed of the machine. When creating Bode plots (run-up or coast-down tests) the 1X amplitude and phase are sufficient, but when we take a close look at Operating Deflection Shape (ODS) analysis, and Modal Analysis, cross-channel measurements that provide amplitude and phase values at all frequencies will be required.

We will now take a closer look at cross-channel measurements.

The most common cross-channel measurements include:

- Cross-channel phase
- Transmissibility or Frequency Response Function

- Coherence
- Orbits (we will discuss orbits separately)

Understanding cross-channel measurements

When an FFT calculation is performed the result is both a spectrum of amplitude and a spectrum of phase values. For example, if the spectrum had 800 values (800 LOR), there will also be 800 phase values. In many machinery vibration monitoring applications the FFT it calculated from a single time waveform. The phase values are referenced to the start of the time waveform, however because the time waveform was not triggered there is no reference. As a result the phase values are meaningless.

Note: Some studies have been performed that suggest that these phase values can actually be of some value – but that is not the subject of this chapter.

Because the phase values are not relative to a meaningful reference they are discarded. We only save the amplitude values.

However, if we collect two time waveforms simultaneously and perform the FFT calculation on both time records, we can use the phase values of both records. The phase values are relative to each other. For example, if the phase value at 1X in channel "1" is 23° and the phase value at 1X in channel "2" was 113°, then it could be said that the vibration at the running speed was 90° out-of-phase between channel "1" and "2". The values of 23° and 113° are meaningless.

Applications of cross-channel phase

When we discussed phase analysis we learned that the easiest way to collect the phase readings was to use cross-channel phase. However in that case we only used the phase at the running speed – we were trying to diagnose unbalance, misalignment, etc. When we collect cross-channel phase we actually have a full spectrum of amplitude and phase readings. This far we have not had an application for this data, however when we discuss ODS and modal analysis we will use this data. The difference is that we will be studying cross channel phase *and* amplitude.

Force-response tests

When studying phase data from a machine, we are measuring the response of the structure to the internal forces that are a result of the shaft turning (and residual unbalance and misalignment, etc.), the gears meshing, balls rolling (in the bearings), vanes and blades being pushed through some medium, etc. In fact, when we perform phase analysis and collect 1X amplitude and phase, we are really only looking at the machine's response to the 1X forces: unbalance, misalignment, bent shaft, etc.

However, we have no way of knowing how much force was used to generate that response. So when performing phase analysis and ODS, we are only able to measure how the machine is vibrating due to the internal forces.

There is a lot more we could learn if we could also measure the force that was used to generate the response. We could then compare the output to the input – the response due to the excitation. We can use two channel measurements such as the Frequency Response Function and Transmissibility to compare the response to the excitation, and we can use cross-correlation and coherence to compare the response to the excitation in order to check the validity of the test.

Transmissibility and FRF

From a mathematical point of view we can use the term transmissibility and Frequency Response Function (FRF) interchangeably. They are both a ratio of the output to the input.

- Transmissibility is often displayed on a linear scale, but can be displayed on a log scale. Transmissibility is typically used to compare two vibration signals, i.e. two sensors mounted on a machine. It indicates how the vibration at position 2 compares to the vibration at position 1. As we will see, it is used in ODS testing with one sensor kept at one reference location while another sensor is moved around the machine or structure.

- The frequency response function is normally used when comparing the vibration at a point due to a measured input of vibration (or force) at another point. It is used to measure the response of the structure to the input excitation. FRF is typically graphed on a log scale.

They are calculated the same way; the output divided by the input.

(In fact, in the electronics world they use the term Transfer Function for the same calculation. In the electronics world it is used to compare the electrical signal at a point compared with the input signal – capacitors, resistors and inductors also create resonances and anti-resonances.)

FRF is the ratio of the output to the input. If we could excite a structure at every frequency with 1 unit of vibration, and we measured 1 unit of vibration at every frequency, then the FRF would be as follows.

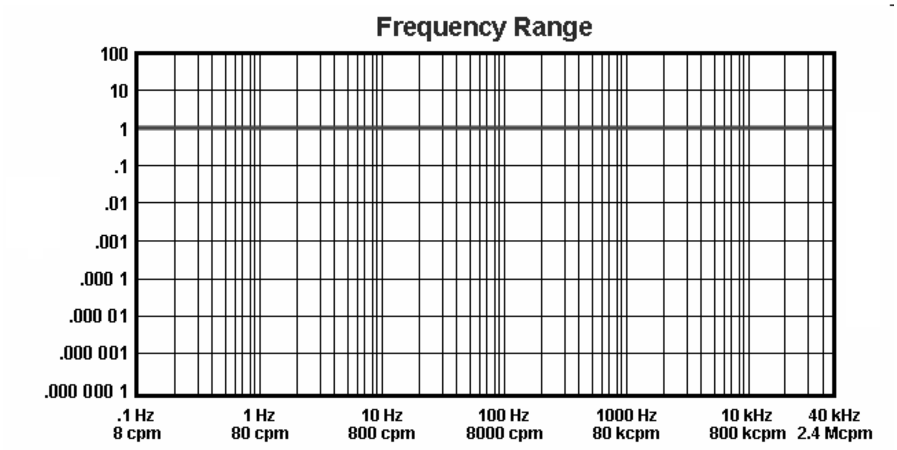


Figure 6-20

When we studied accelerometers and mounting options (in Category II) we learned that at certain frequencies the response was amplified (around the natural frequency) while at other frequencies the response was attenuated (roll-off). This was a frequency response function measurement.

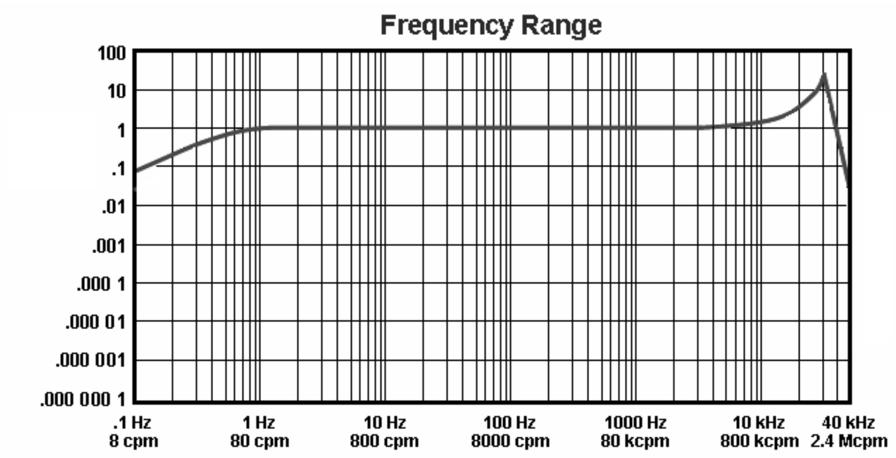


Figure 6-21

The ratio is often expressed in dB and described as 'gain'.

10 Log10 (Output/Input) – for voltage signals
20 Log10 (Output/Input) – for power signals

- Log10 (1) = 0
- Log10 (10) = 1
- Log10 (100) = 20
- Log10 (.1) = -1
- Log10 (.01) = -10

The same type of measurement can be collected on a structure where we do not measure the input forces – as stated previously, if the machine is operating the input forces are coming from the machine itself. But we can still compare the vibration amplitudes and phase from all around the machine or structure to a reference location to see how one location moves (amount of movement and direction of movement) compared to another location.

We will discuss this testing technique in greater detail in the ODS chapter.

Linearity and coherence

All of the cross-channel calculations performed on a multi-channel analyzer assume the system is "linear". What that means is that if you excite a structure with a source of vibration, all of the energy at all of the frequencies will be transmitted to all other points of interest on that structure. We know that resonances will either cause vibration at certain frequencies to be amplified or attenuated, however if a structure is *not* linear it means that:

1. The vibration energy is not transmitted to all points of interest; either because the points are not physically connected, or there is a crack, looseness or some other discontinuity.

2. If you were to double in the input vibration amplitude the vibration at the output would not be doubled.

3. If some of the vibration measured is not related to the vibration input into the system. For example there is electrical noise, vibration from other machines, etc.

Most mechanical systems are "linear enough" for good cross-channel calculations. However it is important to test for linearity.

The coherence measurement

The coherence function γ^2 indicates the degree of the correlation between the input and output of a system. It gives a value from 0 to 1 for each frequency.

- When γ^2 is 1, it indicates that at frequency f all outputs of the system are due to the input.

- When γ^2 is 0, it indicates that at frequency f the output of the system is completely independent of the input.

Typically, an acceptable coherence value will be above 0.9 in a low background noise environment and above 0.7 in a high background noise environment.

Coherence γ^2 is calculated from the cross spectrum between the signals and the power spectra for each of the two signals.

The formula for coherence is as follows: $\gamma^2 = \frac{[cross{-}spectrum]^2}{pwrSpectrumA \bullet pwrSpectrumB}$

Coherence is used to validate the integrity of signals and to identify the source of vibration in the presence of many different (similar frequency) signals.

Coherence is an averaged function. As the number of averages increases, the value of the coherence will decrease. For the first average, all of the data is coherent (coherence = 1.0) therefore, more than one average is required. As a rule, use 4-10 averages for coherence data.

Coherence is an important tool used in modal analysis. If the data is not coherent, it is not possible to draw a conclusion about the response of the structure to the input force.

Coherence is also important in ODS. We are assuming that the vibration we measure at the roving accelerometer is due to the forces generated within the machine (and not from external sources), and that the transmissibility between the roving accelerometer and the reference accelerometer is not compromised. If it is suspected that non-linearities exist, you can test the coherence and physically look for cracks, gaps, etc.

Testing for non-linearities

We can use two accelerometers, one at the reference location and one at the point of interest, and perform a coherence test. Ideally the coherence would be '1' at all frequencies. In this example, the coherence is quite poor.

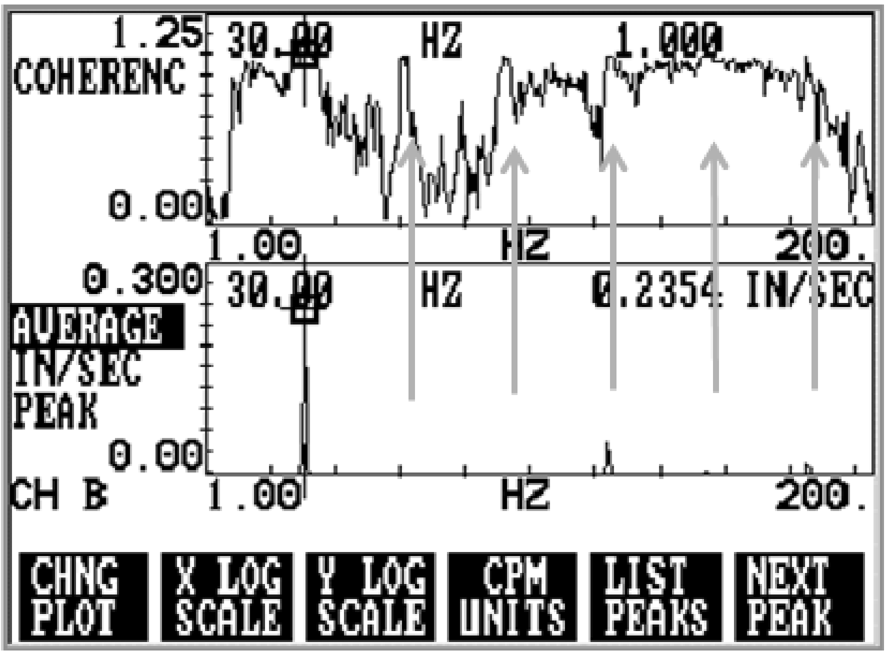


Figure 6-22

It is important to note that the coherence will always be low when the vibration amplitudes themselves are low or noisy. It is difficult to compare two sources of vibration at specific frequencies where there is very little vibration to measure. We are attempting to compare the vibration at point B to the vibration at the same frequency at point A. If there data is noisy or very low in amplitude, the correlation will be very low – thus the coherence will be low.

Using impact testing

Another way to test the coherence is to perform an impact test. The idea is to inject a known force and measure the response. You will need a modal hammer instrumented with the load cell, a response accelerometer, and a 2-channel analyzer.

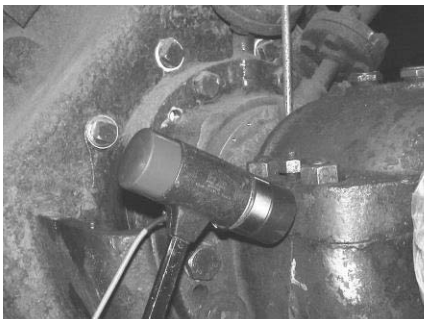


Figure 6-23

In the following example a test was perform to determine if there were non-linearities between the sole plate and the concrete foundations.

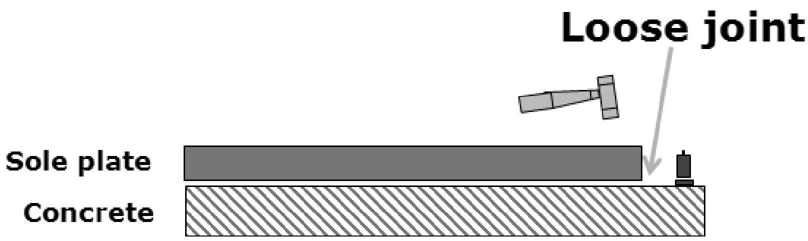


Figure 6-24

The analyzer is placed in coherence mode, and triggering is set up to capture the impact and response.

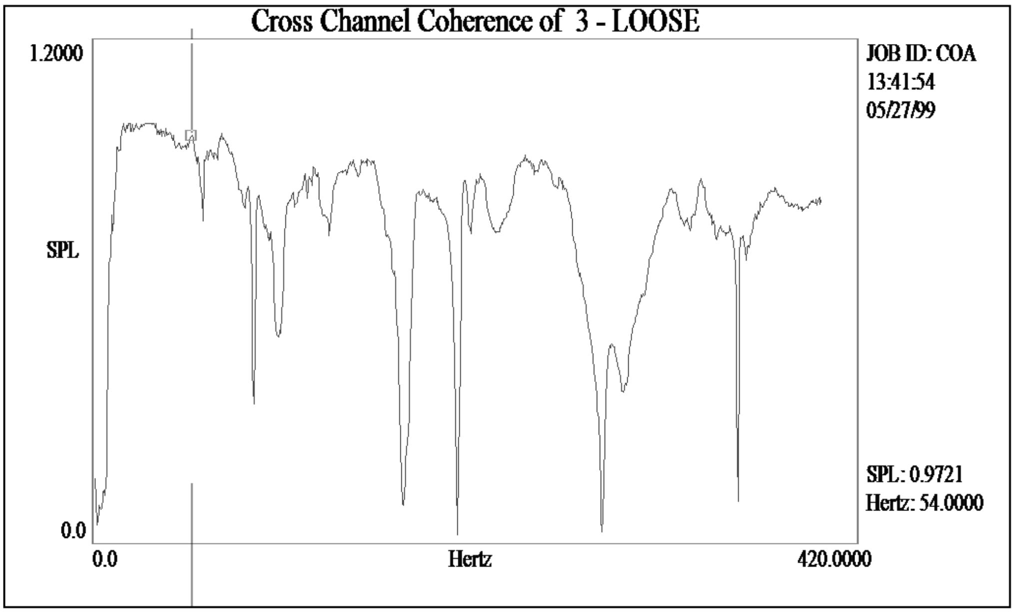


Figure 6-25

We can see that the coherence was quite poor. (Note that the graph is scaled to 1.20 and we are hoping for coherence of 1.0.)

A bolt fastening the sole plate to the concrete foundation was tightened by hand and the test was repeated.

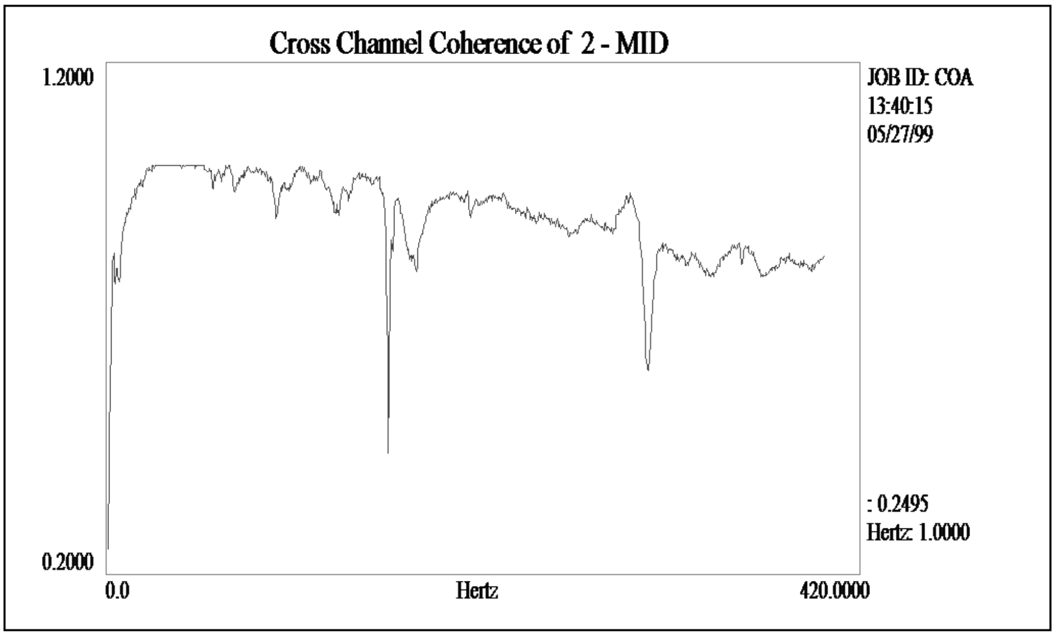


Figure 6-26

You can see that the results have improved. And finally the bolt was tightened correctly and the test was repeated for a final time.

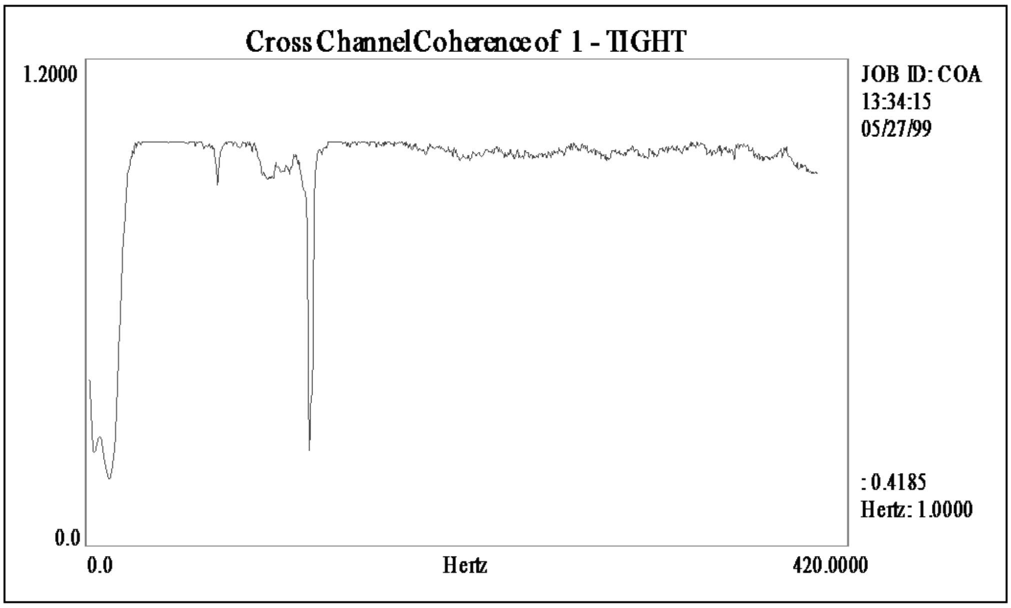


Figure 6-27

You can note a number of factors:

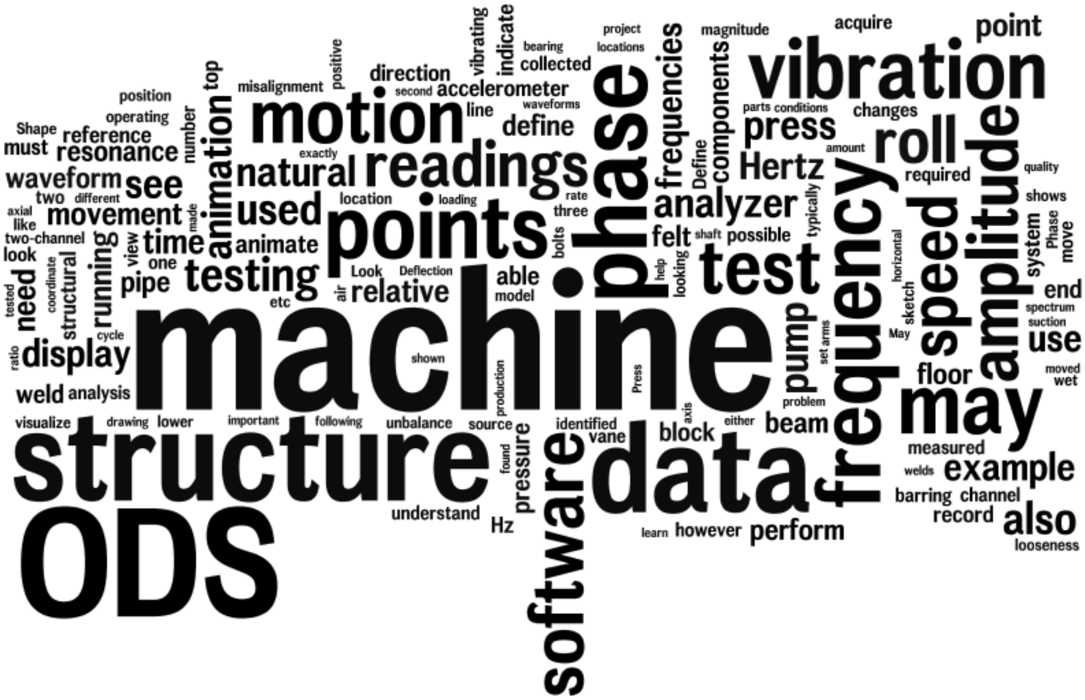
1. The coherence is '1' for most of the frequency range; so there is good linearity.

2. The low frequency response was poor. That could be related to the poor low frequency response of the accelerometer, or the ability of the hammer to inject energy at those low frequencies, or there could have been external noise.

3. At 120 Hz there is a dip in the coherence. That is due to the 120 Hz (twice line frequency in the US) hum from the lights and other sources. Because the hammer did not inject all of the vibration at 120 Hz that was measured on the foundations, the coherence is poor.

Coherence can be poor for a number of reasons:

- Low signal level
- Structural nodal points
- Background vibration
- Non-linearity
- Double hits with impact hammer
- Defective sensor or cable
- Poor sensor mounting



Chapter 7

Operating Deflection Shape Analysis

Objectives:

- Describe operating deflection shapes (ODS)
- Describe the ODS procedure for time based and frequency based ODS
- Provide an example ODS

In this chapter we will learn about the Operating Deflection Shape (ODS) testing technique. We will discuss why we perform ODS testing, and how to perform the test to get optimum results.

Introduction to Operating Deflection Shape (ODS) Testing

In this chapter we will learn about the Operating Deflection Shape (ODS) testing technique. We will discuss why we perform ODS testing, and how to perform the test to get optimum results. You should also refer to the operating manual that came with your ODS software – although we will discuss ways for you to perform basic ODS tests that does not require the use of sophisticated software.

Operating Deflection Shape testing helps us to understand the motion of the machine or structure while it is in operation under normal conditions. ODS testing is an extension of phase analysis techniques; we are acquiring vibration amplitude and phase from around a structure so that we may visualize the motion.

Why use ODS?

We typically perform ODS testing when we suspect that a machine is operating at a natural frequency, or when we suspect that a source of machine vibration (for example blade pass frequency) is exciting a natural frequency. We may observe high vibration levels, or observe structural failures (welds cracking, bolts failing, cracks forming, etc.) or we may witness premature failure of machine components, for example bearings failing quickly.

We can also use ODS as an excellent diagnostic tool. We can diagnose unbalance, misalignment, bent shaft, soft foot, structural looseness, and other faults – all of which can be difficult to diagnose with spectra and time waveform analysis alone.

It must be stressed, however, that ODS can only be used to visualize how a machine vibrates; it cannot confirm the presence of a natural frequency, and it cannot be used to calculate damping, or to test structural modifications in order to estimate how the vibration (and natural frequencies may change). ODS simply involves the acquisition of amplitude and phase readings at the running speed of the machine (or optionally at all frequencies) and provides a visual representation of the relative movement of each point on the machine.

Quick Overview

There are a number of steps required to perform an ODS test. They are summarized below:

Collect magnitude and phase readings on a structure.
Visualize how the machine is moving at each point.
Look for movements of the machine as a whole.
Look for movements between machine components.
Software used to animate the structure

- Create the drawing.
- Define where measurements will be taken.
- Define relationship between data points and points without data.

Step One: Plan the job

We need magnitude and phase readings at each point on the structure – but which points should be used?

What fault types might explain the vibration patterns?
What parts of the machine or structure do not need to be measured?
What points on the machine and support structure must be measured to accurately display these faults?
What will the ODS structure look like when drawn in the ODS software?
How many points should be measured on a beam, base, foundation or floor?
What point should serve as the fixed reference

How many test points should there be?

You will need to decide how many points on the machine/structure need to be tested and animated in order to provide you with the results you desire. There are two issues:

You may like the animation to look realistic; both to help with your visualization of the machine or structure, and to aid in a presentation to management (if you are a consultant, or you need to convince others about your recommendation for structural modifications). It generally helps if the animation looks as realistic (and cool) as possible.

In order to best understand the motion of the machine, it is necessary to test a large number of points. Assumptions about how a structure is vibrating can cost you dearly. For example, if you were testing a structure with a long horizontal beam, you might assume that the vibration is related to its first bending mode; so you might test at each end and in the middle. However it was vibrating in its second bending mode, the vibration in the middle would be minimal (at the natural frequency).

If you are attempting to visualize vibration/motion related to unbalance, misalignment, structural looseness or some other non-resonant condition, you could potentially set up a job with a very small number of test points (an X and Y at each bearing). As we will soon see, it is possible with most software programs to nominate that certain points on the machine vibrate the same way as selected tested points – therefore the model can look very good with just a small number of test locations.

On the other hand, if you are dealing with a large structure, and you are not exactly sure what you will find, then the project could easily involve between 150 and 500 test locations (testing in the X, Y, and Z axis at each location). Once you become efficient with your testing, you may be able to test 3 points per minute.

Make a plan

It is very important to create a sketch of the structure. Map out the scale of the structure, and mark each of the test locations.

You should also define your coordinate system: X, Y, and Z. In this example Z is positive upwards, X is positive to the left of the reference wall on the extreme right, and Y is positive away from us. You must always keep this coordinate system, and the positive directions in mind when you are performing your tests. Phase readings may need to be compensated for direction depending upon the orientation of the accelerometer. It is extremely important to keep the coordinate system, and positive directions, clear in your mind (and in your test date).

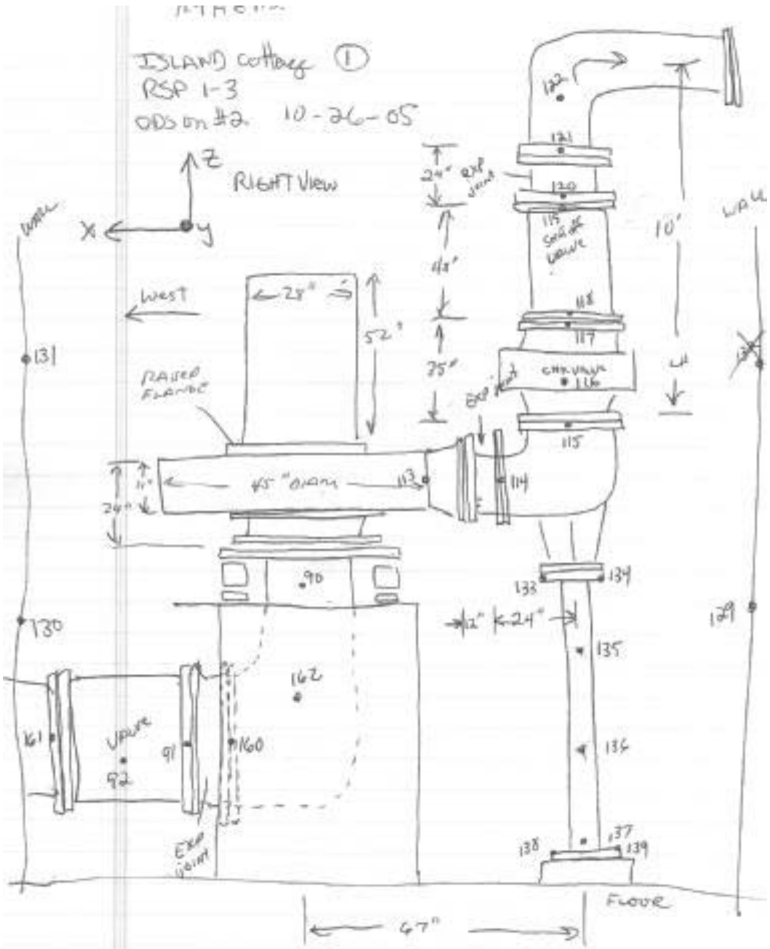


Figure 7-1

The sketch does not need to be perfect; it does not need to be perfectly to scale. But a little effort at this stage will help you to perform a more successful, less confusing ODS test.

It is important to clearly indicate the location of each of the test points. It is also worth adding details that you will add to the final model in the ODS software.

You may also like to create a sketch facing north (for example) and another facing west. Again, this will help you to stay organized, and to define the structure in the ODS software.

If you are "artistically challenged", you could instead take a digital photograph, and either print it and use a pen to add the dimensions, test locations, etc., or to use a graphics program – however that will be more difficult when you are in the field.

Phase readings

Vibration amplitude and phase readings are required to display the amount of movement and the relative direction of the movement. That means that phase readings are required. There are basically two ways to acquire this data when performing ODS testing: use a single-channel system and use a once-per-rev tachometer signal as the reference (and gather amplitude and phase readings at running speed only), or use a two-channel analyzer and gather amplitude and phase readings at all frequencies of interest.

Phase with a single-channel analyzer

Single-channel phase readings require a reference. This is typically a once-per-rev laser or optical tachometer signal. That signal is fed into the "trigger" or "external tach" channel of your analyzer (refer to your manual). An accelerometer is placed at the first test location in the X direction and the vibration reading is acquired. The analyzer will do one of three operations:

It will acquire the amplitude and phase at the running speed and display the readings on the screen. You will record those readings on a table that you print, in a spreadsheet, or simply in a notebook.

It may be able to record the amplitude and phase reading in the memory of the analyzer for later download.

It may acquire the amplitude and phase spectrum at all frequencies, however if an optical or laser tach are used, only the phase readings at 1X, 2X and the other orders are valid, so you should only attempt to view animations at these orders. If you do wish to view animations of your machine at speeds other than running speed, it is recommended that you use a two-channel analyzer.

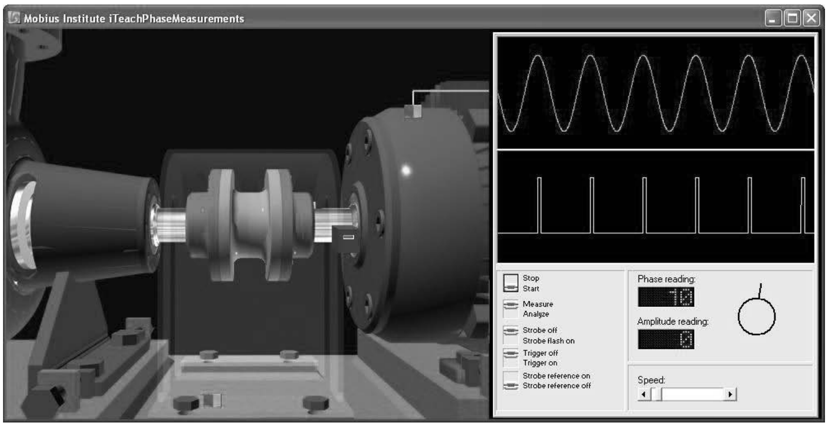


Figure 7-2

You should check with your data collector/analyzer products, and with the ODS software to learn what is possible.

The value of magnitude and phase data

It is important to understand what it is that you are hoping to achieve. Many people assume that it is only the running speed vibration that excites resonance. This is not true. Any strong source of vibration at any frequency will excite a resonance, therefore it is worth, if possible, to use two-channel test methods, described next, to acquire amplitude and phase data at all frequencies.

For example, the following spectra appear to indicate that a resonance may exist in the frequency range 30 Hz – 60 Hz. The running speed of this machine is closer to 20 Hz. Therefore, in this case, we would like to see how it vibrates at approximately 50 Hz (based on the information provided).

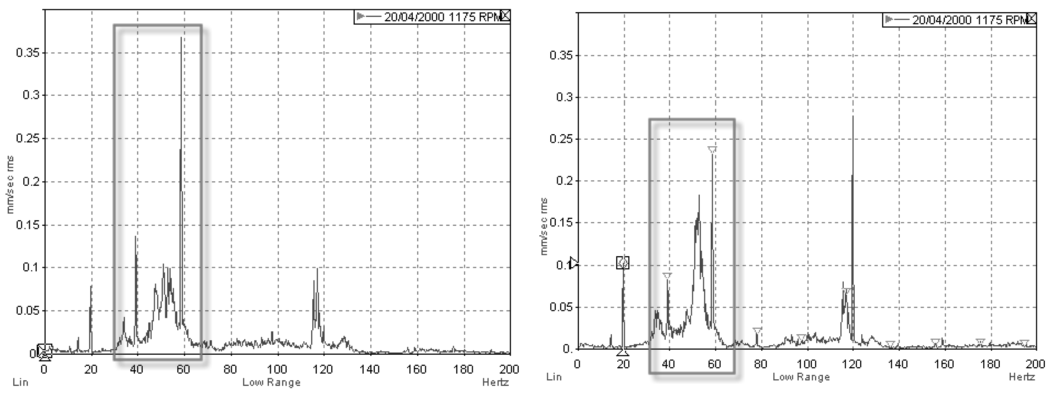


Figure 7-3

Document the readings

As mentioned earlier, the readings can be recorded in the data collector/analyzer for later extraction, or you could draw a table listing the point number and axis with a column for the magnitude (amplitude) and phase readings.

You can even compensate the phase readings at this time. If the accelerometer is pointing in the negative direction, then you would add or subtract 180 degrees.

Phase with a two-channel analyzer

Two-channel testing, also called cross-channel testing, provides amplitude and phase values at every frequency in the spectrum frequency range. Two accelerometers are required. One accelerometer remains at a reference location. A second accelerometer is moved from point to point on the structure/machine, in the X, Y, and Z axis. The analyzer will record the cross-channel spectrum, which contains amplitude (e.g. mm/s or ips) and phase at every frequency, for later download.

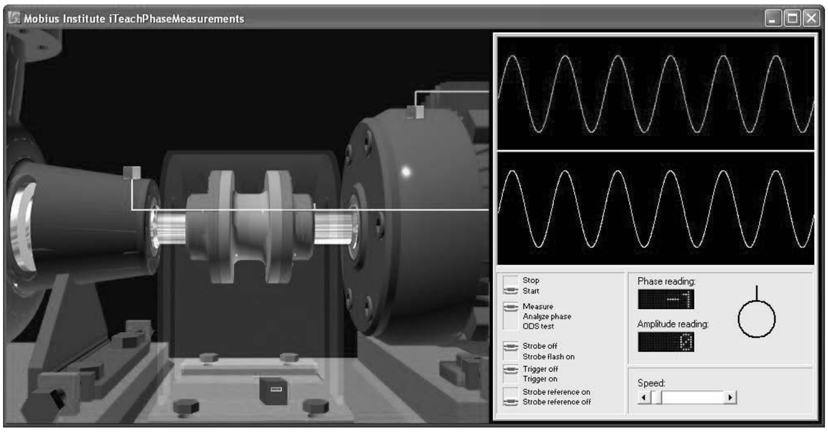


Figure 7-4

You have two choices for how you manage the data:

You can collect the data, display it on the screen, and extract the magnitude and phase readings at the frequencies of interest. You would only do that if option 2 was not available.
The data would be saved in the collector/analyzer and the vendor software would extract the data and transfer it to the ODS software. This provides the greatest automation, and the complete range of analysis options.

Most analyzers will allow you to view the spectra that were acquired at each test location (and axis), along with the reference spectrum collected (i.e. you will be able to see the channel 1 and 2 data). This graph will include the amplitude and phase data. However you should also be able to display a graph that depicts how channel 1 compares to channel 2, as a ratio. Your analyzer may provide the data in one of two formats:

Traditionally the Frequency Response Function was a dB (logarithmic) display. When we discuss modal analysis, we would typically expect to view this data in dB format so that we can easily see the peaks of the modes and the depths of the anti-nodes.

Certain analyzers display this data as the "transmissibility". It is still a ratio, but the display is linear instead. A value of '1' indicates that the vibration amplitude measured on channel 1 and 2 were equal. The linear ratio can be easier to understand than dB values, however the display may be a little confusing (in the author's opinion).

You should check the manual of your analyzer to learn about its display options, and learn how your ODS software will acquire and display the readings.

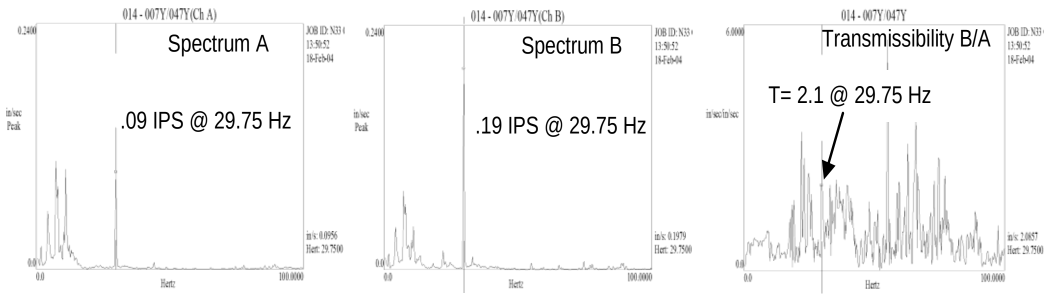


Figure 7-5

Important note:

When you perform an ODS test the analyzer will record the vibration from the accelerometers and the data will be saved. Unlike modal analysis, discussed separately, it is assumed that the machine/structure are vibrating in a consistent manner (speed and amplitude) during the entire test. This may not be the case. If the speed changes, or the amplitude changes (because of load changes, for example), the final animation you see will not provide a real indication of how the machine vibrates. Speed changes are very difficult to manage. Load changes that affect the amplitude will cause the animated motion to be distorted.

One solution, that is generally impractical, is to use a multi-channel data acquisition system to acquire all the data simultaneously.

Otherwise you should simply strive to take all readings under the same operating conditions.

Visualizing the motion

Once the model has been defined and the amplitude and phase data has been acquired, it is time to analyze the data so that we can understand how the machine is vibrating.

We can try to keep it simple by creating a basic diagram of the machine and plotting the amplitude and phase over the diagram. In the following example of a simple horizontal beam, the amplitude is represented as the height of the dot above or below the horizontal line, and the phase angle is represented by its position relative to the line: above or below. We are *typically* looking for 180° phase changes. So you can set 0° to be above the line, and 180° to be below the line. If the mode shape is more complicated, or your structure is more complicated, then you need software designed for the task.

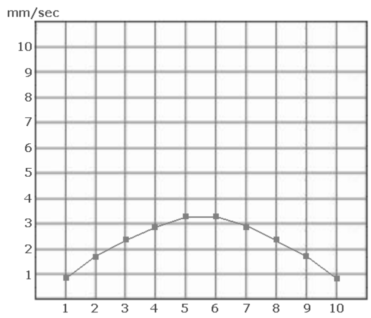


Figure 7-6

The ME'scope software is a popular application that can be used to graph the vibration data, build the model of the machine or structure, and animate the machine or structure at a selected frequency.

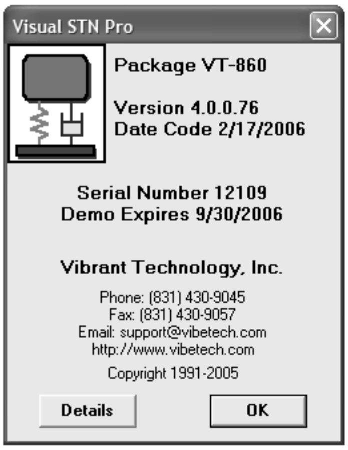


Figure 7-7

The software is developed by Vibrant Technology. (Mobius Institute has no financial relationship with Vibrant Technology, Inc.)

Creating the ODS job

The steps required to set up the project in the ODS software are as follows:

Define the structure in 3D

> You must take your sketch and all of the coordinate information and define the structure inside the ODS software. You can define components, for example motors, pumps, walls,

etc., and re-use them within the same project or future project. You may like to create your own library.

Define the points where data will be assigned

You will define each of the points where data has been collected. Each point will have an identifying number, and can accept data for three axes.

Define interpolated points

You may also define points where data will not be collected. Instead you may define that these points either move the same as another point (where data may exist), and you can define interpolated points, where a point will act according to the data collected at two points – the software will compute new values according to the relative distance between the points.

Define graphical elements

You can also define colors, color fills, and lines that are drawn between points. These elements help to make the structure more realistic, and more appealing to managers who may see the final report.

Import/enter the data

The data must be imported into the software (or manually entered). You will need to look at the supported file formats and to see if the software can read the data directly from your vibration monitoring system.

Animate the machine/structure

And finally, you can animate the structure. If you used a single-channel analyzer, the structure can only be animated at a single frequency – typically the running speed of the machine.

If you used a two-channel analyzer and acquired cross channel data you will be able to select the frequency of animation. While the structure is animating you can move a cursors through the test data in order to see how the structure moves at different frequencies.

The software will look at the relative amplitude levels to assess the amount of movement, and the relative phase angles to determine the motion of the structure. The software will them cycle through 0° to 360° allowing the structure to animate. You will be able to change your perspective and you can zoom in to get a closer look.

Here is an example:

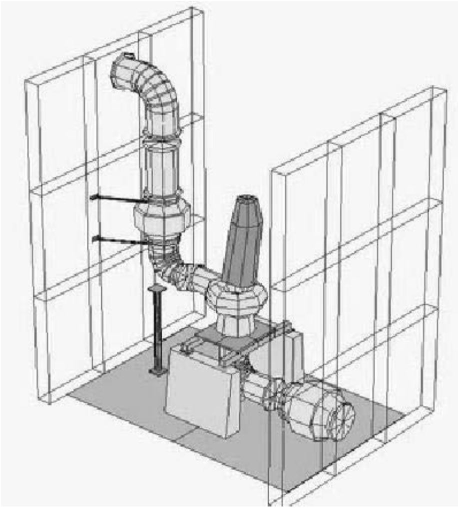


Figure 7-8

Interpreting the animation

There are generally three features you are looking for:

Look for motions of the whole machine:
- Rocking, bouncing, twisting
- May *indicate* resonance
- May indicate structural looseness

Look for movement on or between the components:
- May indicate unbalance or misalignment
- May indicate cocked bearing, bent shaft, etc.

Look for movement between machine components and base/foundation:
- Looseness, broken bolts, broken welds, grout problems.
- Soft foot.
- Diagnosing common machine fault conditions

When you view the animation at the running speed of the machine you can view the motion of the bearings, potentially in three dimensions (i.e. vertical, horizontal and axial). If a machine was out of balance, you may see a circular motion in the radial direction and minimal motion in the axial direction – unless it is an overhung machine. By looking at the axial movement you may detect misalignment, cocked bearing, or bent shaft.

You can also see how the components move in relation to each other. Looking at the relative radial and axial movement, you may detect misalignment.

And if you study how machine components move relative to the foundations, you may detect soft foot, looseness, broken bolts, cracked welds, cracked grout, and more.

If you study the Phase Analysis chapter carefully, you will be able to see how different fault conditions will cause the machine to move in different ways. ODS testing enables you to create a visual animation of that movement.

Checking for resonances

By looking at the motion of the structure at the running speed, and at other frequencies where the machine generates vibration, you may see the structure swaying, bouncing, twisting, or rocking. These motions would indicate that the machine may be resonating at those frequencies – but it is not proof of the resonance. These movements can be confused with looseness/weakness.

It is not possible to prove that resonance exists with ODS testing alone. Even if resonance is suspected, it is not possible to identify the exact natural frequency as you can also study vibration amplitudes at frequencies where the machine itself generates vibration.

For example, the natural frequency may be 168 Hz (10,080 CPM), however the pump vane rate may be 175 Hz (10,500 CPM). The pump vane rate vibration will excite the natural frequency, so the amplitude at 175 Hz (10,500 CPM) will be much higher than normal, but you cannot tell if the natural frequency is actually higher or lower than the pump vane rate.

The only way to prove the existence of a resonance is to perform a modal test or a run-up/run-down test and observe a 180° phase change.

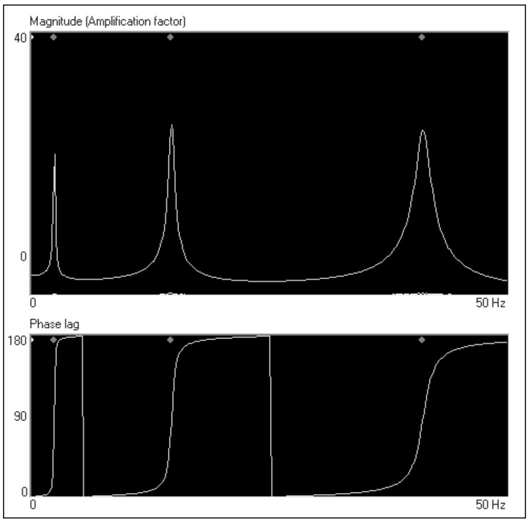


Figure 7-9

Time-based ODS

There is a completely different way to visualize the operating motion of the machine or structure using time waveform data. The time waveform represents the exact motion of the machine; as the waveform rises and falls we know that the structure rose and fell (if the waveform is in units of displacement).

If we collect multiple time waveforms simultaneously from points on a structure, we have a record of exactly how the structure moved from one moment to the next. If we collected waveforms from either end of the center-hung fan and saw that one waveform was sinusoidal that rose at the beginning of the record, and the other waveform was also sinusoidal but it fell at the beginning of the cycle, then we could visualize a rocking motion – couple unbalance.

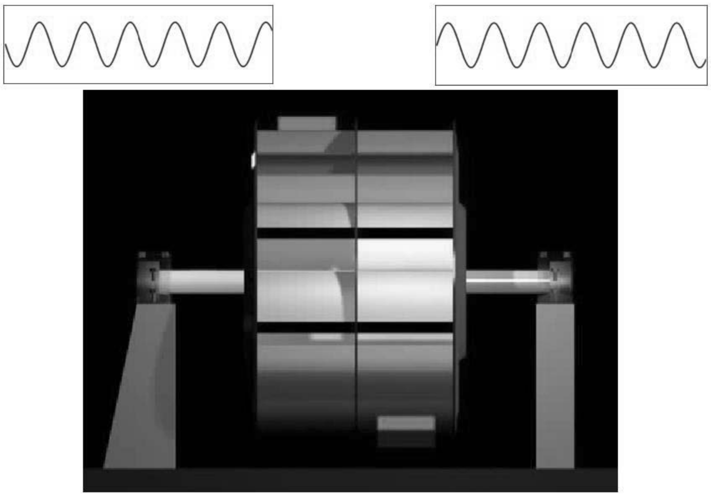


Figure 7-10

We could perform this function at multiple points and then use software to animate the machine. We must have multiple points tested *simultaneous* – and that is where the challenge lies.

This application can be used to understand the motion of a machine or structure, however it is probably best applied to situations where you are trying to analyze the exact motion of a process such as a hole punch, drill press, or some other mechanical system where you are looking for motion that is out of the ordinary: bounces, wobbles, scrapes, etc.

The following illustration was taken from the ME'Scope software program. The four waveforms were collected from four points on the structure. The other points are instructed to vibration in-phase with the tested points. The software is able to move a cursor through the data, automatically updating the model of the satellite dish.

What you see in the movement of the dish is exactly what happened during the test. Rather than animating the machine at a specific frequency (for example, the running speed of a machine, or a frequency that was suspected to be a natural frequency), it shows the full motion of the structure.

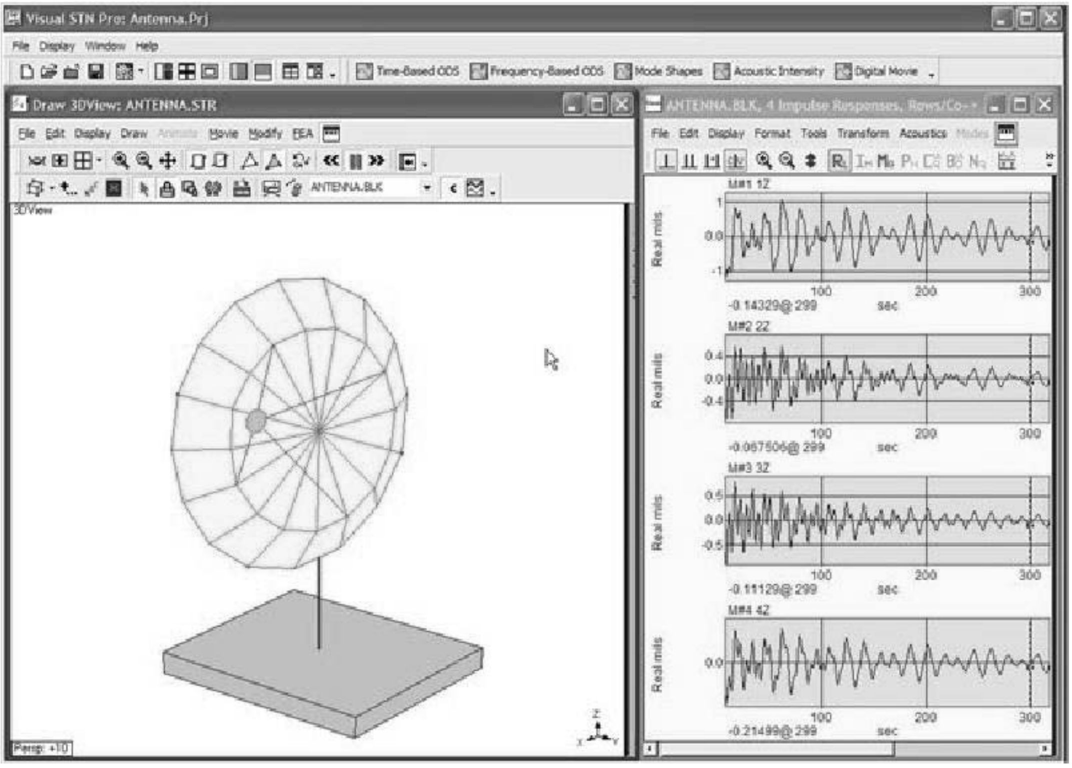


Figure 7-11

Example 1 ODS – Pipe Hanger Looseness

Background: The #7 waste pump is a three vane vertical pump direct coupled to a 1500 H.P. motor. A 10' drive shaft connects the motor to the pump. The unit runs at a constant 592 rpm (9.86 Hertz). The pump is located in the basement. The 20" discharge pipe is supported from the bottom and bottom and the runs vertically through the building. Five floors above the pump is a concrete and steel structural beam. A welded steel pipe hanger is bolted to the bottom of the beam and steadies the pipe at the top. The pump has been used infrequently since installation, several decades ago, because it shakes the whole building. Similar pumps in the building do not shake the floor.

Figure 7-12 - #7 Waste Pump

Investigation: Vibration levels on the motor and pump were smooth. No problems were found on either component. The discharge pipe was vibrating 3 mils peak-peak at the third harmonic of rotational speed (29.71 Hertz). This vibration could be felt on all decks but it was highest on the 5th floor. The source of the vibration is pump vane pass frequency *(Vane Pass frequency = rotational speed x 3 vanes on the pump impeller)*.

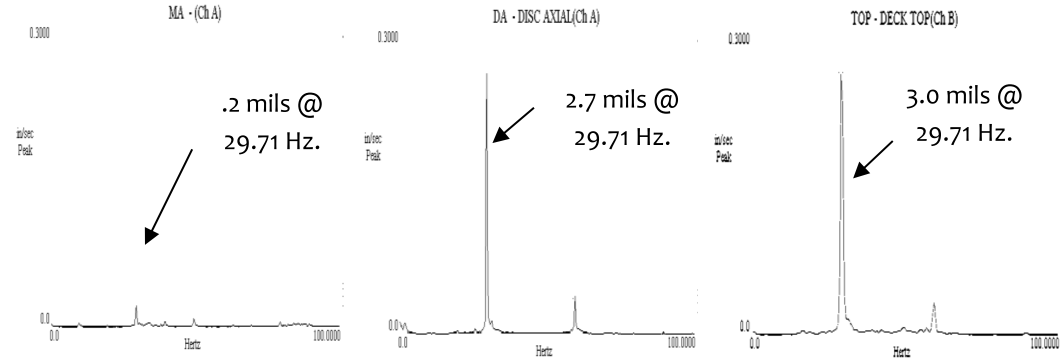


Figure 7-13 - Vertical Vibration at pump vane pass frequency (29.71 Hertz) Motor (left), Discharge pipe near the motor (center) and 5th Floor deck (right)

Impact testing on the 5th floor beam identified a natural frequency at 29 Hertz – very close to vane pass frequency.

An ODS was made of the entire pipe structure. Scaffolding costing $10,000 was required to measure along the length of the pipe. The ODS structure drawing is shown in Figure 1.3. Measurements were made along the length of pipe sections and at each flange, elbow, valve and pipe supports.

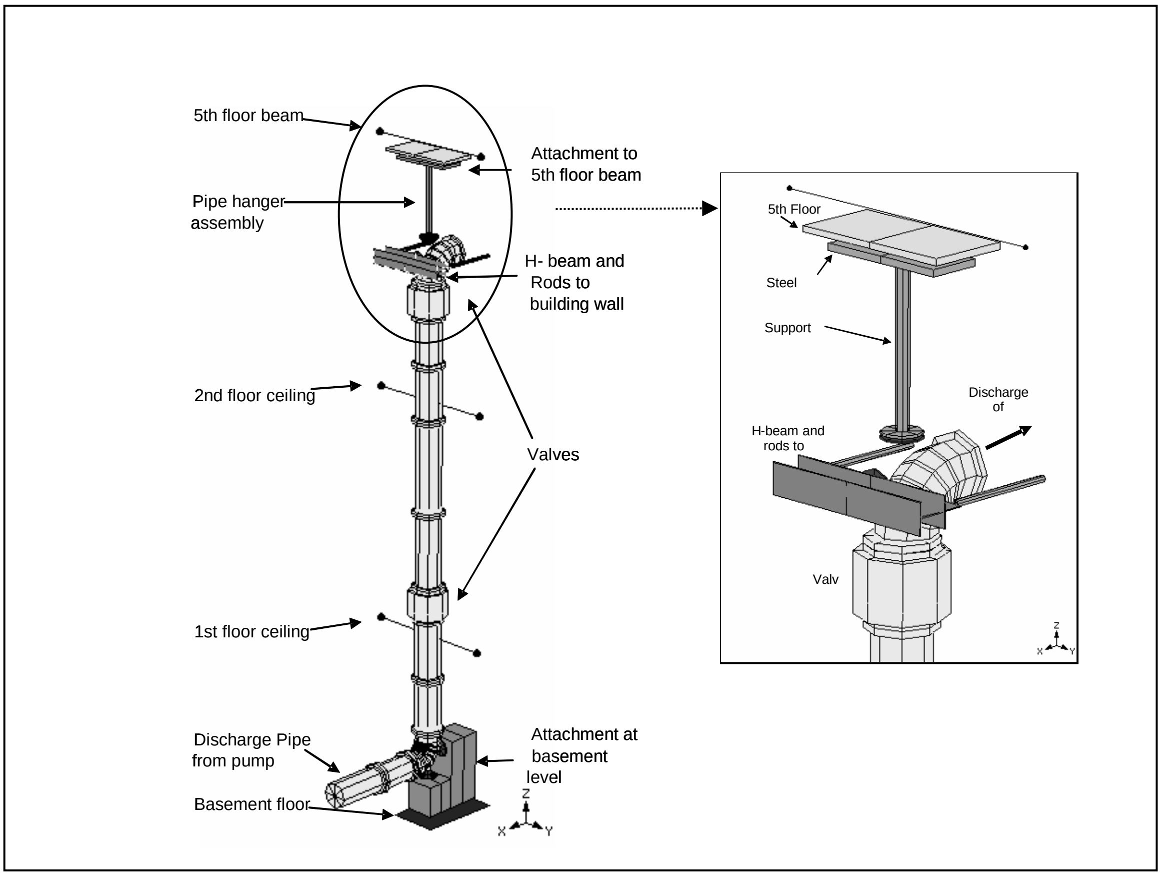


Figure 7-14 - ODS Structure Drawing

The ODS animation at 29.71 Hertz identified looseness where the pipe hanger bolts to the underside of the 5thth floor beam. Mechanics found that the hex nuts that attach the hanger assembly to the beam were loose. After the bolts were tightened, the pump was operated with very little vibration.

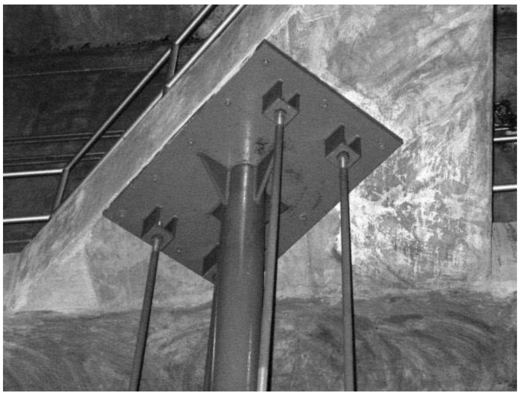


Figure 7-15 - Pipe Hanger Assembly Attached to 4th floor Beam

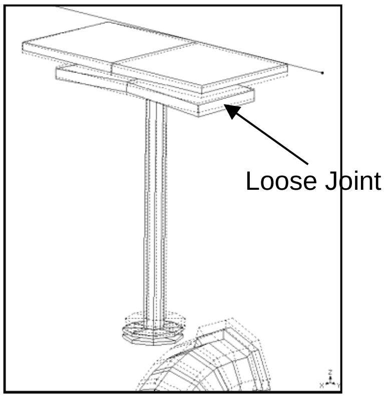


Figure 7-16 - Loose Joint in ODS Animation

Example 2 – ODS of a Papermill Press Section

Background: The wet press section of a paper machine was experiencing a high failure rate on felts. The felt became barred after a short amount of run time. Felt barring is the term used to describe the condition where the felt thickness is no longer uniform. Felt barring is caused by periodic nip forces such as roll eccentricity, roll unbalance, or relative motion between rolls. A felt can also become barred if a barred roll transfers its pattern to the felt.

There was no shortage of theories at the plant about source of the felt barring. Some thought that the wet press machine frame might be weak or resonating while others believed that the problem was related to the bellows (air bags) which control the nip pressure. The press section components are identified in Figure 7-17.

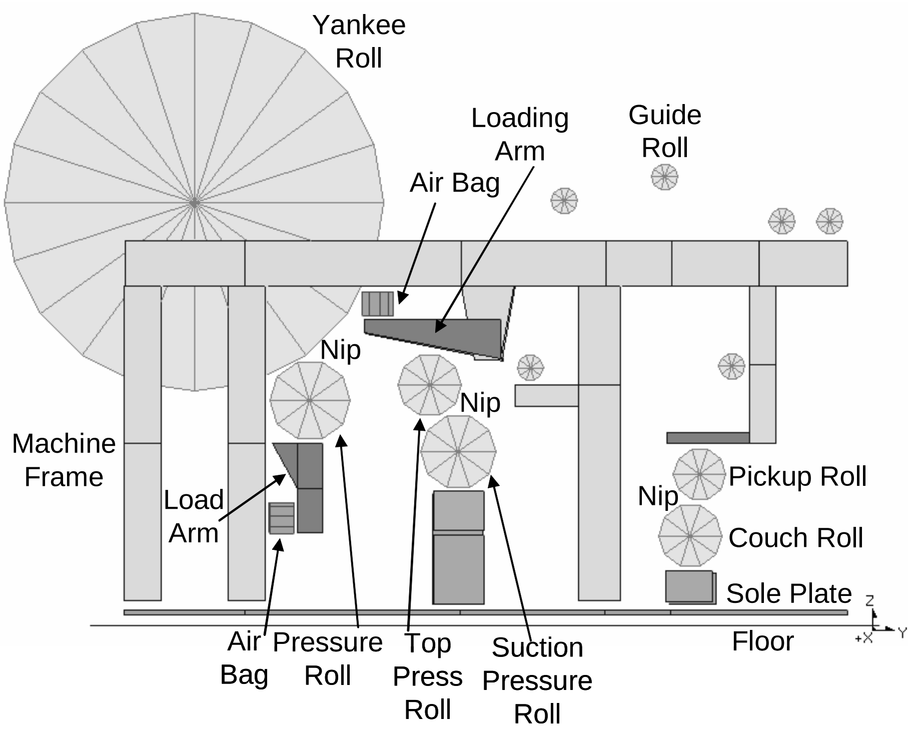


Figure 7-17 - Wet Press Components

Much effort was put into elimination of the felt barring but the problem persisted. It was decided that an Operational Deflection Shape (ODS) was needed of the entire wet press section.

The ODS consisted of 253 measurements on the floor, sole plates, machine frame, roll bearings and loading arms. The reference accelerometer was placed vertically on the suction roll bearing. A second accelerometer was moved to each degree of freedom where the transmissibility was measured. Transmissibility is the ratio of the two signals at each line of resolution. The transmissibility data set was used to animate the wet press ODS structure drawing shown in Figure 7-18.

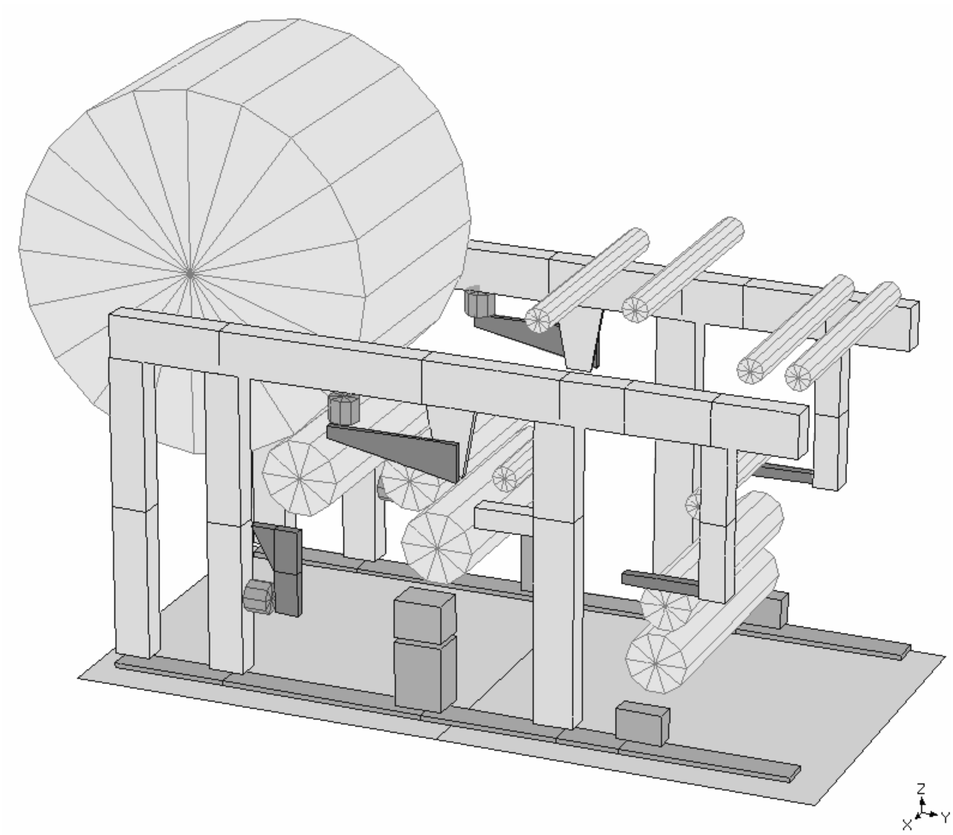


Figure 7-18 - 3D Structure Drawing of the Wet Press

The largest vibration found on the machine was at 19.2 Hertz. From the analysis of the ODS at this frequency, the following observations were made:

The loading arms on each side of the top press roll were out of phase with each other. The vibration was 1.4 inches per second (IPS) vertically on the loading arms. As a result, the motion of the top press roll was out of phase from end to end. The roll was not maintaining contact with the suction pressure roll. The top press roll speed was 6.72 Hertz. The suction pressure roll speed was 5.52 Hertz.
The pressure roll was rocking out of phase from end to end and not maintaining contact with the Yankee roll. The vibration on the pressure roll bearings was .25 IPS. The pressure roll speed was 5.2 Hertz
The ODS animation showed the machine frame, sole plates and concrete floor were not deforming or vibrating.

Figure 7-19 is a picture of the wet press ODS animation at 19.2 Hertz. It shows the motion of the top press roll and loading arms. Slight relative motion between the pressure roll and Yankee roll is also visible. The machine frame is not shown in the picture in Figure 7-19 because it was not moving.

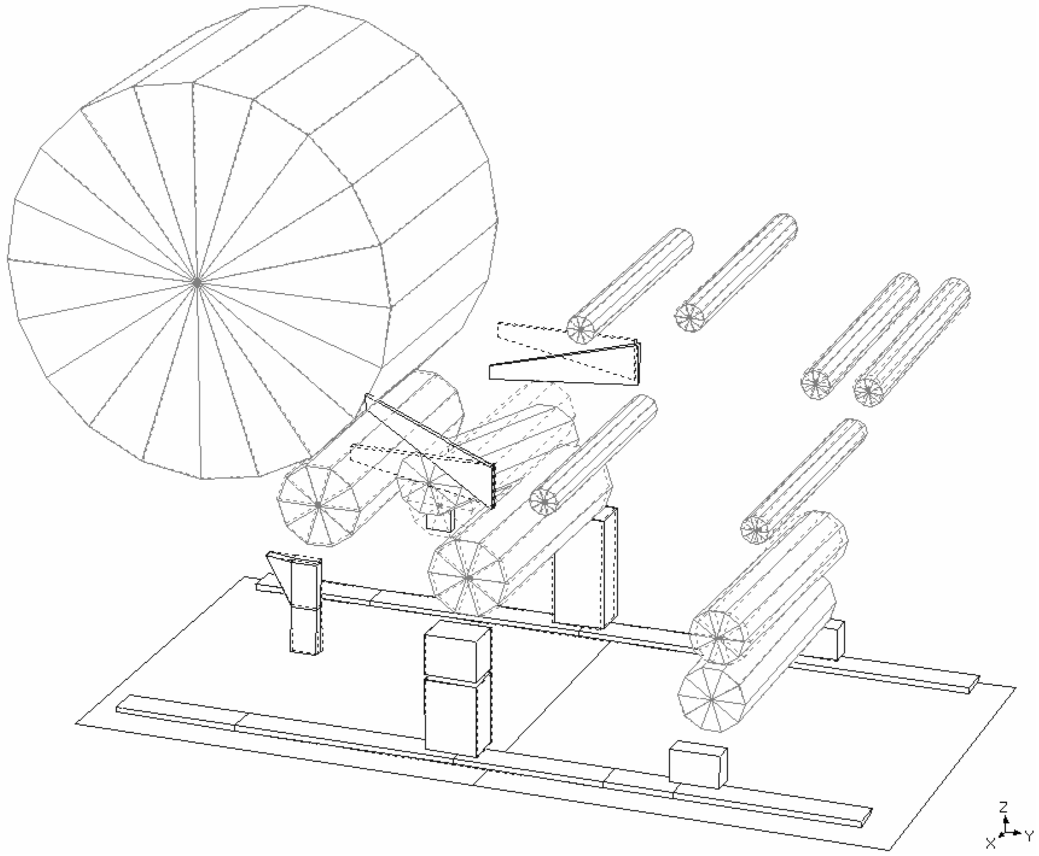
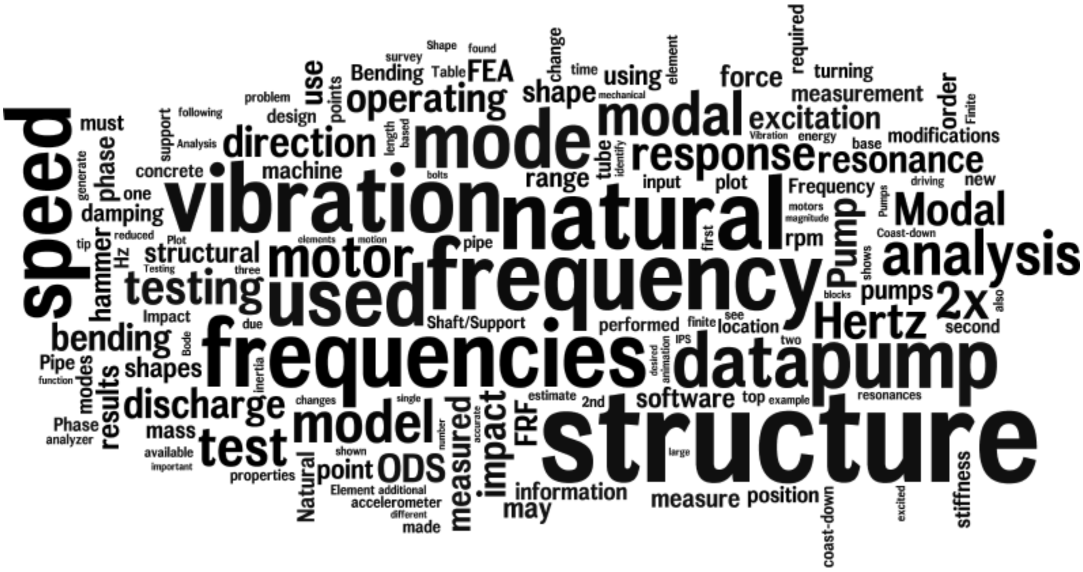


Figure 7-19 - Wet Press Shape at 19.2 Hertz

The motions of the pressure roll and top press roll caused the felt barring. The source of the 19.2 Hertz vibration was thought to be resonance of the air bags excited by a harmonic of the top press or suction rolls.

To correct the problem, the plant engineers decided to fill the air bags with half water and half air. The modification was successful in eliminating the felt barring problem. The air over water modification increased the bag's stiffness and moved its natural frequency away from forced vibrations.



Chapter 8

Modal Analysis

Objectives:

- What is modal analysis?
- Measurement collection
- Creating mode shape animations
- Modal analysis and finite element analysis

An introduction to modal analysis

Modal analysis is a testing method that provides data that provides proof of the existence and frequency of resonances, and an estimation of damping and mode shapes. We can use this data to understand how a structure vibrates we can use this information to design the changes required to minimize the vibration under normal operating conditions.

Modal analysis is used in a wide range of industries.

It is used to test hard drives to ensure that the head of the drive can be quickly moved to the required location of the disk without vibration that would generate data read/write errors.
Modal analysis is used to test automobiles to minimize noise as the driver operates the car through a range of speeds (you may remember that cars used to vibrate far more than they do today).
Modal analysis is used on precision objects, such as the Hubble space telescope, to ensure there is minimal vibration; thus delivering the clearest images.

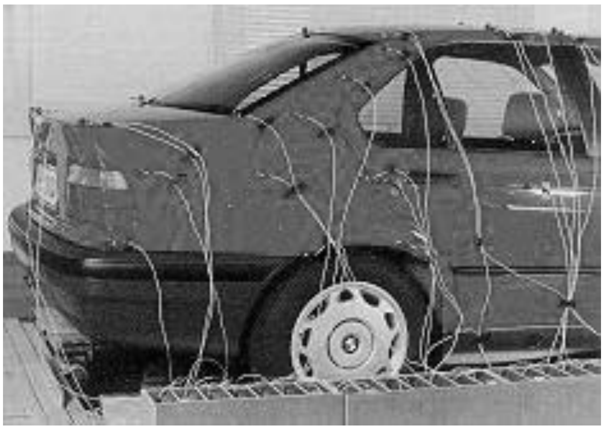


Figure 8-1

In industrial applications modal analysis is typically used to understand the mode shapes of a structure so that vibration levels are minimized across the operating speeds of mounted machinery; which in turn delivers great reliability (reduced structural failure or increased bearing wear due to excessive vibration.)

Modal analysis is not the same as Operating Deflection Shape (ODS) analysis. Whereas ODS provides an indication of the existence of resonances and the mode shapes, Modal analysis provides proof of the existence of resonances, and an accurate estimation of the mode shapes. If performed correctly, modal analysis can provide the data necessary to enable design changes to be calculated that will change the natural frequencies so that they do not coincide with machine operating speeds (or sources of forced vibration).

These results were generated from modal test data. Using sophisticated software, test data collected from numerous points can be used to show the animated mode shapes at each of the natural frequencies.

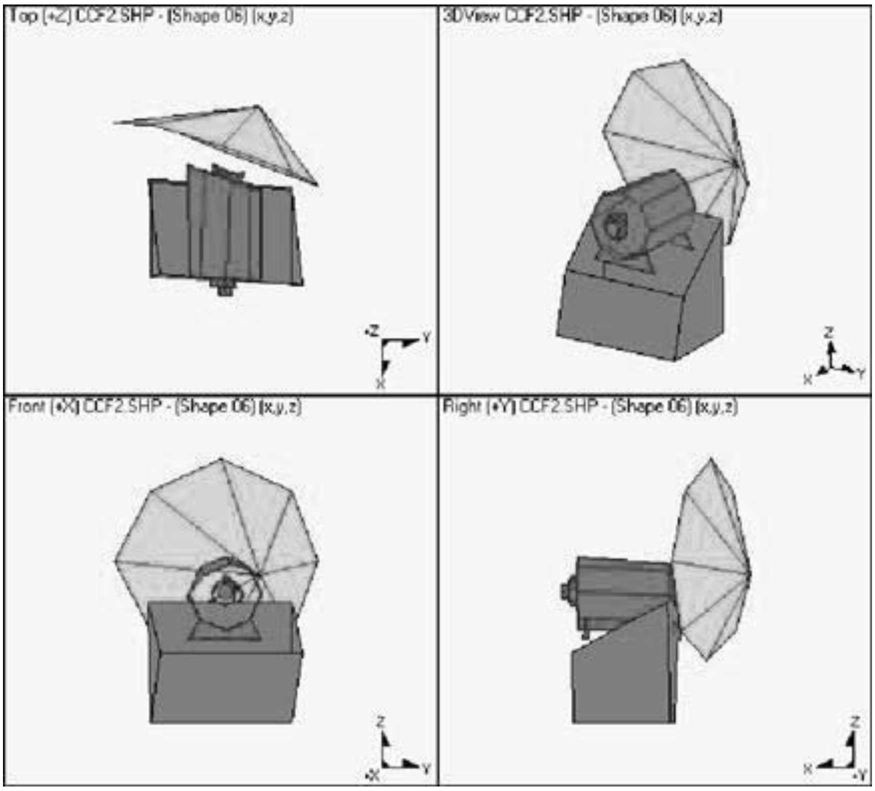


Figure 8-2

Note that ODS may be used to generate an animation of the vibration of the machine based on vibration amplitude and phase information. The animations generated from ODS data do not prove the existence of resonances.

Here are a few key facts about experimental modal analysis:

Modal analysis is performed on a structure when it is not operating.
Modal analysis reveals the information that proves where the resonant frequencies will be located, and how the structure will move/deform when it is excited.
Modal analysis can be used as the basis to compute the required structural modifications required to alter the structure

- Change the mass, stiffness or damping to move the resonances to a frequency that will not be excited by our machine's forcing frequencies.

There are a number of steps required to perform the modal analysis test:

1. Determine the geometry/test points
    a. Enter the plan into the modal software

2. Take data at each test location
    a. Feed it to the analysis software
3. Analyze the results
    a. Find the resonant frequencies
    b. Curve fit the frequency response function to extract damping
    c. Animate the structure
4. Make the required modifications
    a. Modify the structure so resonances won't be excited

Defining the geometry of the structure under test

In order to organize the testing process, and to perform the animation, you must create a model of the structure being tested. There must, as a minimum, be a point on the model for every single point on the structure that you will test. In order to create a model that most closely resembles the structure, it is necessary to define a large number of points.

The good news is that it is not actually necessary to collect a measurement at each and every defined point. It is possible to define points that move the same as other points on the structure. This allows for a better looking model (and animation) without requiring as many test points.

It is important not to use too few test locations. You must be careful not to make assumptions about the location of nodes and the mode shapes that you may uncover.

For example, if you were to define a small number of points along a beam you would not be able to correctly identify the third (or higher) bending mode – you may be mislead into thinking that a first or second bending mode exists.

Most modal analysis programs make it relatively easy to define the structure, and provide predefined model components (either shapes like rectangles, cylinders, rods, etc or actual motors, pumps, etc.) that enable you to quickly assemble the models.

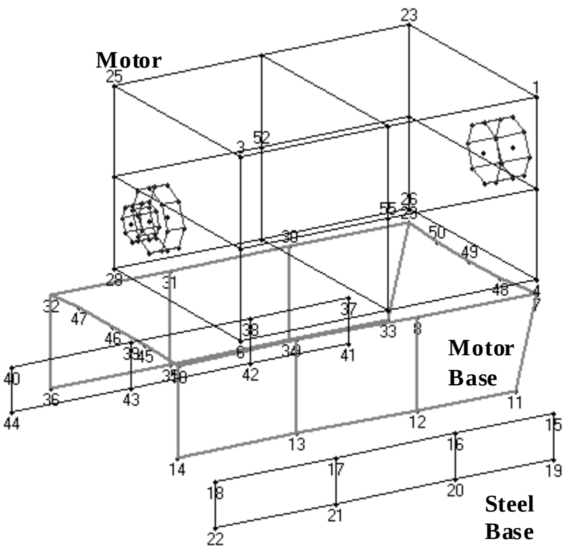


Figure 8-3

Modal testing

Modal testing involves the injection of a known force and measuring the response. A number of methods are used to excite the structure, however the machinery reliability field will almost always use "impact testing" – we use a "modally tuned" hammer (equipped with a load cell) to excite the structure. We measure the force used to strike the structure with the load cell, and we measure the response with an accelerometer. We would normally repeat these tests in all three axes so that we have a complete picture of the modes. We will discuss this procedure in greater detail shortly.

The machine must not be operating at the time of the test because we make the assumption that the measured vibration is generated as a result of the input force, or "excitation". Any external vibration or noise will compromise the test results. The coherence and cross-correlation tests are used to identify the presence of external noise/vibration (as well as identifying possible non-linearities in the structure, and the effectiveness of the excitation force.)

Frequency Response Function (FRF)

In order to learn about the resonant characteristics of the structure, we compute the Frequency Response Function (FRF), also known as the Transfer Function. The FRF is the ratio of the measured response to the input force.

If we were able to inject a force and measure a response that was exactly the same as the excitation force, then the FRF would be a flat line. However we know that at certain frequencies we will find that the vibration levels are amplified. We also know that the amplification will peak at certain frequencies (the natural frequency) and drop off at high and lower frequencies – the rate of reduction depends upon the damping and how closely spaced the modes are located. Remember, we are testing multi-degree of freedom (MDOF) structures so there will be more than one mode.

We will also observe that at other frequencies the output response will be very low in comparison to the input force. These frequencies represent the natural frequencies. For example, let's assume that we test a beam and find that the frequency of the second bending mode is 30 Hz. Now let's assume that we are measuring the response in the middle of the bar – right at the nodal point for the second bending mode. The value in the FRF at 30 Hz would be VERY low.

The following illustration shows three FRSs taken in the X, Y and Z direction, along with a simple structure being animated.

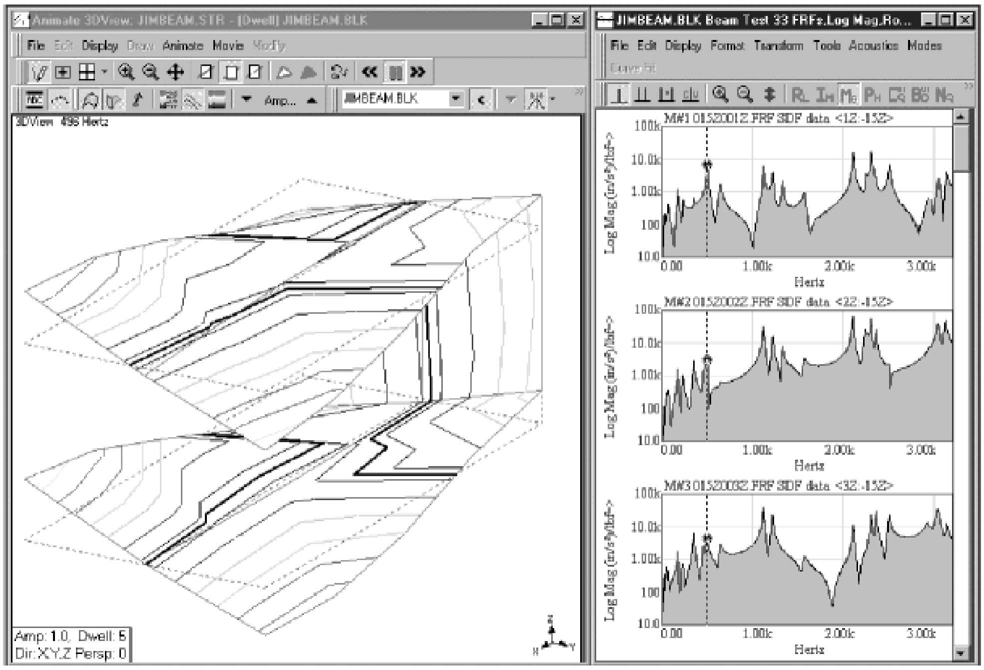


Figure 8-4

It should be noted that because we have a response at all frequencies, we can animate the structure at any frequency. We can certainly identify the natural frequencies and view each mode shape. Understanding the mode shape is the first step to understanding how the structure may be altered in order to move one (or more) offending modes away from the source of vibration that is being amplified (typically the running speed of the machine).

Sources of excitation in modal testing

There are commonly two methods used to provide the excitation: we can inject a "pulse" of energy (which includes energy at frequencies in our area of interest), or we can use a "shaker" to inject energy into the structure.

As mentioned previously, we typically use the first method via a modally tuned hammer. This is will be discussed in depth shortly.

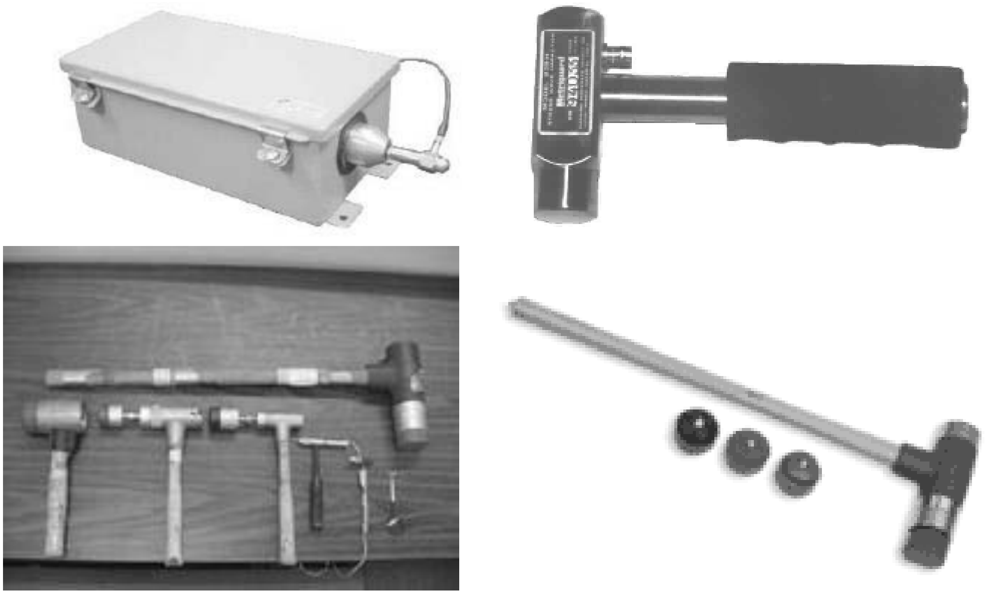


Figure 8-5

Using shakers in modal testing

In order to generate the FRF we need to compare the output response vibration to the input excitation vibration, at each frequency of interest. Shakers can be used to inject vibration energy into a structure. A shaker can be thought of as a large audio speaker with an attachment that enables the sound to be injected into the structure. The excitation force is generated by playing sounds into the shaker.

1. The sound may be random noise (or pink noise) – vibration at all frequencies (or in a specific range of frequencies)

2. Or vibration at a single frequency that is swept through the desired range of excitation frequencies

A great deal more could be said about shakers, however they are rarely used in the "machinery health" application, so they will not be discussed further.

Impact testing for modal testing

The most common form of excitation is the impact hammer. When a structure is struck with a solid object (such as a hammer or block of timber), an impulse of energy is injected into the structure. As we will learn, the energy covers a limited range of frequencies. It is our job to choose the most suitable hammer for our application.

Hammers are available that are designed specifically for this task. They are modally tuned to generate the desired input, and they are instrumented with a load cell so that the excitation force can be measured. The input force is measured in units of volts per kilogram force (volts/kgf) or volts per pound force (volts/lbf).

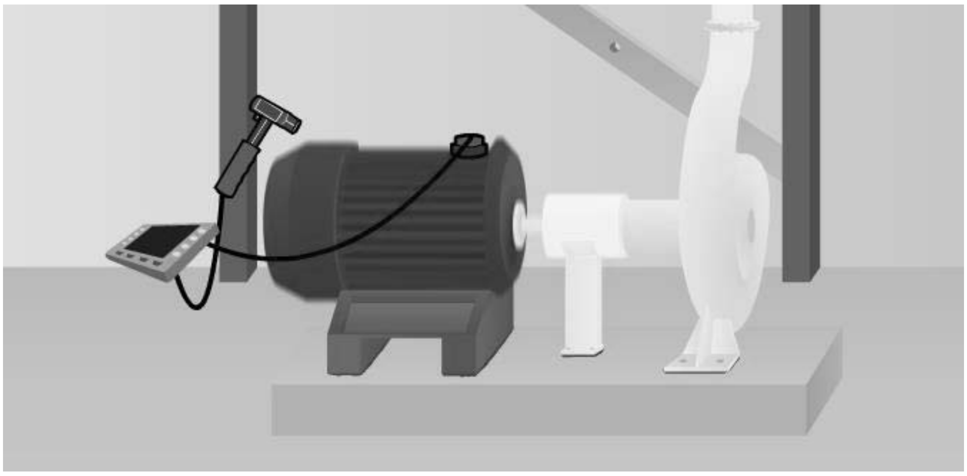


Figure 8-6

An accelerometer is used to record the response of the structure to the excitation force – in units of “g” (or mm/s^2).

The importance of selecting the correct hammer tip

The size and weight of the hammer, and the stiffness of the tip on the hammer, greatly affects the results obtained in the modal test.

We need to consider the size of the structure we are trying to excite. If it is a large, massive structure, such as the walls or floors of a building, then we must use a very heavy hammer. As shown in the following illustration, modally tuned hammers are available the size of a sledge hammer. A light hammer simply will not generate a sufficiently large force. A large structure is likely to have modes with a low frequency. We will use a soft tip on the hammer. The soft tip ensures that the input energy is concentrated in the lower frequency range.

If the structure is not as massive, such as the support structures for machines, then a lighter hammer may be used, and the tip can be stiffer. Again, it is easier to impart a sufficient force into the structure, and a slightly stiffer tip will ensure that higher frequencies are excited – the natural frequencies are likely to be higher in frequency.

And finally, if we were testing the rotor of a mid-size fan, or blades from a fan, then we would use a smaller hammer (otherwise we may damage the fan blade being tested), and we would use a stiff tip as the natural frequencies will be higher in frequency.

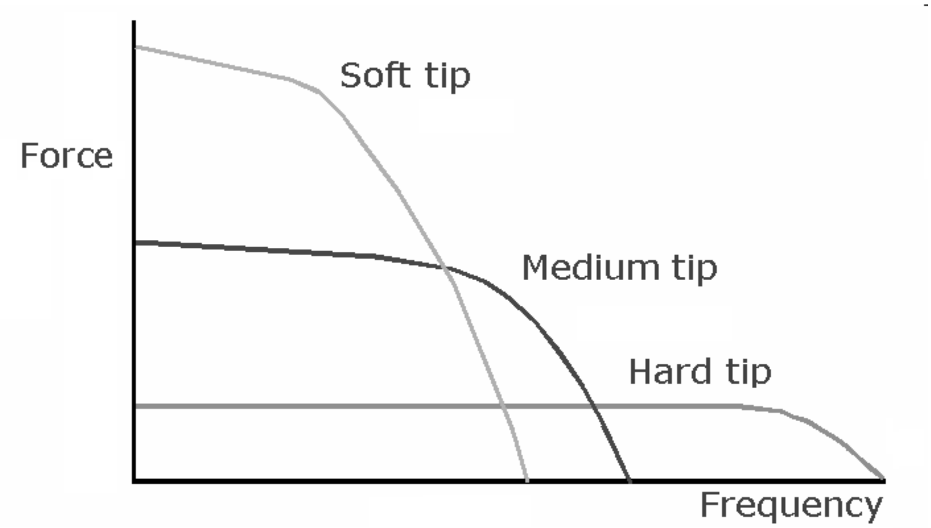


Figure 8-7

The following is an example of the input excitation and response. You can see the pre-trigger used on the input. You can also see how the structure has responded. As we will learn soon, the frequency and damping can be determined from frequency and rate of decay of the response.

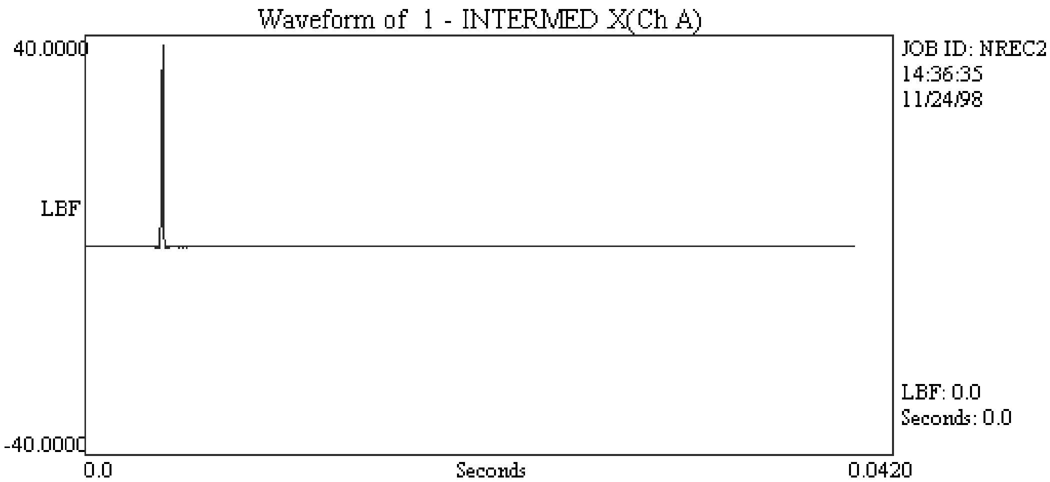


Figure 8-8

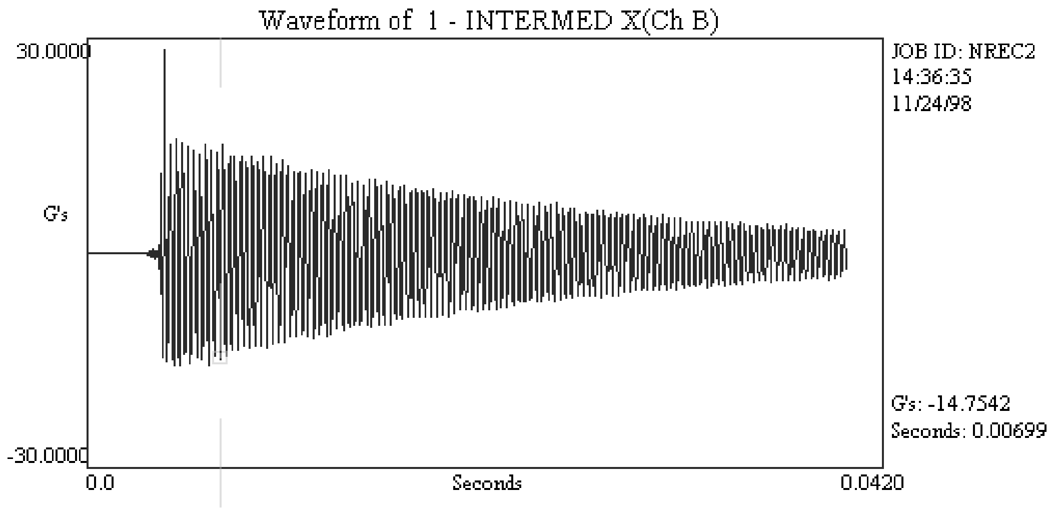


Figure 8-9

The modal testing sequence

We are attempting to determine how the entire structure behaves. We therefore need to compute the FRF between every point on the structure (in 3 axes) and a reference. This can typically be achieved in one of two ways:

1. We select a single location where we will strike the structure, and we move the accelerometer to every point on the structure. The data is captured simultaneously. For example, we will strike the structure at position "1" and read the response at position "1" in the X, Y, and Z axis (repeating the test for each axis). And then the accelerometer is moved to position 2 (X, then Y, then Z) while we always impact the structure at position "1".

2. Alternatively we can always measure the response in a single location while we strike the structure at each test location and axis. It can be difficult to use this method in industrial applications because it is difficult to impact the structure at each location and axis.

The collection of the data is critically important. The analyzer must record the excitation vibration and the response in its entirety. Therefore the impact must be wholly contained within the time record of one channel and the response must be wholly contained in the time record of the second channel.

Analyzer setup

The analyzer must be setup so as to capture the data as described above. Triggering is used so that the analyzer will wait until you strike the structure. We use pre-trigger so that the impact does not begin at the very beginning of the time record. We turn off the window function as there will be no leakage (assuming the impact is entirely captured within the time window.

The greater challenge lies with the response channel. Depending upon the damping, the structure may continue to vibrate (or "ring") after the analyzer completes the data acquisition. The result will be leakage and very poor results. We have two options available to us:

1. We can change our collection settings so that the time window is longer. We can to this by lowering the Fmax or increasing resolution.

2. We can use the "exponential" window. The exponential window is designed to effectively dampen the response signal – the aim is to add enough damping so that the signal decays to zero within the time record. The FRF calculation is then compensated for this additional source of damping.

As stated earlier, the FRF is derived by dividing the output response in volts/g (e.g. 100 mV/g) by the input excitation in volts/kgf or volts/lbf (or volts/N). Therefore the units of the FRF are g/kgf or g/lbf (or g/N).

The driving point measurement

In order to calibrate the system, and to enable us to compute how the structure may respond after structural changes are made (e.g. additional stiffness added between two points), we need a "driving point measurement".

The driving point measurement is the test where the response is measured at the same location as the input excitation. For example, if we always impact the structure at position "1", the driving point measurement is the single test performed when the response accelerometer is located at position "1".

Figure 8-10 shows a driving point measurement. This is a required measurement for modal analysis.

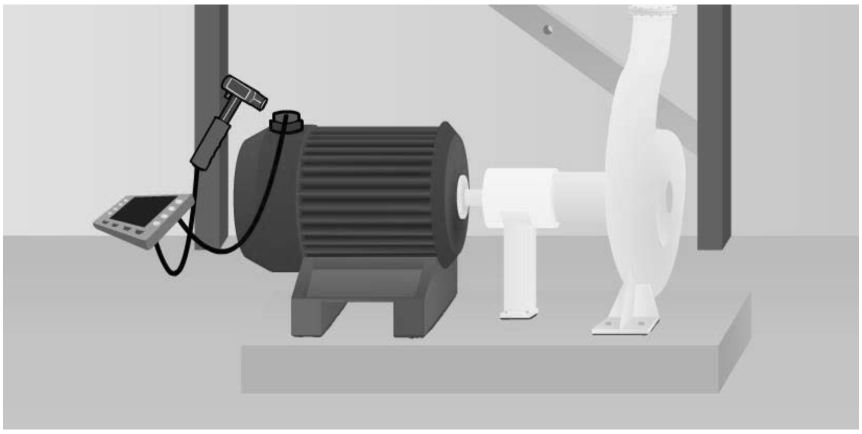


Figure 8-10

The rule of reciprocity

There is a rule that is assumed to be true when performing modal tests – the rule of "reciprocity". If we impact a structure at position "1" and measure the response at position "2", the FRF calculated will be the same is we were to impact the structure at position "2" and measure the response at position "1".

For this reason:

1. We can impact the structure at just one location and measure the results at each of the other locations and get the same results as if we impact the structure at each of the test locations and always measure the response in one location.

2. We do not need to perform the driving point measurement at every single test location.

Displaying modal test data

The FRF function is complex - it has both magnitude and phase. The FRF can be displayed in a variety of formats like MAGNITUDE, REAL and IMAGINARY. Three typical display types are:

- Bodé Plot
- Nyquist
- Co-quad

The Bodé Plot

The most common graph is the Bodé plot. Here we can see magnitude versus frequency and phase versus frequency. We look for 180 degree phase shifts in the phase plot, and corresponding peaks in the magnitude data. If modes are "closely coupled" i.e. their frequencies are closely spaced, it can be difficult to accurately estimate the frequency or the damping. Techniques are available to assist in this situation.

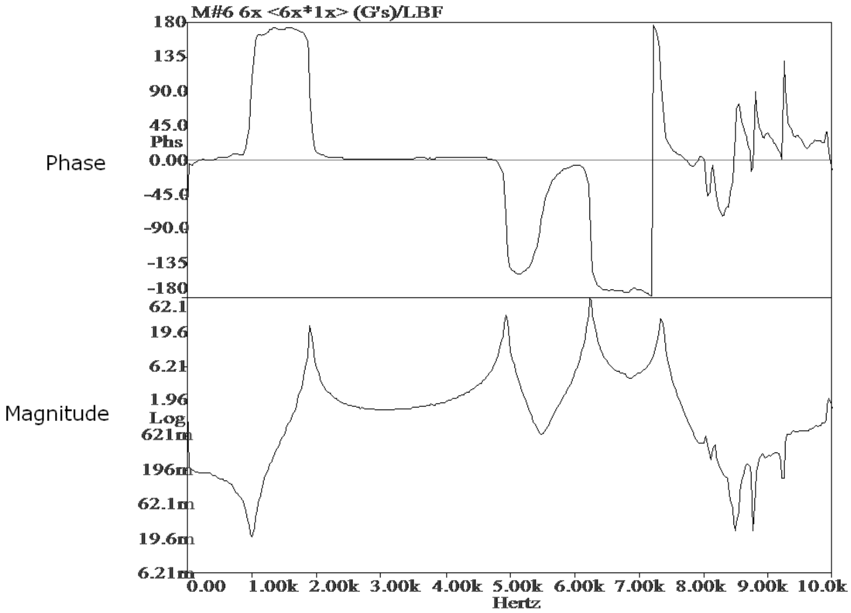


Figure 8-11

The Nyquist plot

The Nyquist plot graphs the real versus imaginary data with changing frequency. Because of the 180 degree phase change witnessed across a resonance, circles will plot out where natural frequencies exist.

In this example we did not have high resolution data so the circles are not very smooth, but with better data we will clearly see the circles (and as we will see, this shape can be used in a "circle-fit" curve fitting technique to estimate the damping).

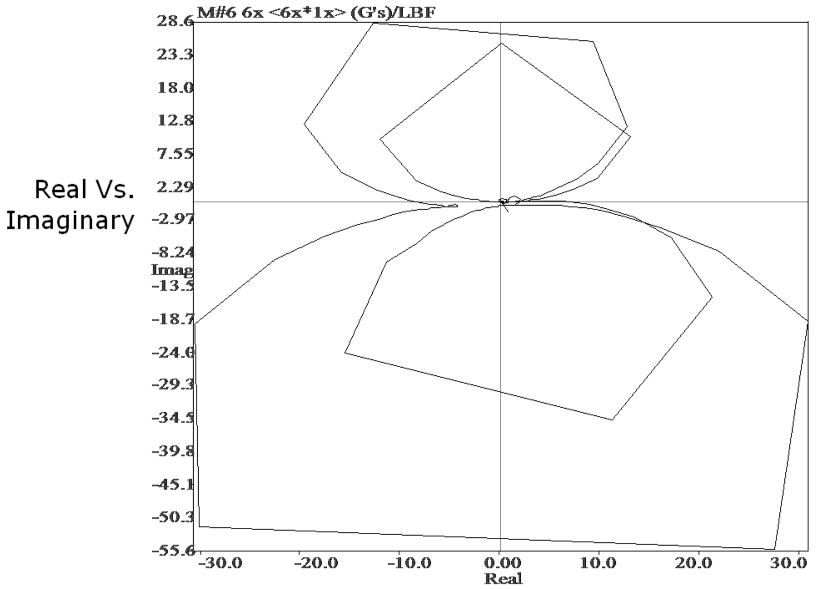


Figure 8-12

The co-quad plot

The co-quad plot displays the data in real and imaginary format. Among other things, we can get an accurate estimate of the natural frequency from these plots – the zero-crossing in the real data and the peaks in the imaginary data.

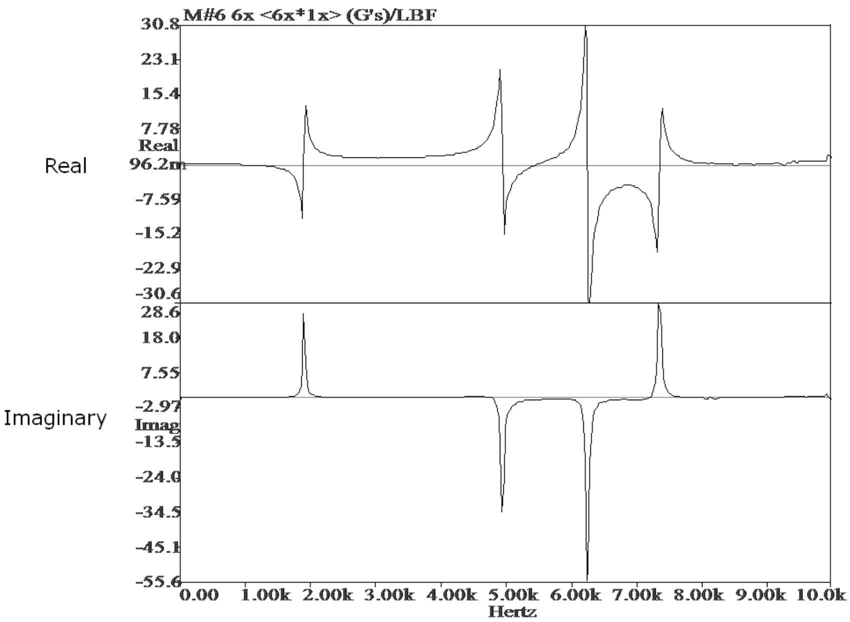


Figure 8-13

Curve-fitting modal test data

Modal analysis software is ultimately used to identify the natural frequencies and to animate the structure so that you can better understand the mode shapes. By definition the phase must change 180 degrees in order to validate the presence of a resonance. Therefore we can use the magnitude and phase data to identify the resonances and to estimate the damping. We use curve fitting techniques to get the most accurate estimate of the natural frequency and to assess the slope of the curve – it is the slope that indicates how heavily damped each mode is. A number of curve fitting techniques are available. They can utilize the circular shape of the Nyquist plot, or the gradient of the slope on the Bode plot, and there are other techniques. In modal analysis this is a very important step – but it is beyond the scope of this Category III course.

Animating the mode shapes

The modal analysis software will take the natural frequency, magnitude and phase information from each tested point and generate an animated display of the defined structure. You may display an animation of each mode shape. You can use this information visually in order to assess how the structure could be modified in order to move the natural frequencies away from machine operating frequencies (that are exciting the resonances), or you can utilize sophisticated software available is some of the packages to calculate the new natural frequencies when modifications are made. This is commonly called “structural dynamic modification” or SDM.

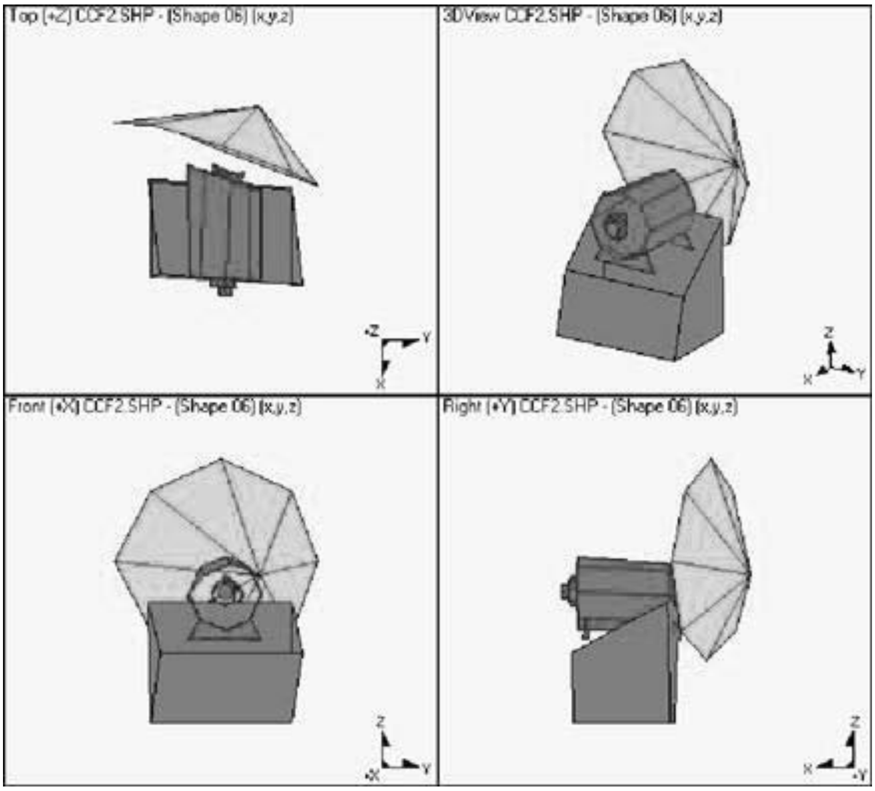


Figure 8-14

FRF Modal measurement tips

Here are a few additional testing tips:

- FRF modal testing requires that all inputs and responses be measured simultaneously.
- FRF's impact testing cannot be measured on operating machines where internally generated forms of energy exist.
- FRF modal testing requires calibrated force hammers and response sensors for accurate estimates of system properties.
- Make a test sketch before beginning testing.
- Try exploratory impacts at the beginning of the testing. Find a driving point where all modes can be seen from all measurement locations.
- Always measure the Driving Point.
- Use the appropriate hammer size and tip hardness.
- Use 4-6 averages for impact testing.
- Check for good coherence (> 80%).

The benefits of modal analysis

Although the animations displayed in the modal analysis software may look the same as the animations seen in ODS software, the data used is very different, and the information gained is very different. Modal testing does prove the existence of natural frequencies and can estimate the damping characteristics.

Further, modal analysis can help you to estimate how various structural modifications will affect these mode shapes. This can save you a great deal of time and money as compared with the experimental techniques used with ODS testing – i.e. visually estimating the best places to make design changes, making them, and retesting to see if the changes were successful.

Finite Element Analysis (FEA)

Modal analysis is based on experimental data – we must have the structure in order to perform the test. It is possible to generate the same information about a structure while it is still being designed.

FEA is a purely numerical method of estimating the natural frequencies, damping and mode shapes. Modern FEA packages are very powerful. In addition to the mode shape data, FEA can be used to model the stresses that will exist in a structure (to determine whether it will be able to perform the required function without failing). It can be used to model heat transfer; it can be used to simulate failure modes (collisions and more); and it can be used to model fluid dynamics. FEA allows the designer to be confident that a structure will be fit for purpose before the first prototype is built. Modal analysis may then be used to test the object, compare the test data to the FEA model, and refine the model accordingly. With this review in place, further design modifications may be made.

Finite Element Analysis is performed using “finite elements” – building blocks that model the material used in the object, and the joints/bonds used in construction. The model understands the material properties, the mechanical properties (how joints/bonds moved), and the dynamic properties (natural frequencies, damping and mode shapes).

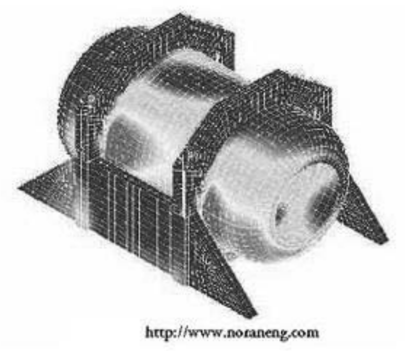


Figure 8-15

Figure 8-16 is an illustration of an FEA model of a SAAB automobile. A simulation has been performed to determine the damage that will be done (both to the automobile and passengers) when the car collides with a wall.

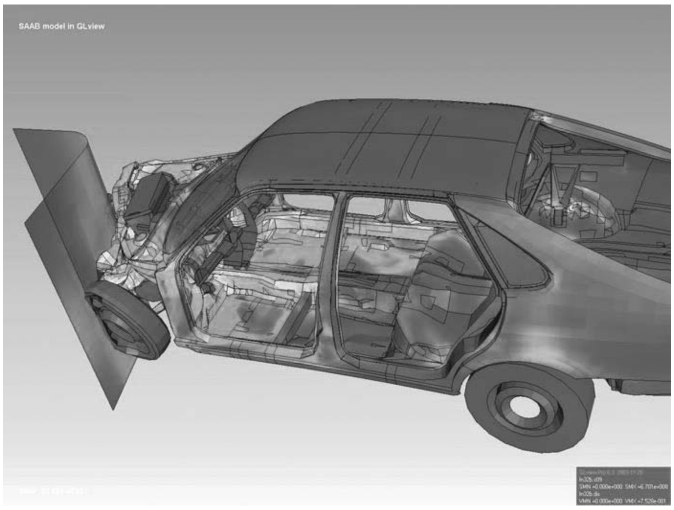


Figure 8-16

Of course, considerable computer power is required in order to create such a complicated simulation.

In Figure 8-17 we can see a model of a fan with a simulation of a blade failure. A test was performed and the resultant damage was compared to the simulation – the results were very close. Of course, it is far less expensive to simulate failures, and to experiment with design changes in order to arrive at the best design.

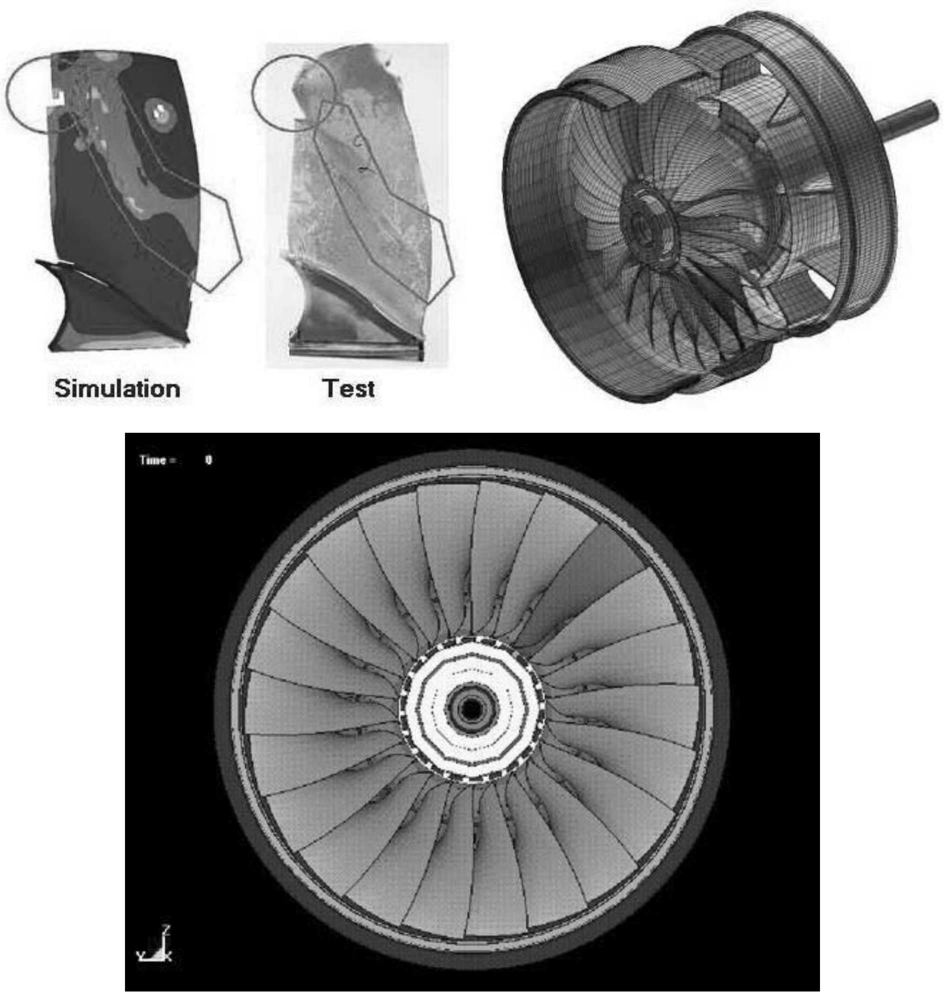


Figure 8-17

Combining modal analysis with finite element analysis

As mentioned previously, modal analysis is often performed on the prototype in order to test the accuracy of the finite element model. Changes can be made to the model until it matches the experimental results. Once the model is refined, additional simulations can be performed, and further improvements can be made to the design.

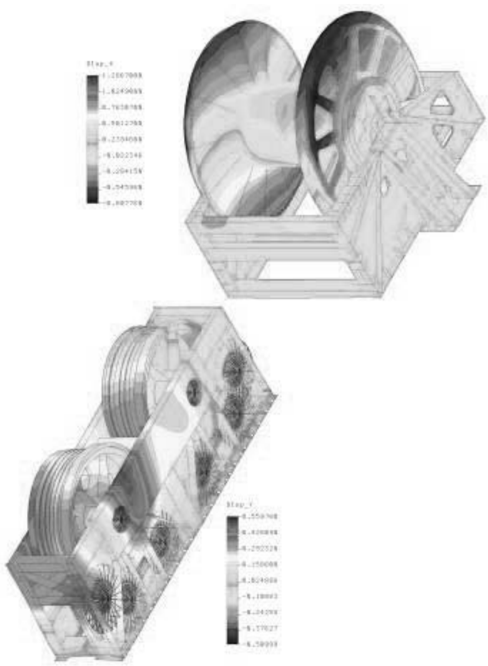


Figure 8-18

Vertical Waste Water Pump

This vertical waste water pump, like so many other machines in industry today, had high vibration due to resonance. The problem began after the old motor failed and a new motor was installed in its place.

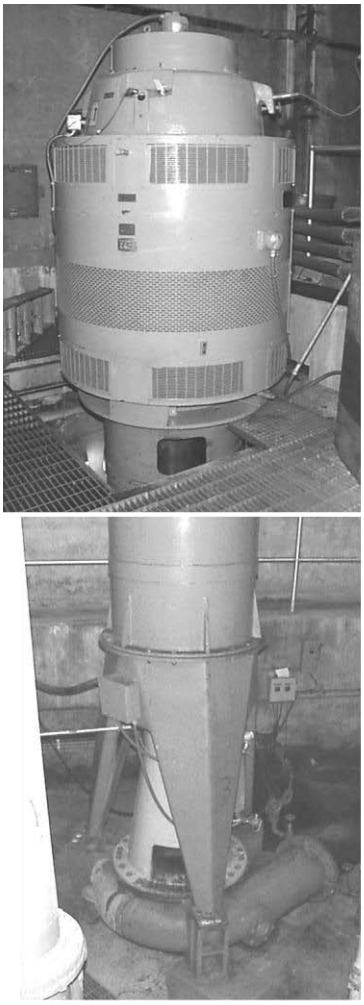


Figure 8-19

This case history details the diagnostic testing which began with route based vibration measurements and progressed to resonance testing, Operational Deflection Shape, modal analysis and finally Finite Element Analysis.

ODS and Modal analysis are powerful tools that enhance an analyst's ability to understand the sources of vibration. Resonance problems are easy to identify but difficult to solve. In the case of this vertical waste water pump, a Finite Element Analysis was needed to evaluate potential structural modifications to move the offending natural frequency away from the forced vibration – thus eliminating resonance.

Vertical Pump Case History

The Problem: Vibration on Pump #1

Background: Pumps 1-4 are sewage pumps. The pumps are Morris two-vane, vertical, centrifugal pumps. The drive motors on pumps 1 & 2 are new, 800 HP, U.S. induction Motors, on variable frequency drives. The motors on pumps 3 & 4 are Reliance motors with liquid rheostat speed controls and have been in service for decades. The operating speed range of the pumps is 700-890 rpm. The pump is bolted to the basement floor and the motor is supported on top of a tube enclosing a 15 foot long drive shaft. Figure 1 is a picture of the motors on pumps #1 and #3.

Figure 8-20 - Left – New Motor (pumps 1 & 2), Right – Old Motor (pumps 3 & 4)

The vibration problem on Pump #1 began after its original motor *(an 800 HP, wound rotor, Reliance motor with a liquid rheostat speed control)* was replaced with a new U.S. motor. Since installation of the new motor, the vibration on the machine has been extremely rough. Other characteristics of the problem are listed below.

The new motor is a different design and weighs about 800 pounds more than the old motor. A mounting plate was fabricated to connect the new motor to the top of the tube.

The highest vibration was over .9 inches/second–peak (IPS) at twice rotational speed
The problem direction was in-line with the discharge piping
The worst vibration occurs when the pump speed is above 820 rpm
The pump could not be operated at the desired speed due to excessive vibration
The VFD had to be programmed to exclude the 820-890 rpm speed range

Baseline Data – Baseline vibration data was measured with the pump operating at 870 rpm (14.5 Hz.). The largest vibration was .7 IPS – peak in the discharge direction on the motor. **Error! Reference source not found.** shows spectra measured on the motor, tube and pump.

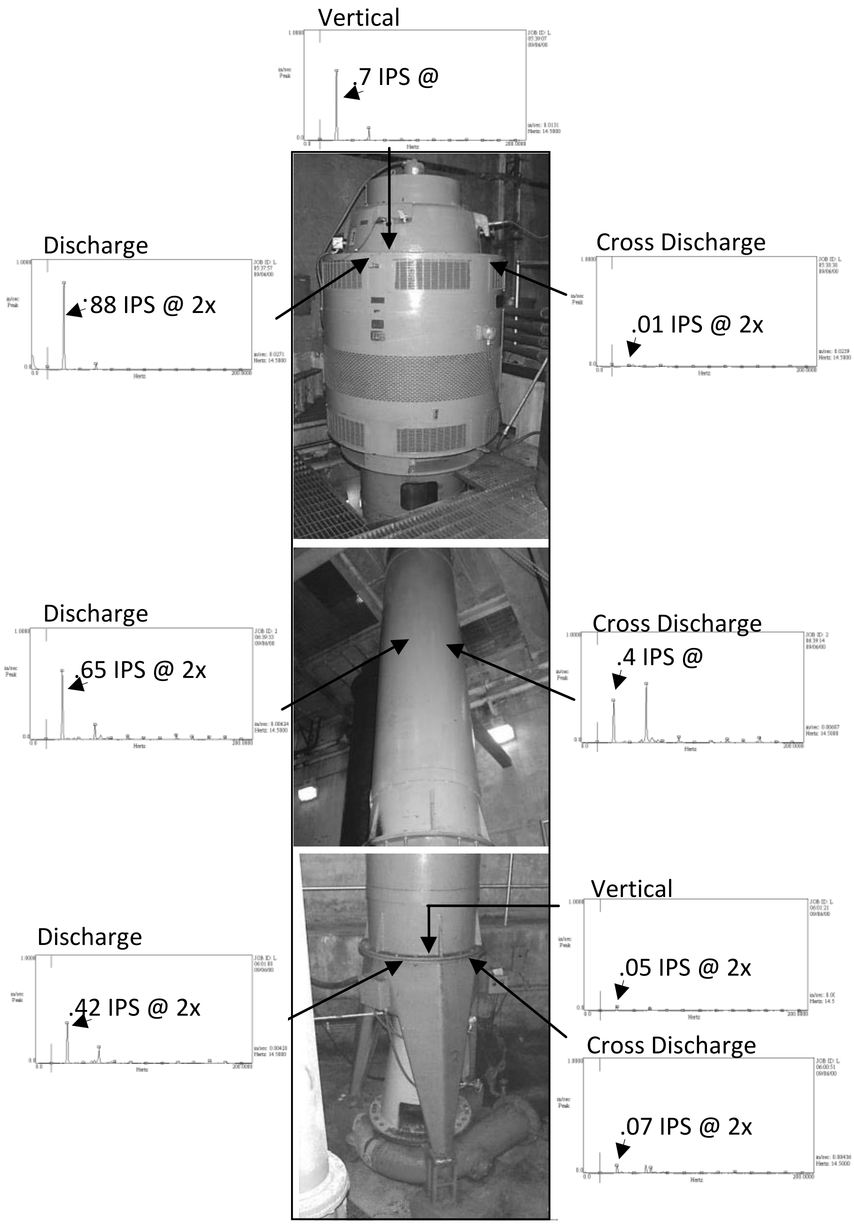


Figure 8-21 - Baseline data

Coast-down Testing:

The pump was operated at maximum seed (899 rpm or 15 Hertz). A tachometer was used to measure motor speed. Accelerometers were placed at the top of the motor in the discharge and cross discharge directions. The machine speed was decreased slowly using the VFD control. A Bode Plot of vibration and phase versus rpm was recorded during coast-down using the **Analyze | Monitor | Monitor Peak/Phase** function on the 2120 Analyzer.

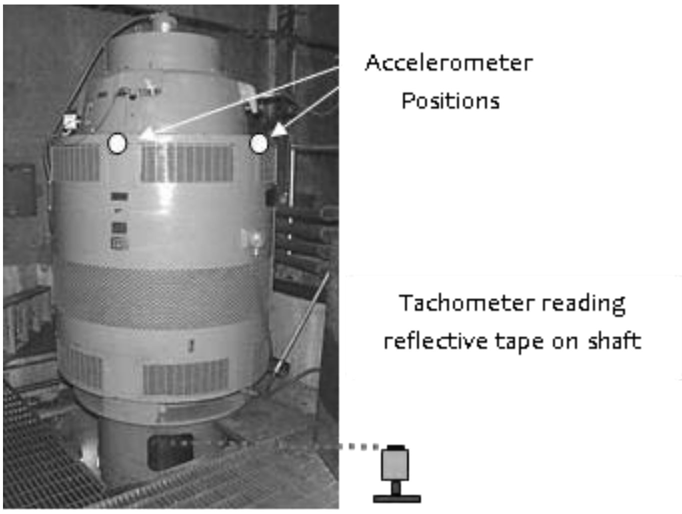


Figure 8-22 - Setup for Coast-down Testing

The Monitor Peak/Phase function was configured to measure coast-down vibration and phase at 2x turning speed. The screen below shows the set-up.

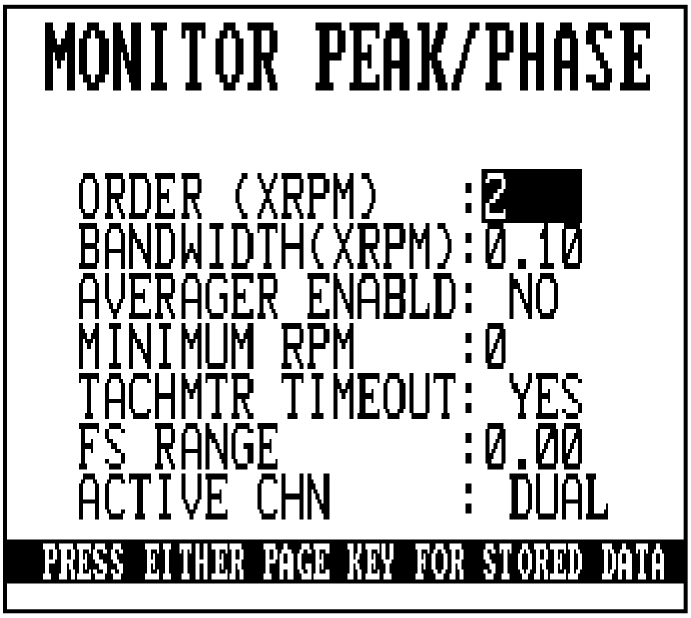


Figure 8-23 - 2120 Setup for Coast-down Testing

The Bode plot in Figure 8-24 is a trace of 2x turning speed vibration and phase during coast-down. It indicates that at speeds below 830 rpm, the vibration on the motor was smooth with less than .1 IPS in the discharge direction. Above 830 rpm, the 2x vibration increased rapidly and peaked at 0.9 IPS when the motor speed was 876 rpm. The 2x vibration frequency at this point was 1752 CPM or 29.2 Hertz. A phase changed of about 180° was noted through the amplification area.

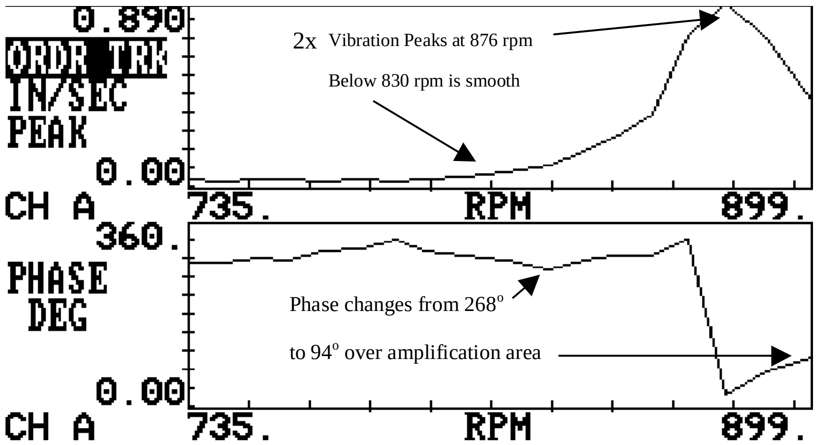


Figure 8-24 - 2x Coast-down Bode Plot data (discharge direction)

The Bode plot in Figure 8-25 shows the 2x turning speed vibration on the motor during coast-down in the perpendicular to discharge direction. It appears that there is a natural frequency just above maximum speed. The vibration increased rapidly as the pump approached top speed. The maximum vibration in this direction was about .44 IPS.

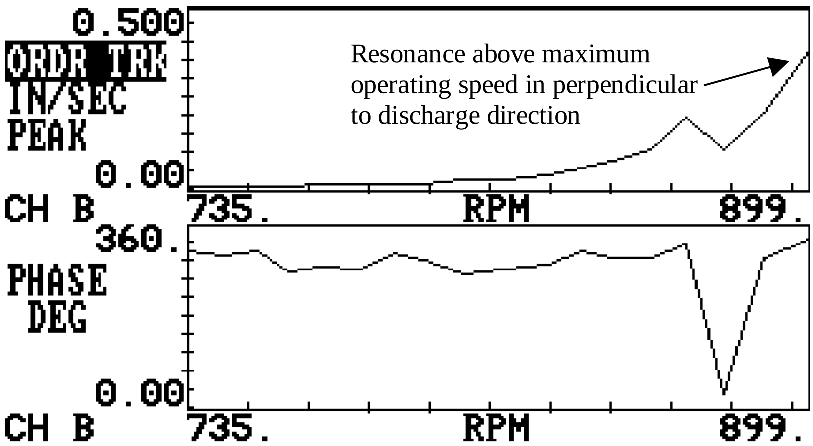


Figure 8-25 - 2x Coast-down Bode Plot data (perpendicular to discharge direction)

Coast-down Testing Results:

The coast-down data identified a natural frequency[1] at 29.2 Hertz. Amplification due to resonance occurs when the pump speed is between 820 – 890 rpm. The resonance is coincident with 2x rotational speed. The 2x turning speed vibration level is amplified to .9 inches/second – peak at 29.2 Hertz when the pump speed is 876 rpm. The direction of the vibration is in line with the discharge pipe. The coast down data indicates vibration at 2x rotational speed is exciting a natural frequency on the machine.

[1] A natural frequency is the frequency at which a part likes to vibrate. Resonant amplification results whenever forced vibrations coincide with natural frequencies in a system. At resonance, a small amount of excitation can be greatly amplified. The amount of amplification depends on the system damping characteristics.

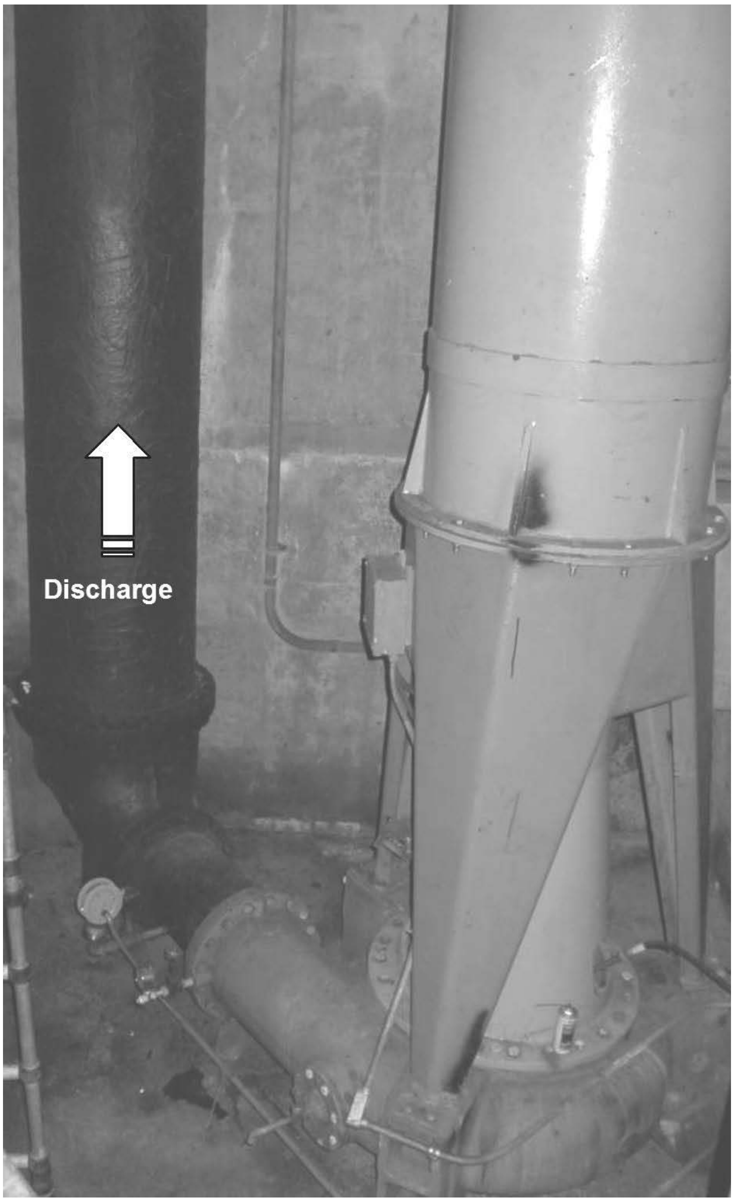


Figure 8-26 - Pump 1

Impact Testing:

Once the pump was shut down, impact testing was performed on the motor. One accelerometer was used to measure response to an impact made with a three pound impact hammer. The impact and response were measured at the same position and direction on the motor. Both the discharge and cross discharge directions were tested.

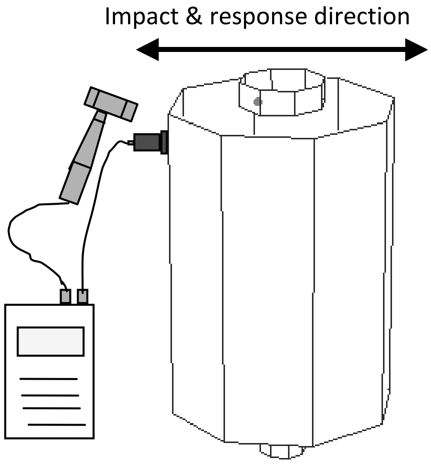


Figure 8-27 - Impact test setup

Impact testing found natural frequencies in the discharge direction at 5, 30 and 39 Hertz (Figure 8-28)

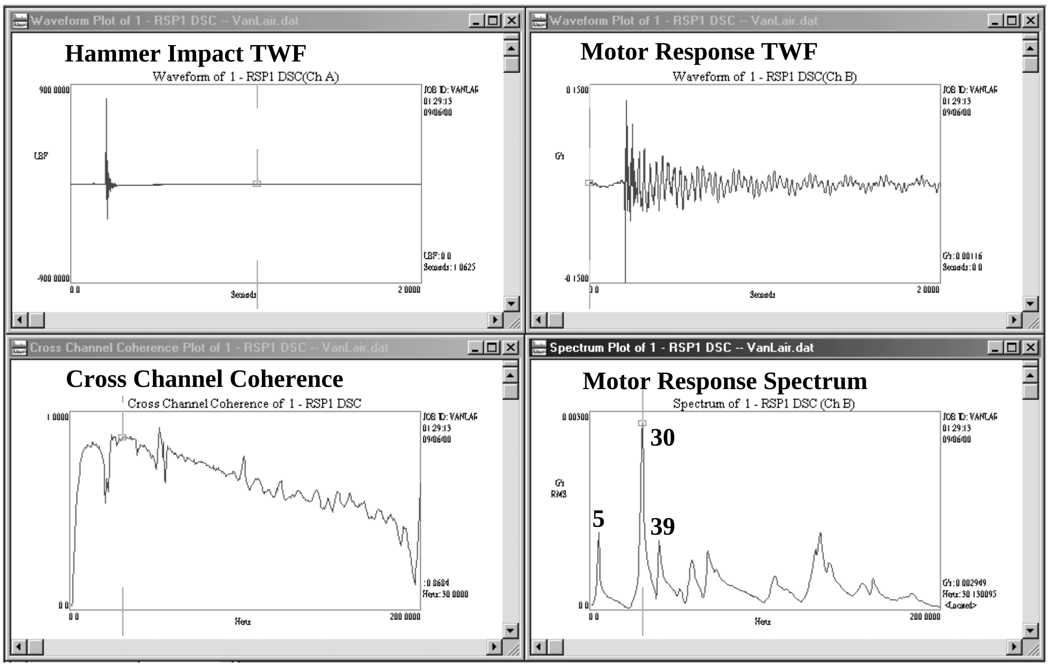


Figure 8-28 - Discharge Direction Impact Test

Impact testing identified natural frequencies in the cross-discharge direction at 5.1, 31 and 40 Hertz.

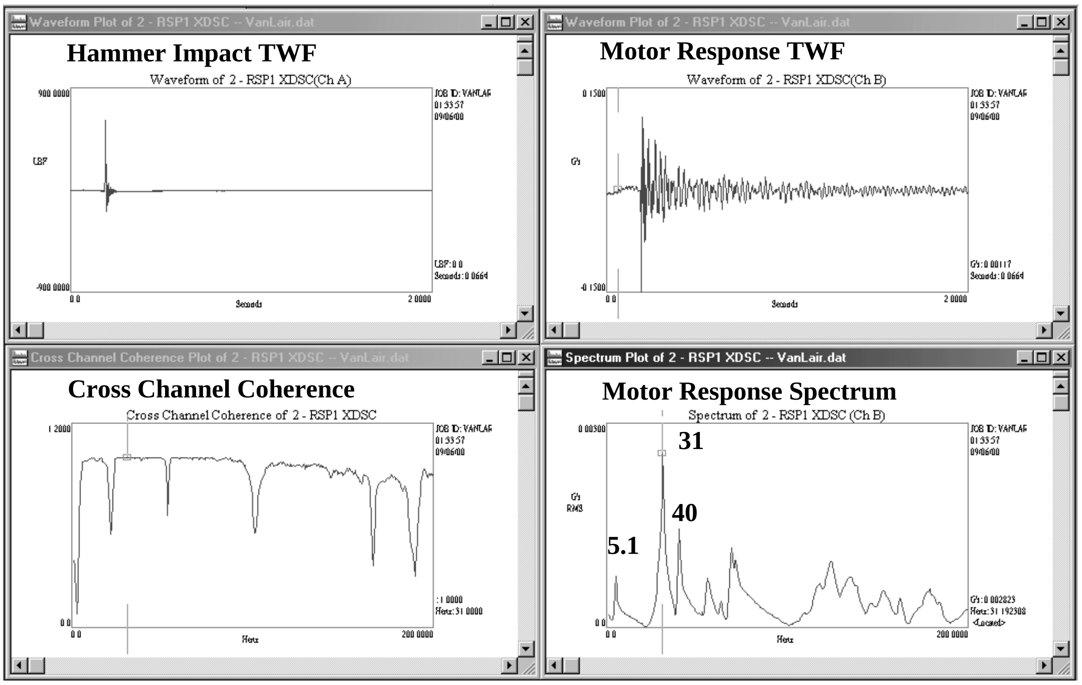


Figure 8-29 - Cross-Discharge Direction Impact Test

Impact Testing Results:

Natural frequencies were found at 5, 30 and 39 Hertz in the discharge direction. When the pump is operating in the optimum speed range of 830-890 rpm, the second harmonic of turning speed is coincident with the 30 Hertz natural frequency and results in resonance.

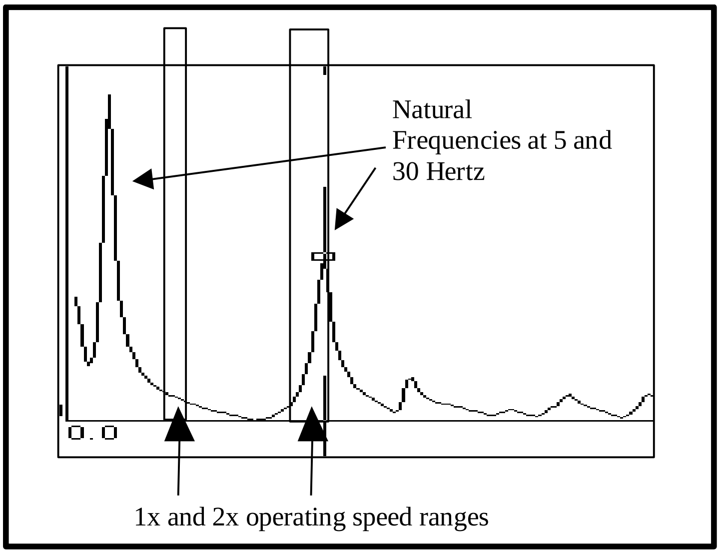


Figure 8-30 - Cross-Discharge Direction Impact Test

The pump has two vanes. Vane-pass frequency *(pump speed x # vanes)* and misalignment are potential sources exciting the resonance at 2x turning speed. The coast-down data indicates that the amount of 2x turning speed vibration is very small below 830 rpm. This means that vane pass or misalignment is very small and probably cannot be reduced any further.

The natural frequency at 5 Hertz is not excited by any forced vibration over the operating speed range of the pump. Vibration at this frequency is noticeable only during coast-down of the machine when the entire building shakes as the pump speed passes through the 5 Hertz natural frequency.

The natural frequencies in the cross-discharge direction were very similar to those found in the discharge direction.

Natural Frequency Comparison – pumps 1-4:

Impact tests were completed on three other pumps. Both the discharge and perpendicular to discharge directions were measured. The natural frequencies are listed in table 1. The pumps were offline during impact testing.

	P 1	**P 2**	**P 3**	**P4**
Natural Frequency (Discharge Dir) - ***cpm***	**30.0**	**22.5**	**33.7**	**28.6**
Natural Frequency (X-Discharge Dir)- ***cpm***	**31.2**	**22.7**	**36.8**	**28.5**

Table 8-1 - Natural Frequencies and Stiffness Values for Pumps 1-4

Note: All machines have the same pump and configuration. Pumps 1 and 2 have new US Motors. Pumps 3 and 4 have the old style Reliance Motors.

On each of the four pumps, the natural frequency near 2x operating speed appeared lightly damped, as indicated by a tall peak with a narrow skirt. This means that a small amount of forced vibration at 2x will be greatly amplified. It also means that a small change in pump speed has a great impact on vibration. The graph below compares the natural frequencies near 2x operating speed for each pump. The difference between the pumps may be due to different boundary conditions. Pumps 1 and 2 have new motors. Pumps 3 and 4 have old motors.

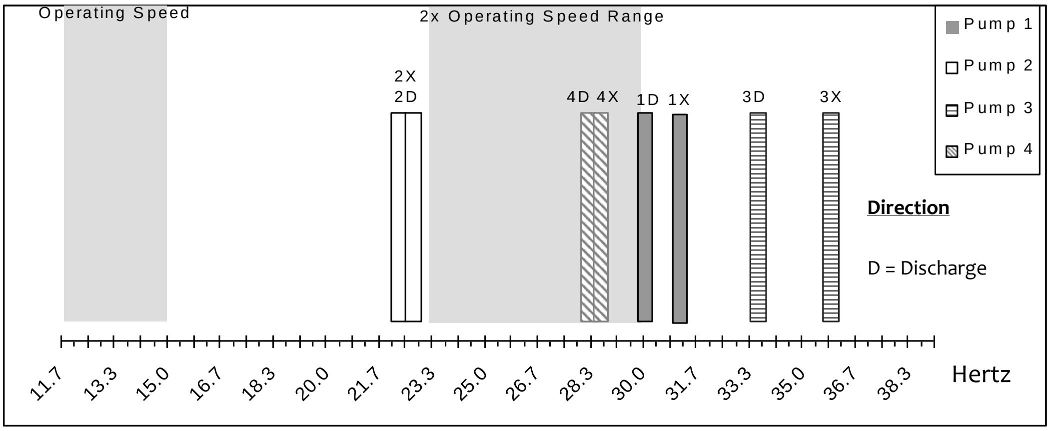


Figure 8-31 - Natural Frequencies near 2x Operating Speed Range for Pumps 1-4

Pump 1 *(discharge direction)* and Pump 4 *(both directions)* have natural frequencies in the 2x operating speed range. Based on the impact data measured on Pump 1, amplification factors of 10x-15x can be expected when 2x pump speed is near the 30 Hertz natural frequency.

The natural frequencies for pump 4 are below the upper end of the 2x operating speed range. Pump 4 usually operates at full speed and 2x operating speed is far enough away from the natural frequency that resonance will not occur.

The natural frequencies for Pumps 2 and 3 fall outside of the 2x operating speed range.

Operational Deflection Shape Test:

An Operational Deflection Shape Test[1] (ODS) was completed on the #1 Pump. The purpose of the ODS was to show the shape of the structure when operating at resonance. ME'scope Visual ODS-pro software was used for the test.

In order to generate operational deflection shapes, the ODS program requires:

1) Phase and magnitude data for each measured position

2) A structure drawing of the machine

Data were collected on the #1 pump using a CSI 2120 2-channel analyzer with the Advanced 2-Channel downloadable program (DLP). The DLP gives the 2120 additional cross channel

[1] An **O**perational **D**eflection **S**hape (ODS) is a method of analyzing the vibratory motion of rotating equipment and structures under normal operating conditions. An ODS is an extension of phase analysis. A computer generated model of the machine is animated with order tracked phase and magnitude data, transmissibility measurements or simultaneously measured time waveforms.

capability and allows storage of cross channel measurement data to the analyzer memory. The Advanced 2-channel DLP facilitates ODS data collection.

The pump was operated normally during the ODS testing. An operating speed of 870 rpm was chosen because it was slightly below the actual natural frequency and did not cause an excessive amount of vibration.

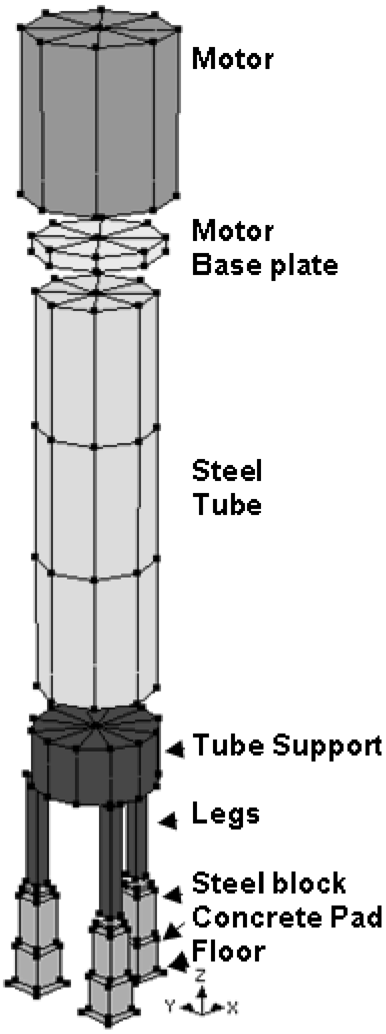


Figure 8-32

One accelerometer was positioned at the top of the motor, in the discharge direction, and used as the reference. A second accelerometer was moved to each position and direction where cross-channel phase transmissibility measurements were made. A total of 250 measurements were made on the machine. All bolted or welded joints were measured so that broken welds, looseness, soft joints and weakness could be identified.

The structure to the right was created in ME'scope ODS software. The black dots indicate the measurement points.

After completing the measurements, the data were downloaded from the analyzer to a computer using CSI's VibPro software.

VibPro is a CSI software product used to download, analyze and print data collected with the Advanced 2-Channel and Advanced Transient DLP's. It is also used to export measurement databases to format compatible with ME'scope software.

In the ODS software, the measurement file is linked with a structure drawing to produce animation at the desired frequency. The graphic below shows the structure's shape at 2x turning speed *(29 Hertz)*. It *indicates* that the pump tube is bending. The shape approximates the second bending mode of a cantilevered structure.

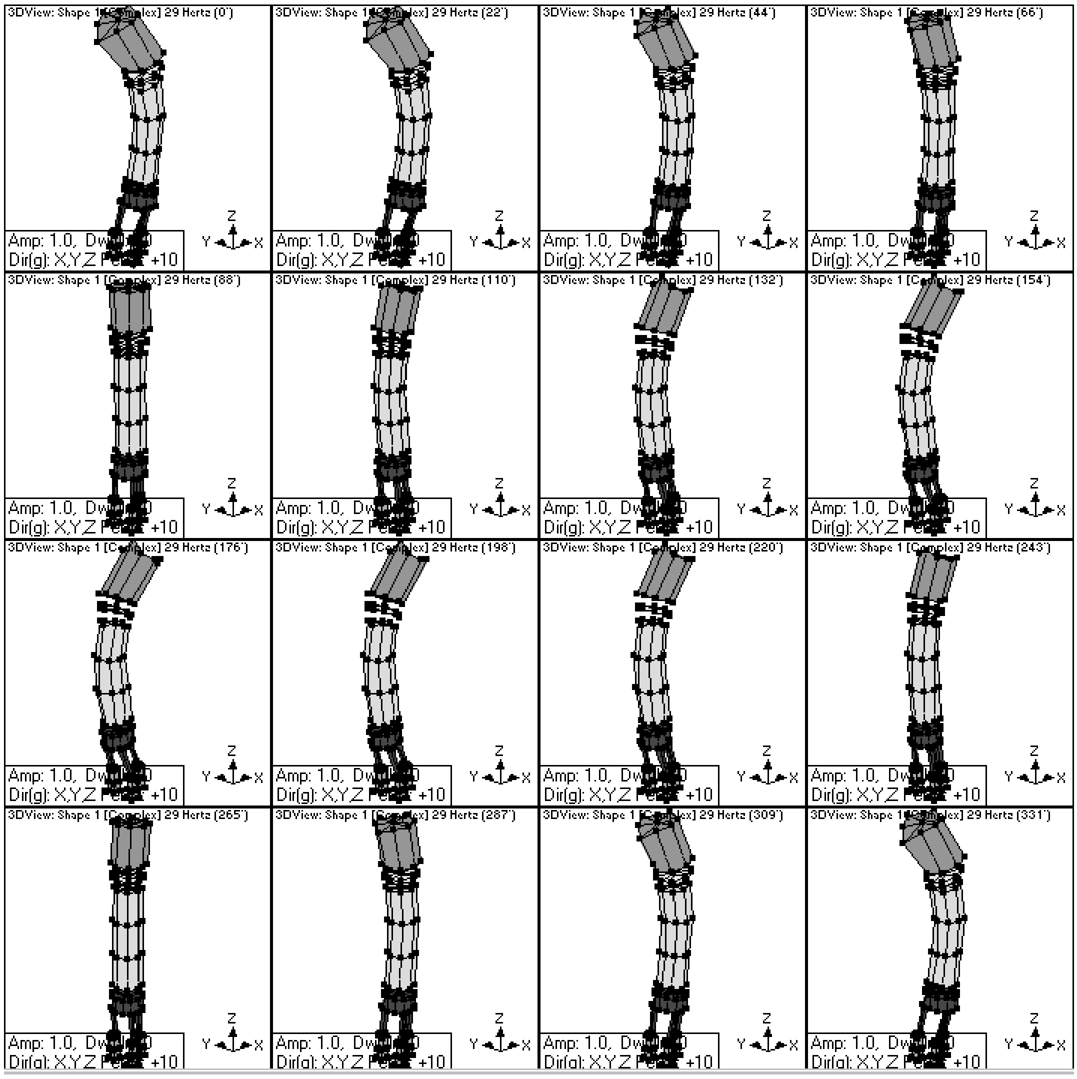


Figure 8-33 - ODS Animation of Natural Frequency at 29 Hertz

The ODS showed looseness on one leg of the pump base. Another ODS study was made of the three legs. Many additional points were measured to get a better view of the problem. The new structure drawing of one leg is shown in Figure 8-35.

Figure 8-34 - Picture of one Pump Foot

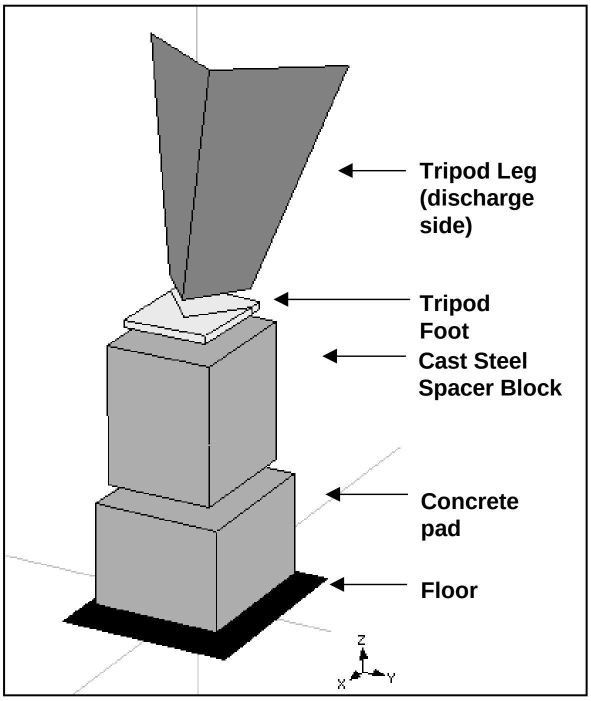


Figure 8-35 - ODS Structure of the Pump Foot

The animation of the leg data confirmed looseness. Relative motion was observed between the cast steel spacer block and the concrete pad. The motion was largest on the leg closest to the discharge. The predominant motion was in the vertical direction.

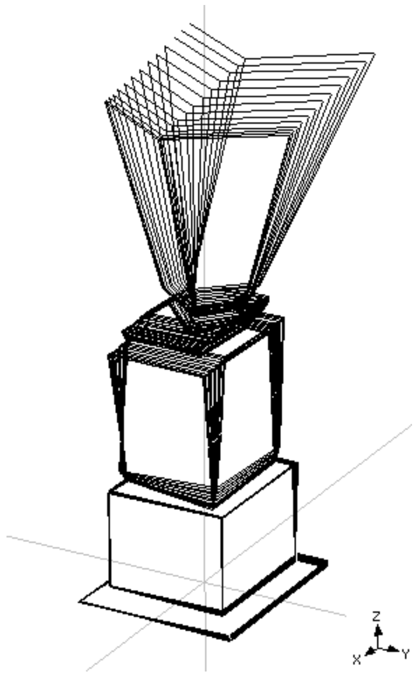


Figure 8-36 - ODS Animation of Discharge Side Leg

It was thought that the loose foot might have lowered the natural frequency of the pump structure and resulted in resonance at 2x operating speed range.

The bolts holding the steel shim block to the concrete were checked and tightened. No change in vibration was noted. The vibration did not change even when the bolts were loosened.

ODS Test Results:

At 2x turning speed (29 Hertz) the shape of the pump structure approximates the second bending mode of a vertical cantilevered structure. The greatest motion was in the discharge direction.

The ODS test identified relative motion between the cast steel block and the concrete pier on the discharge side foot. The motion on this leg was noticeably larger than the other two legs. All bolts were tightened and the base was inspected. No obvious problems were found and it was determined that the loose foot was a separate, unrelated issue that was not contributing to the resonance.

Attempt to Stiffen the Structure:

An 8" x 8" timber was wedged between the concrete wall of the building and the tube. The timber was placed about half way down the tube in the discharge direction where the bending was greatest.

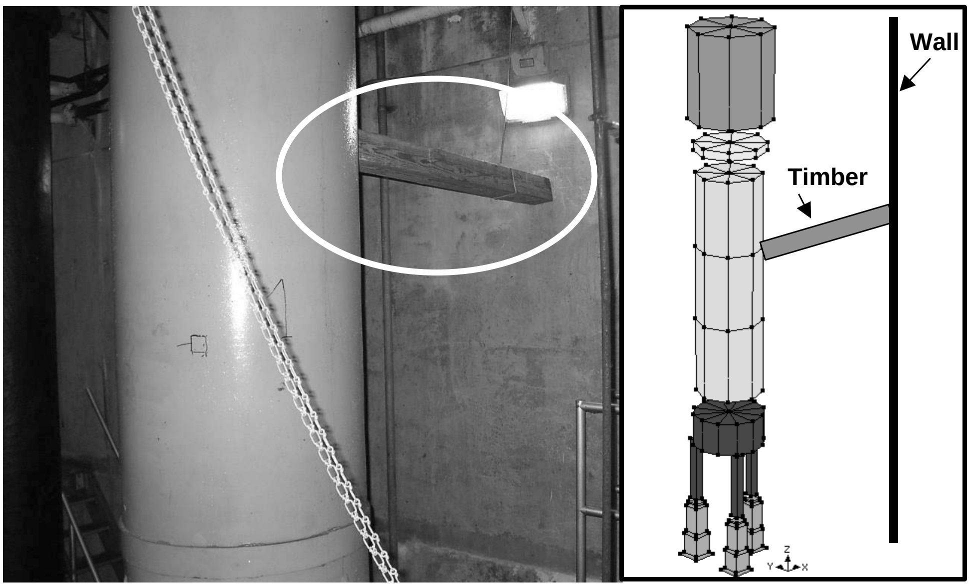


Figure 8-37 - Timber used to stiffen tube

Once the timber was wedged tightly in place, the pump was operated and another coast-down test was performed. The data showed that the natural frequency near 2x operating speed didn't change in frequency, but the amplitude was cut in half. The vibration was reduced from .89 IPS to .45 IPS. The vibration in the perpendicular to discharge direction also decreased. The timber was left in place after the test and secured with a chain-fall. The speed restriction was removed from the drive control and the Operating Department began to use the pump at the desired speed.

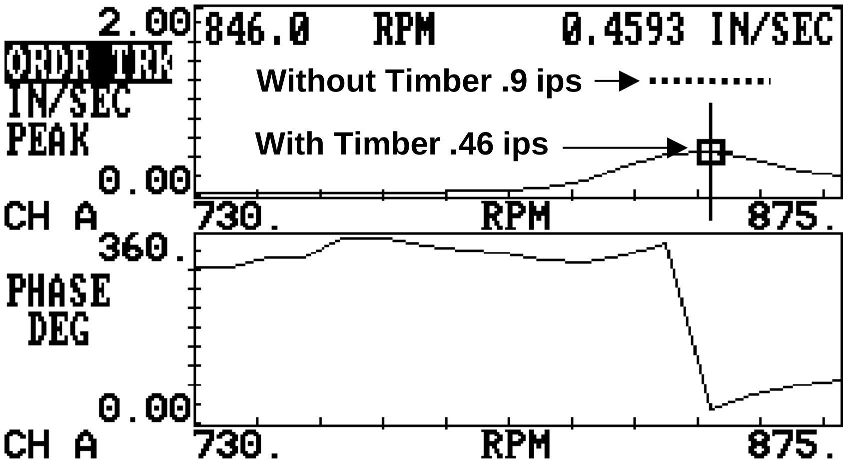


Figure 8-38 - 2x Coast-down Bode Plot of #1 Pump with an 8" x 8" Timber *Wedged Between the Building Wall and the Pump Tube*

Conclusions Based on ODS, Coast-down and Impact Testing:

The ODS animation showed that the structure was bending at 2x turning speed. Bending is an indication of resonance or structural weakness. The coast-down and impact tests confirmed a natural frequency at about 29 Hertz. The source of excitation causing resonance is 2x turning speed vibration from misalignment or vane pass frequency. The amount of 2x vibration was very small below the amplification curve. Reducing resonant amplification by correcting the source of 2x excitation is therefore not an option. Changing the speed of the pump is also not practical. Correcting resonance must be accomplished by changing the mass or stiffness of the pump structure.

How should mass or stiffness be changed to get the desired results? Sometimes, the modifications needed are obvious, simple and inexpensive to implement. In most cases, they are not. Making structural modifications by trial and error is a roll of the dice resulting in one of the following:

- Vibration is reduced at the intended frequency
- Vibration is not reduced at the intended frequency
- Vibration is reduced at the intended frequency, however changing mass or stiffness results in excitation of a different natural frequency.
- Vibration is reduced but the modification element itself begins to resonate.
- The modification results in mechanical or structural failure, machine down time, costly repairs or safety issues.

Natural frequencies cannot be eliminated. Changing mass or stiffness moves all of the natural frequencies. To correct the resonance problem on this pump, the natural frequency at 29 Hertz must be moved away from 2x turning speed.

Modal Survey:

The next step to correcting a resonance problem is to complete a Modal survey[2].

Modal data were collected on the #1 pump using the CSI 2120 2-channel analyzer with the Advanced 2-Channel DLP. The Advanced 2-channel DLP is required for modal data collection. Frequency Response Functions (FRF's) were collected at 194 degrees of freedom (DOF's).

At each measured point:

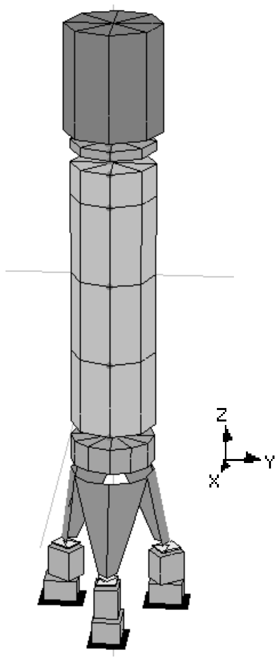


Figure 8-39

a) Natural frequencies were excited by striking a 12 pound impact hammer against the top of the motor *(discharge (X) direction*)

b) An accelerometer was used to measure a 0-100 Hertz response at each DOF. Four averages were acquired at each DOF.

The figure below is a Modal Peaks plot. It is a summation of all 194 FRF's and makes it easy to see the response of all points and directions in one view. Pump #1 had natural frequencies at

[2] Modal Analysis is an experimental method of determining the natural frequencies, damping values and mode shapes of a machine or structure. Modes are fundamental properties of a structure and are not dependent on forced vibrations. Mode frequencies are controlled by the mass, stiffness, damping and boundary conditions of the structure. Each mode will have a unique bending or twisting shape.

4.8, 29.81, 39.3 and 39.9 Hertz. Additional natural frequencies were found above 50 Hertz. Each natural frequency has a unique shape called a mode shape. After curve fitting the FRF data, the mode shapes were animated and analyzed.

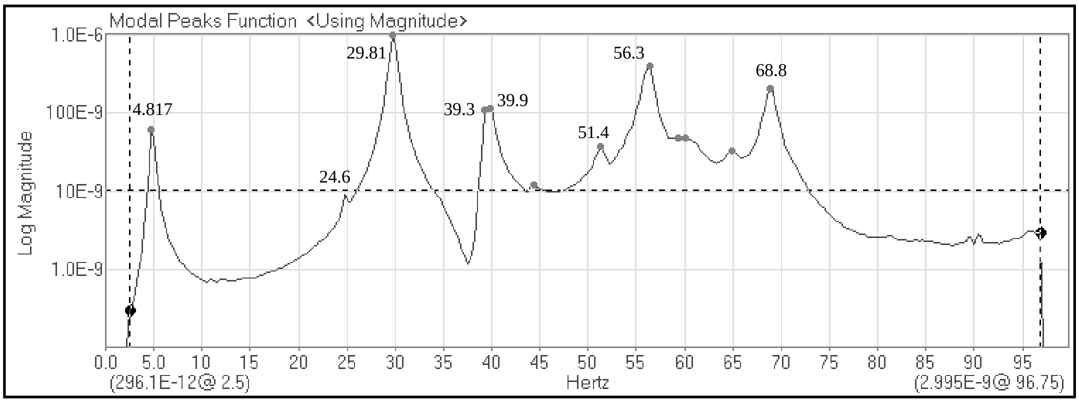


Figure 8-40 - Modal Peaks Plot is a Summation of all FRF Measurements

Modal Survey Results:

The first bending mode was 4.82 Hertz. No mechanical vibrations are exciting this mode into resonance except during start-up and coast-down.

The *frames view* of the 4.82 Hertz mode shape is shown in Figure 8-41 - Frames View of 4.82 Hertz Natural Frequency.

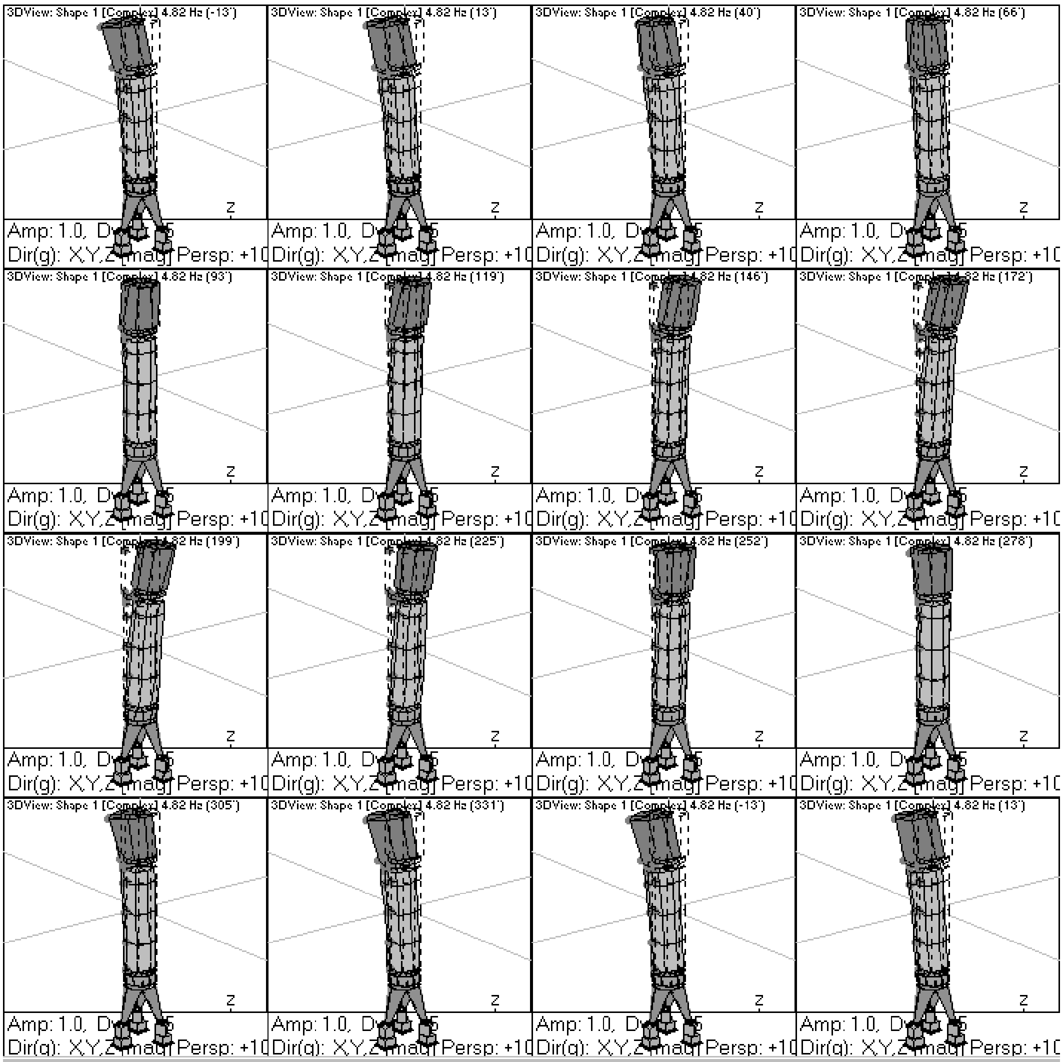


Figure 8-41 - Frames View of 4.82 Hertz Natural Frequency

The modal survey confirmed that the mode shape at 29.8 Hertz is the second bending mode of the pump structure.

A *frames view* of the deformation at 29.8 Hertz is shown in Figure 8-42.

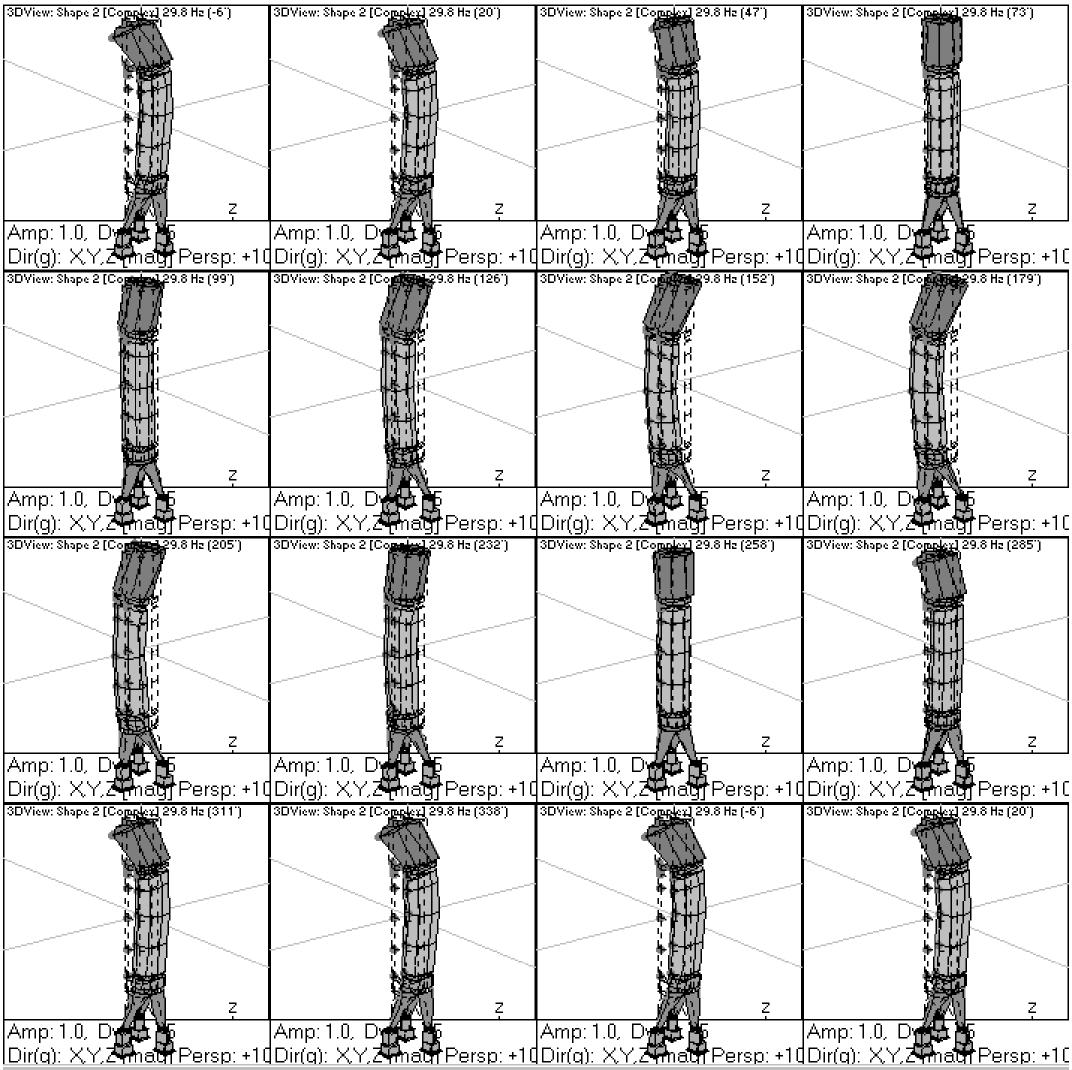


Figure 8-42 - Frames View of 29.8 Hertz Natural Frequency

Table 8-2 lists the Modal Survey results and characterizes each mode shape observed. Seven modes were found between 0-60 Hertz.

Frequency (Hz)	Mode Description
4.82	First bending mode
25	First torsional mode – not sufficiently excited in the modal survey
29.81	Second bending mode, Shaft/Support Pipe in Phase
39.3/39.9	Second bending mode, Shaft/Support Pipe out of Phase
51.4	Vertical mode
56.3	Third bending mode

Table 8-2 - Pump #1 Modal Frequencies

Modal analysis does not solve resonance problems. The SDM version of ME'scope software has some imbedded Finite Element Tools that can be used to estimate the required structural modifications. Once a modal survey has been completed and analyzed, determining the appropriate structural modifications may be intuitive. Often, it is not and the next step in correcting resonance is Finite Element Analysis.

Finite Element Analysis (FEA):

A Structural Engineer was contracted to complete a Finite Element Analysis[3] of the #1 Pump. The FEA is used to evaluate the effectiveness of potential structural modifications.

A finite element model of Pump 1 was created. The model is built after gathering all the available information about the structure. The information in the model includes all dimensions, material properties, discrete stiffnesses, component masses and boundary conditions. Some of the information is easy to obtain. Other information, like the motor's stiffness and inertia, is an educated guess.

Included in the FEA model were the concrete pads, metal blocks, pump base, pump volute, support pipe, motor shaft, motor adapter plate, and the motor. The motor was constructed using a combination of rigid and concentrated mass and inertia elements. Given the relatively large dimensions of the motor and the cantilevered configuration, the mass moments of inertia are important in the dynamic behavior of the structure. The only mass property data available on the motor was its total weight of 9200 pounds. Therefore, estimates of the moments of inertia were required to accurately predict the natural frequencies of the pump. The mass moments of inertia were calculated by approximating the motor geometry as a cylinder with uniformly distributed mass. The actual values used in the model are shown in Table 8-3.

		Mass Moments of Inertia^		
Weight (lbf)	Mass ($lbf\text{-}sec^2/in$)	I_{XX} ($lbf\text{-}sec^2\text{-}in$)	I_{YY} ($lbf\text{-}sec^2\text{-}in$)	I_{ZZ} ($lbf\text{-}sec^2\text{-}in$)
9200	**23.8**	**21950**	**21950**	**13750**

Table 8-3 - Motor Mass Properties and Weight

The principal mass moments of inertia are calculated at the motor center of gravity.

The pump assembly included numerous bolted interfaces between the structural components. Considerable effort was made to accurately model these connections. While the bolts themselves are considered to be rigid, only assumed preloaded areas, which are local to the

[3] A Finite Element Analysis (FEA) is an analytical technique that utilizes a mathematical model of a structure to predict its natural frequencies, damping values and mode shapes. Once the FEA model has been created, it is used to evaluate the effectiveness of proposed structural modifications. Sigmadyne, Inc. (http://www.sigmadyne.com) completed the FEA on Pump 1.

bolts, are used to connect the parts. These preloaded areas were modeled using special constraint elements that do not add stiffness to the bolted components themselves. This technique was used to model the connections between the concrete and metal blocks, the top of the metal blocks and the base plates of each leg of the pump base, and the top flange of the pump base and the bottom flange of the support pipe.

The finite element model was created for use in MSC/NASTRAN, a commercial finite element code. The model includes a total of 6575 beam, shell, and solid elements. Solid base, support pipe and adapter plates were modeled using shell elements. The thickness and internal design of the volute was not known. It was modeled as a uniform shell structure with a thickness of 5/8 inch. The motor shaft was modeled using beam elements connected to the model at the base of the motor, the top of the volute, and the top of the conical structure mounted on the volute. The conical structure inside the legs of the pump bases was modeled using a rigid element.

The entire model is constrained at the bottom surface of each of the three concrete blocks. Physically this represents the interface between the concrete blocks and the floor beneath the pump. The complete model is shown in Figure 8-43.

The FEA model is only as accurate at the information used in its construction. The FEA prediction of the natural frequencies will not be accurate if any of the information is incorrect. In addition, the estimates for structural modification may not be valid. Examples of things that can lead to inaccurate FEA results include:

- Incomplete or inaccurate information about the structure
- Weakened structural members due to rusting, corrosion and cracks
- Loose or weakened concrete base
- Stretched or loose bolts
- Incorrect estimates of material properties, masses, stiffness and boundary conditions

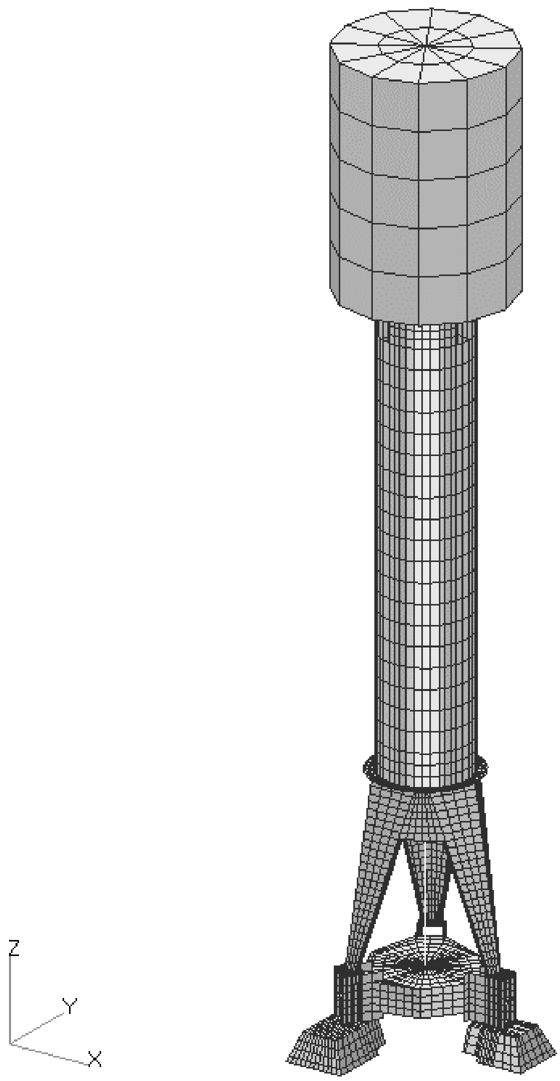


Figure 8-43

The assumptions used in the FEA model of the #1 Pump are listed below.

1. The weight of the motor, specified by the customer, was 9200 lbf. Since the mass moments of inertia are unavailable, they are calculated from this weight based on a diameter of 67.5 inches and a height of 92.5 inches. The mass moments of inertia are calculated using the equations for a solid cylinder with uniformly distributed mass.

2. The motor is assumed to be rigid over the frequency range of interest.

3. Concrete is assumed to be a homogeneous and isotropic material and to behave in a linear elastic manner. The following mechanical properties are used in the analysis:

   Modulus of elasticity: 2.0 MSI.

Poissonʼs Ratio: 0.2
Weight density: 145 lbf/ft^3

4. The rebar used in concrete does not significantly alter its mechanical properties.

5. Steel is assumed to be a homogeneous and isotropic material and to behave in a linear elastic manner. The following mechanical properties are used in the analysis:

   a. Modulus of elasticity: 30 MSI.
   b. Poissonʼs Ratio: 0.3
   c. Weight density: 481 lbf/ft3

6. The pump is fixed at the bottom surface of the concrete block. Any additional compliance due to the surrounding structure is ignored.

7. All dimensions used in the creation of the model were measured on Pump 1 as part of the modal survey or were supplied by Plant personnel.

As previously stated a modal survey is an experimental procedure used to measure both the natural frequencies and mode shapes of a structure. The accuracy and validity of the finite element can be evaluated and enhanced by correlating the predicted natural frequencies with the measured values. This correlation is based on comparison of the mode shapes of vibration rather than a simple comparison of the frequency values. This ensures that similar modes (i.e. the first bending mode of a cantilevered structure) are being appropriately compared.

Once the finite element model was evaluated and tuned using the modal survey results, a design optimization analysis was performed using the model. The objective of this analysis was to determine the changes in the natural frequencies due to alterations of the pump geometry.

FEA Results:

Both predicted and measured values for the first eight natural frequencies of Pump 1 are summarized in Table 8-4. Natural frequencies of a symmetric structure occur in orthogonal pairs. The physical significance is that the pump can actually vibrate (bend) in any direction based on the direction of the applied excitation. It is difficult to experimentally measure multiple modes with nearly identical frequencies. During the modal survey, the frequency resolution was insufficient to resolve the orthogonal pairs for the first two bending modes. The frequency separation of the next mode pair (motor shaft and support pipe oscillating out of phase) was sufficient and both were identified in the test.

Mode Shape Description	Measured Value (Hz)	Analytical Prediction (Hz)
1st Bending	4.8	4.4
1st Bending		4.4
Torsion	24-25	23.5
2nd Bending, Shaft/Support Pipe in Phase	29.8	29.2
2nd Bending, Shaft/Support Pipe in Phase		29.5
2nd Bending, Shaft/Support Pipe out of Phase	39.2	41.8
2nd Bending, Shaft/Support Pipe out of Phase	39.9	42.2
Vertical Extension	51.4	60.3

Table 8-4 - Correlation of Measured and Predicted Natural Frequencies of Pump 1

The correlation between the analytical prediction and the measured frequency values was very good.

A picture of the FEA mode shape at 29.2 Hertz is shown below. The drive shaft is vibrating in-phase with the support tube.

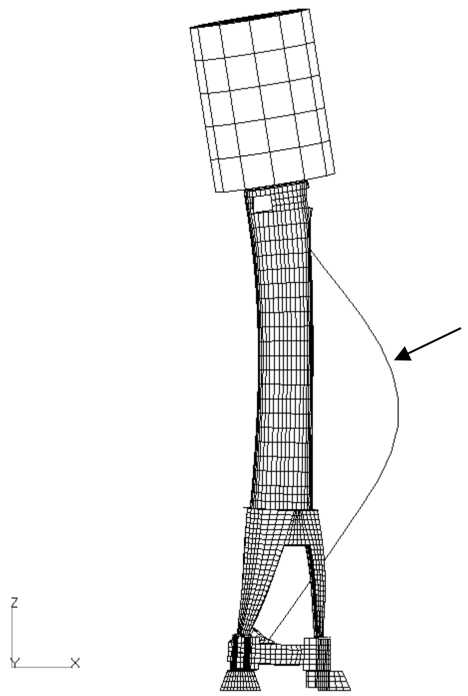


Figure 8-44 - Second bending mode of the Pump structure at 29.2 Hz.

Since the natural frequency at 29 Hertz is at the upper end of the 2x operating speed range, the structural engineer recommended increasing the machine stiffness. Increased stiffness results in a higher natural frequency. To accomplish the change, several potential solutions were evaluated using the FEA software including:

- Shortening the length of the pump tube
- Increasing the wall thickness of the tube
- A different motor
- Improving the attachment to the floor
- Bracing the tube to the building wall

The stiffness of the support pipe *(T_p)* was considered by treating the support pipe thickness as a design variable. It was found that this parameter has little effect on the modes near 29 Hz. The only feasible option was to reduce the length of the pump tube. The customer agreed that changing the tube length was an acceptable solution. Tube length reductions of 48, 62 and 72 inches were evaluated. All three models produced good results. Table 8-5 shows the change at each natural frequency based on pipe length reductions.

			Length Reductions		
Mode Shape Description	Baseline (Hz)	T_p=0.75" (Hz)	48" (Hz)	62" (Hz)	72" (Hz)
1st Bending	4.4	4.8	5.4	5.8	6.2
1st Bending	4.4	4.9	5.4	5.9	6.3
Torsion	23.5	25.5	24.6	25.1	25.5
2nd Bending, Shaft/Support Pipe in Phase	29.2	30.0	37.1	39.9	41.8
2nd Bending, Shaft/Support Pipe in Phase	29.5	30.2	37.3	40.2	42.2
2nd Bending, Shaft/Support Pipe out of Phase	41.8	39.2	49.6	54.1	58.5
2nd Bending, Shaft/Support Pipe out of Phase	42.2	39.6	49.8	54.2	58.6
Vertical Extension	60.3	64.9	64.4	66.0	67.3

Table 8-5 - Effect of Support Pipe Length on the Predicted Natural Frequencies of Pump 1

Decreasing the length of the support pipe and correspondingly, the overall height alters the natural frequencies of the pump. A length reduction of 48 inches increases the frequencies of the first pair of second bending modes (shaft and support pipe bending in phase) to approximately 37 hertz. An important assumption in this prediction is that all other stiffnesses remain unchanged. This is especially important given the number of bolted connections in the pump assembly. Figure 8-45 shows the change in 2nd bending mode natural frequency with respect to operating speed and 2x operating speed.

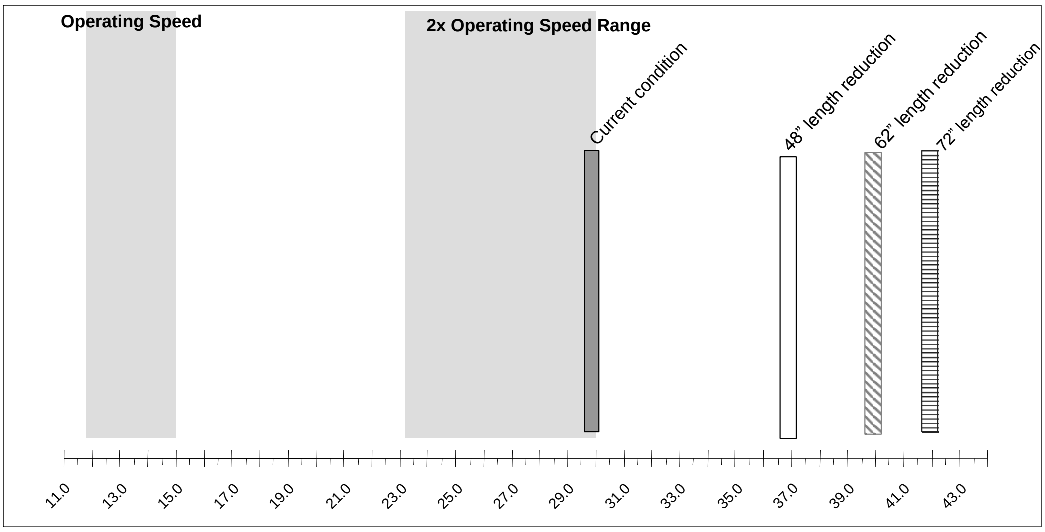


Figure 8-45 - Pump 1 Second bending mode Frequency Shift with Pipe Length Reduction

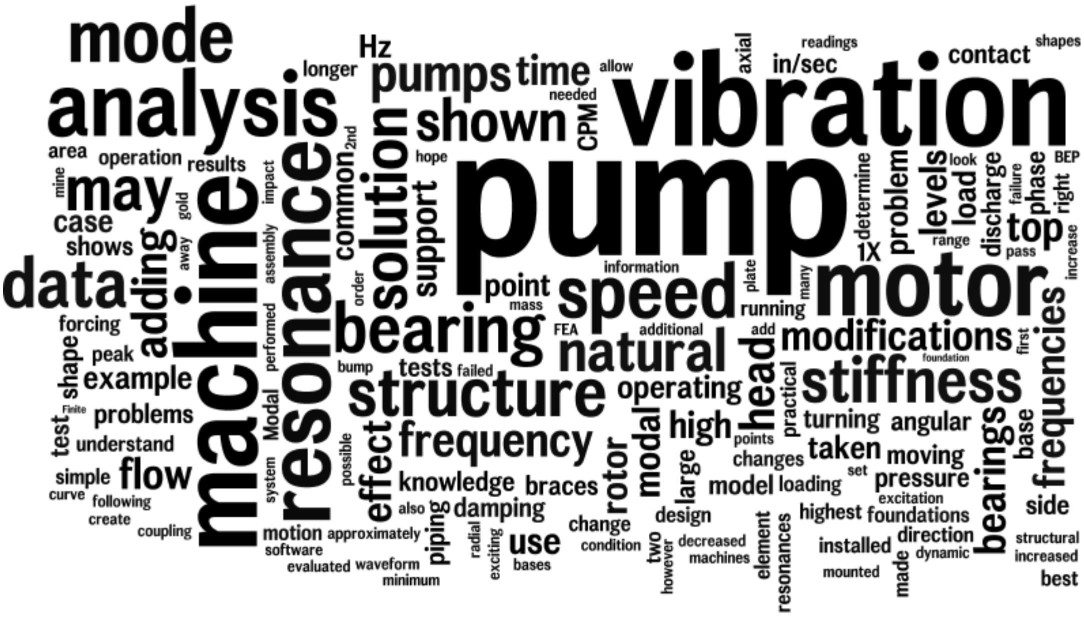
Measurements of the machine vibration did not indicate any forced vibrations in the area of 37 Hertz which might excite the (new) second bending mode natural frequency.

Summary:

Resonance testing and Operational Deflection Shape studies are useful tools for analyzing vibration problems. When resonance is identified as the problem, a Modal Survey is necessary to identify all of the natural frequencies and evaluate each mode shape. Without knowing the mode shape, it is impossible to know how to correct the resonance.

Resonance is best corrected by using FEA tools to evaluate the effectiveness of potential structural modifications.

The recommendations from this job were well received by the customer. Maintenance planned to reduce the length of the tube by 48". The work order was cancelled when engineering decided to replace all of the pumps within two years.



Chapter 9

Correcting Resonance Problems

Objectives:

- Describe a number of methods to deal with resonance problems
- These include: changing speed, altering forcing frequencies, changing stiffness, mass or damping or isolating the vibration.

In this chapter we will learn about various methods to deal with resonance problems. These include changing the input forcing frequencies, isolating the input forcing frequencies from the structure or altering the characteristics of the structure in terms of mass, stiffness or damping.

Correcting Resonance Problems

Utilizing phase measurements, run-up or run-down tests, bump tests, ODS tests or modal analysis, you may believe that a machine is exciting a resonance which is generating high vibration and reducing the life of the machine. Now it is time to consider alterations to the machine or structure so that the forcing frequencies of the machine no longer coincide with the natural frequencies, or so that the resonances are so heavily damped that they do not amplify the vibration to such an extent.

There are a number of possible methods for determining the best solution and then implementing the changes.

Change the machine speed

A temporary solution is to change the speed of the machine so that it no longer excites the resonance. For example, if the vane pass frequency of a pump matched a natural frequency, then you could reduce the speed of the machine by 15% so that the mode was no longer excited. Of course, in many situations this will not be possible.

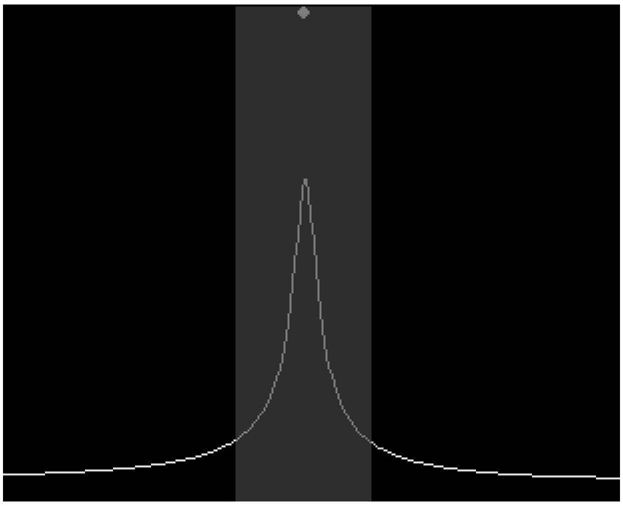


Figure 9-1

Change the stiffness

Typically, however, the aim is to change the stiffness so that the natural frequency is increased (or decreased) to a point where it is a minimum of 15% away from the forcing frequency.

For example, if the vane pass frequency was 8880 CPM (148 Hz) and it exciting a natural frequency, we would attempt to increase the stiffness of the machine so that the natural frequency was a minimum of 148 x 1.15 or 171 Hz. 15% is a rule of thumb – some people recommend 20%.

In brief, the typical approach is to stiffen structures with additional braces and support structure. But how do you determine how much is required?

There are three basic approaches:

- Use finite element analysis to model the structure and then make changes to the model in order to determine the best solution. This is major undertaking and would only be justifiable in special cases.

- Use modal analysis to test the structure, and use the "structural dynamic modification" software to estimate the effect of adding stiffness (mass or damping).

- Study the structure closely and engineer some modifications that will change the dynamics of the machine or structure. This takes some skill and experience, and a good knowledge of the existing mode shapes.

Finite element analysis (FEA)

If you have time and expertise, you can create a "finite element model" of the machine and structure which will allow you to model the effect of adding braces and/or stiffening.

Finite element analysis (FEA) is summarized at the end of the modal chapter and we cannot go into great detail here. Suffice to say that it is possible to create a mathematical approximation of the structure that enables you to study its various mode shapes. You can then model the effect of adding braces, isolators, etc. until you achieve the desired effect. This is a very challenging area…

Modal analysis

Modal analysis provides all the information you need to design modifications that will minimize the effect of the resonance(s). Modal analysis is discussed in greater depth in the modal analysis chapter, however we can summarize by saying that the tests provide information about the natural frequencies, mode shapes, and damping. You can use this information to visually determine the best modifications, or you can use special software.

Most modal analysis software program have an option that enables you to estimate the effect of adding stiffness, damping or mass to a structure.

Estimating the required structural modifications

If you are not able to use FEA or perform a modal analysis test *and* use the structural dynamics modifications module, then you are on your own! You must use your practical understanding to determine the best solution to the problem at hand. In most cases the most practical solution is to add stiffness. Adding mass or damping is often impractical. We will soon discuss damping, isolation, and the use of tuned absorbers, but for the moment we will focus on adding stiffness.

The most practical solution is typically to add stiffness between a point that is not moving (e.g. the foundation, the base of a column, etc.) and the point on the structure where vibration is highest – i.e. the anti-node.

For example, this structure probably has a first bending mode in the center of the span. However it you added stiffness to the center of the span (assuming it is mounted on a table-like structure), and the resonance was actually second bending mode, the brace would have no effect.

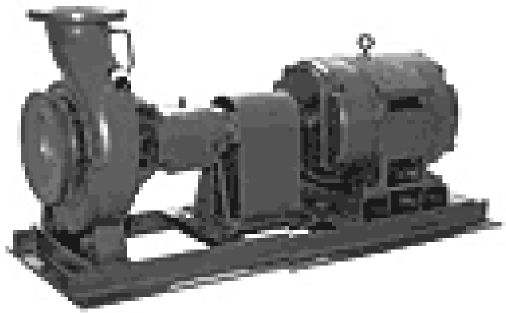


Figure 9-2

With this machine you may suspect that the motor is "bouncing up and down". However it may be twisting or rocking side to side.

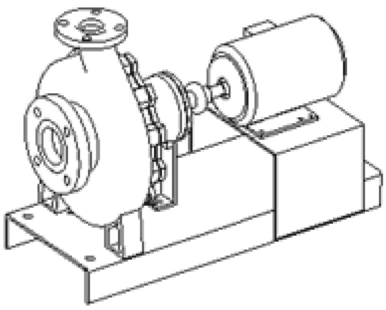


Figure 9-3

It is therefore important to understand the mode shape before adding braces or ribbing otherwise you may spend a lot of time and money with poor results.

Also be careful when welding on machine bases that you do not induce misalignment or soft foot – the heat and additional forces applied could alter the base flatness and position.

Case Studies

The following case studies provide some examples of problems encountered in industry with solutions that solved the problems.

Motor – Pump on Bent-Steel Base

In this common example we have a simple bent sheet steel base for pumps. It is common for these foundations to have natural frequencies in the 10 to 35 Hz range – right around the running speed.

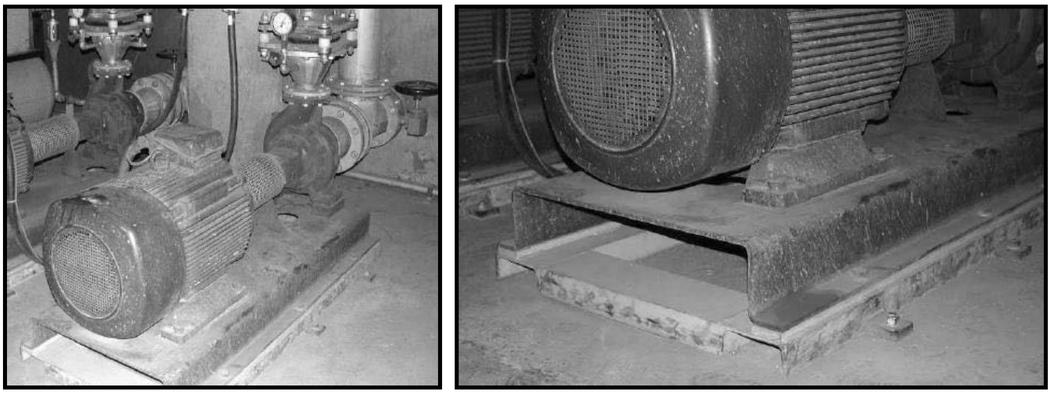


Figure 9-4

One solution is to fill the bases with concrete (and separate common foundations), as shown in Figure 9-5.

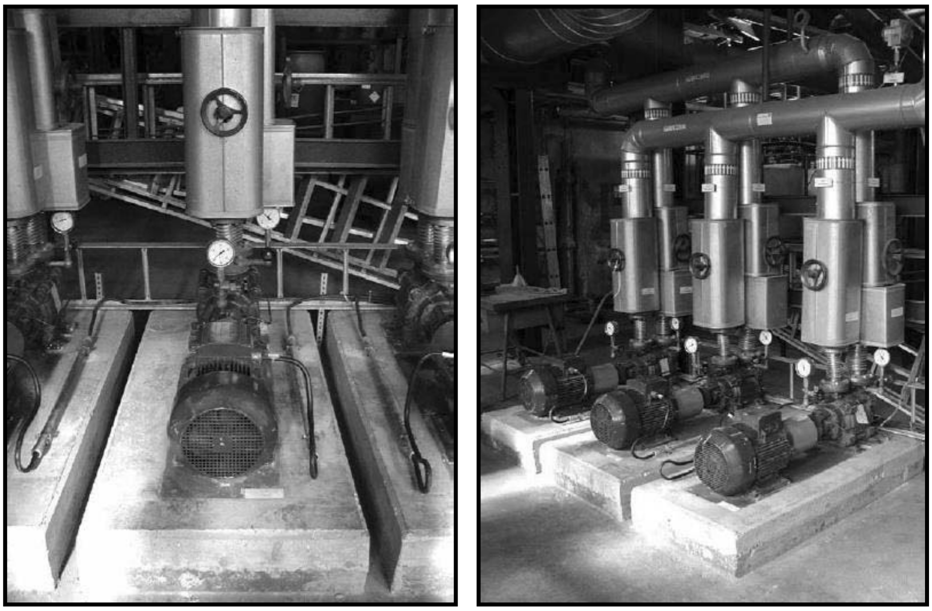


Figure 9-5

Overhung Pumps

Overhung pumps without adequate support can have a rocking resonance, with high vibration vertically and possibly axially. The solution is quite simple – add support under the flange. This greatly stiffens the machine vertically, thus the resonance was no longer a problem.

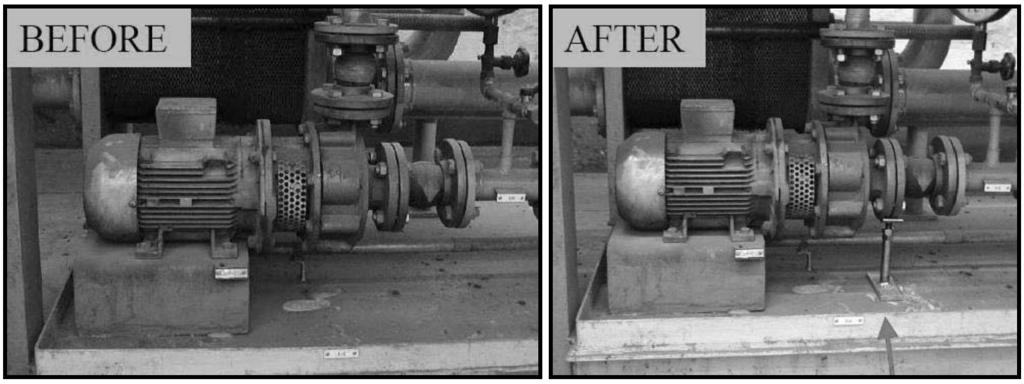


Figure 9-6

The support does not have to be welded in place or bolted to the floor; it can be created so that the height can be set to match the application perfectly.

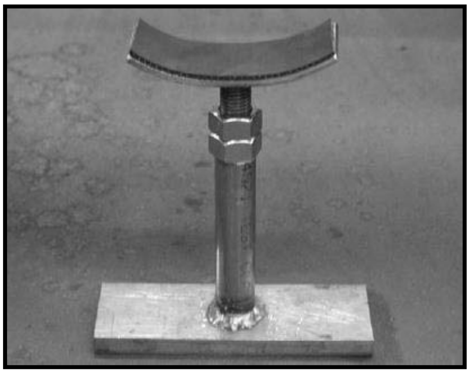


Figure 9-7

Horizontal Flexibility

A great many machines have high vibration in the horizontal axis due to flexibility in that direction. By increasing the stiffness we reduce vibration levels.

In this example, four triangular pieces of steel have been added to the outside of the base-plate, and one inside. The stiffness was increased dramatically, and the vibration levels were reduced.

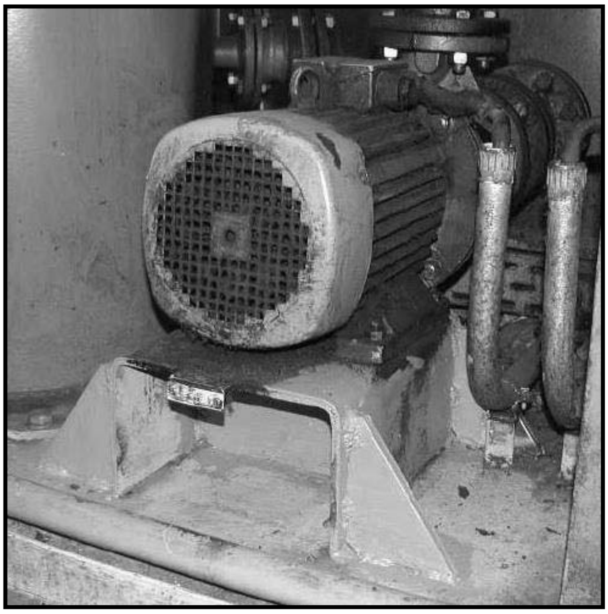


Figure 9-8

As a general comment, there are many overhung pumps with inadequate stiffness on the coupling side; little to no support on the coupling side; and they are mounted on foundations with natural frequencies in their operating speed range.

Base resonance

In the following example, it was believed that the foundations had a resonance in the operating range of the machine. As you can see from the following bump tests data, there was a resonance at approximately 100 Hz (it was moved to 159 Hz). The machine needed to be stiffened to increase the natural frequency away from any of the machine forcing frequencies.

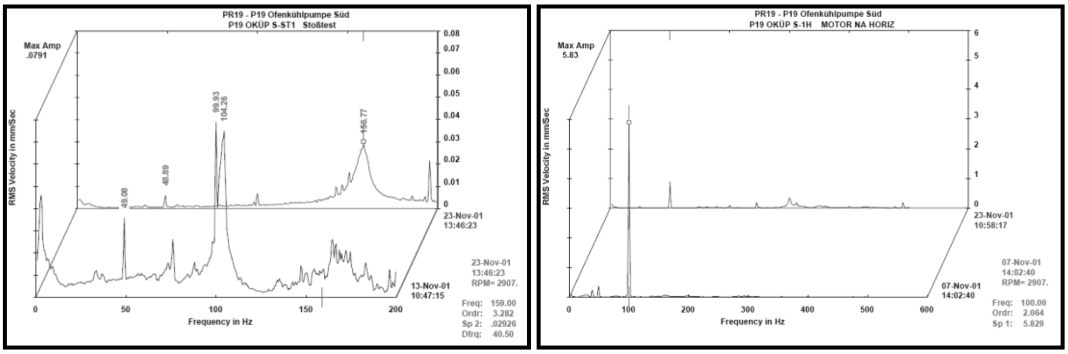


Figure 9-9

The graph on the right above shows before and after spectra. The peak at 100 Hz has dropped considerably.

Three modifications were tried: adding braces to the foundation; adding support to the motor; and the change that had the greatest effect, adding support to the pump.

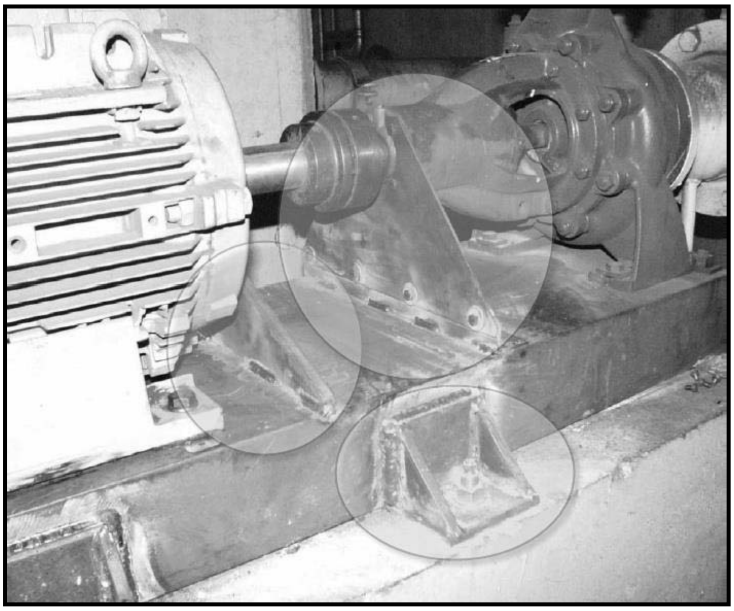


Figure 9-10

Vibration from Sister Machines

Another common situation is where pumps are mounted on the same metal structure – with at least one unit operating as a stand-by. In addition to the resonance problem, you face the risk that the standby machine is constantly experiencing the vibration from the other pump(s). This will damage the bearings (brinelling). It is not uncommon for the standby machine to fail quickly after starting, even though the last time it was run vibration levels were OK.

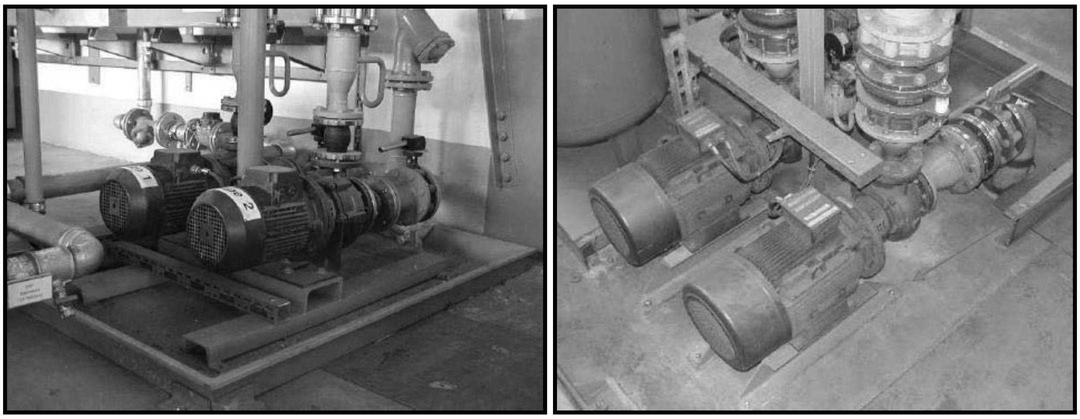


Figure 9-11

The solution is to cut the base so that the two machines are mechanically isolated, or create new, completely isolated bases.

Case Study: Minera Yanacocha- Newmont Gold in Peru

Introduction

Pump problems are common, and troubleshooting them often involves more than just simple spectrum analysis. In this case history, time waveform analysis, phase analysis, pump system knowledge, bearing loading and motor construction knowledge were all necessary to diagnose and fix the problem. Vertical pumps have a unique set of problems and most of the time analysis is done with vibration readings taken on the motor, as the pump is usually inaccessible.

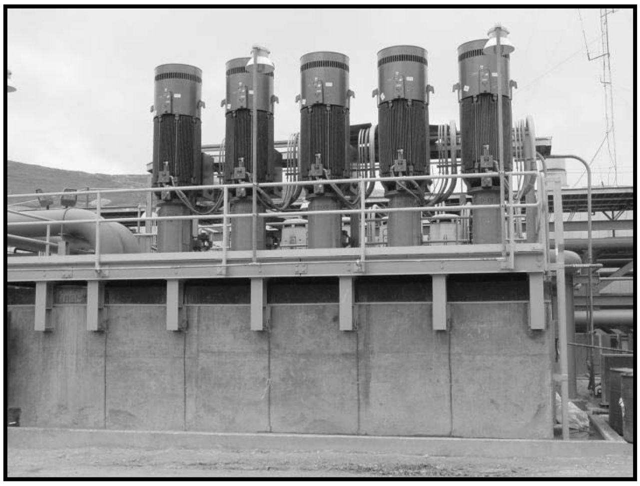


Figure 9-12

The pumps analyzed here were installed in order to raise production of gold at a large gold mine in South America. The mine is situated at an altitude of 15,000 feet! Five pumps were installed and commissioned. The pumps are known as "barren solution pumps", as the gold, silver and other elements have been removed from the solution. The solution is then circulated through a leach pad of crushed rock to become "pregnant" with elements. The mechanics and operators noted "high vibration" but an investigation was not initiated until a motor failed. A consultant was called in and the diagnosis of "resonance at running speed" was the outcome of his testing. No actions were taken to correct this condition. Within three months, two more motors had failed, so another consultant was summoned. He also diagnosed resonance, and recommended a design modification to the pump head.

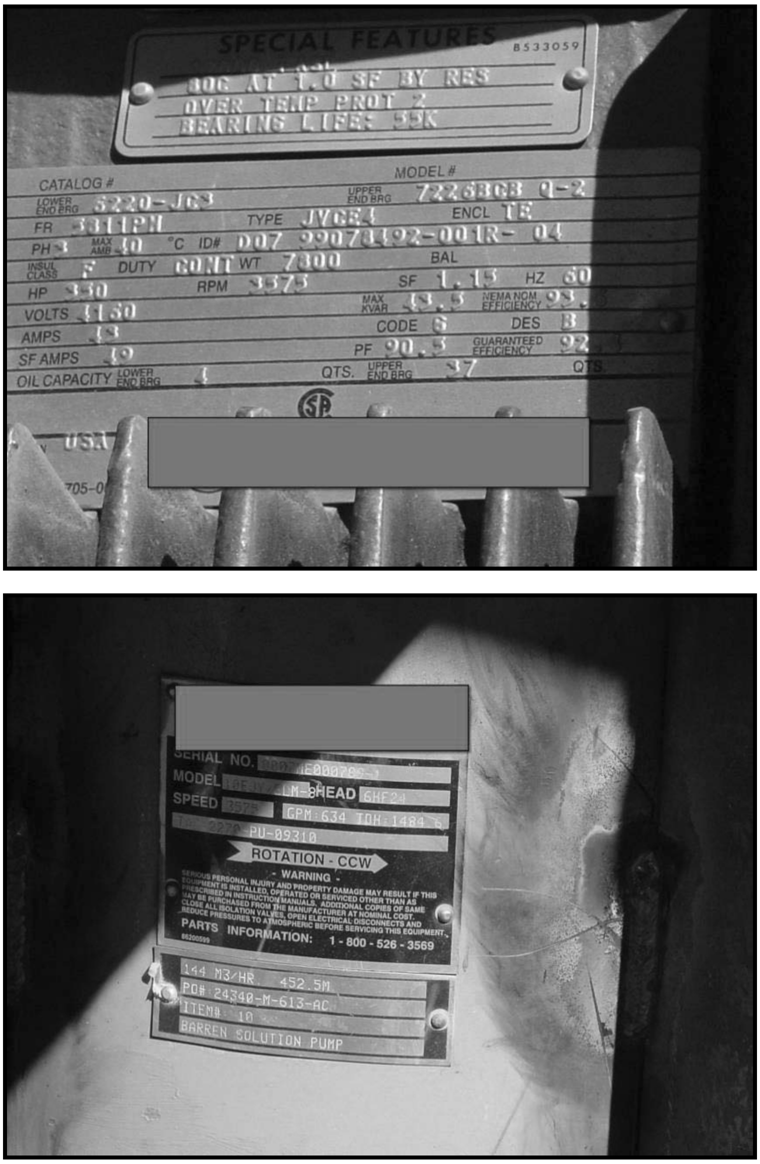


Figure 9-13

A set of vibration data was taken for reference before modifications. The data is shown on #8 pump and #10 pump for all motor points. The highest levels are at 1X turning speed, in line with the discharge head. It is interesting to note that the highest vibration levels are at the bottom of the motor. This is abnormal, and phase analysis is necessary to understand the deflection shape of the pump.

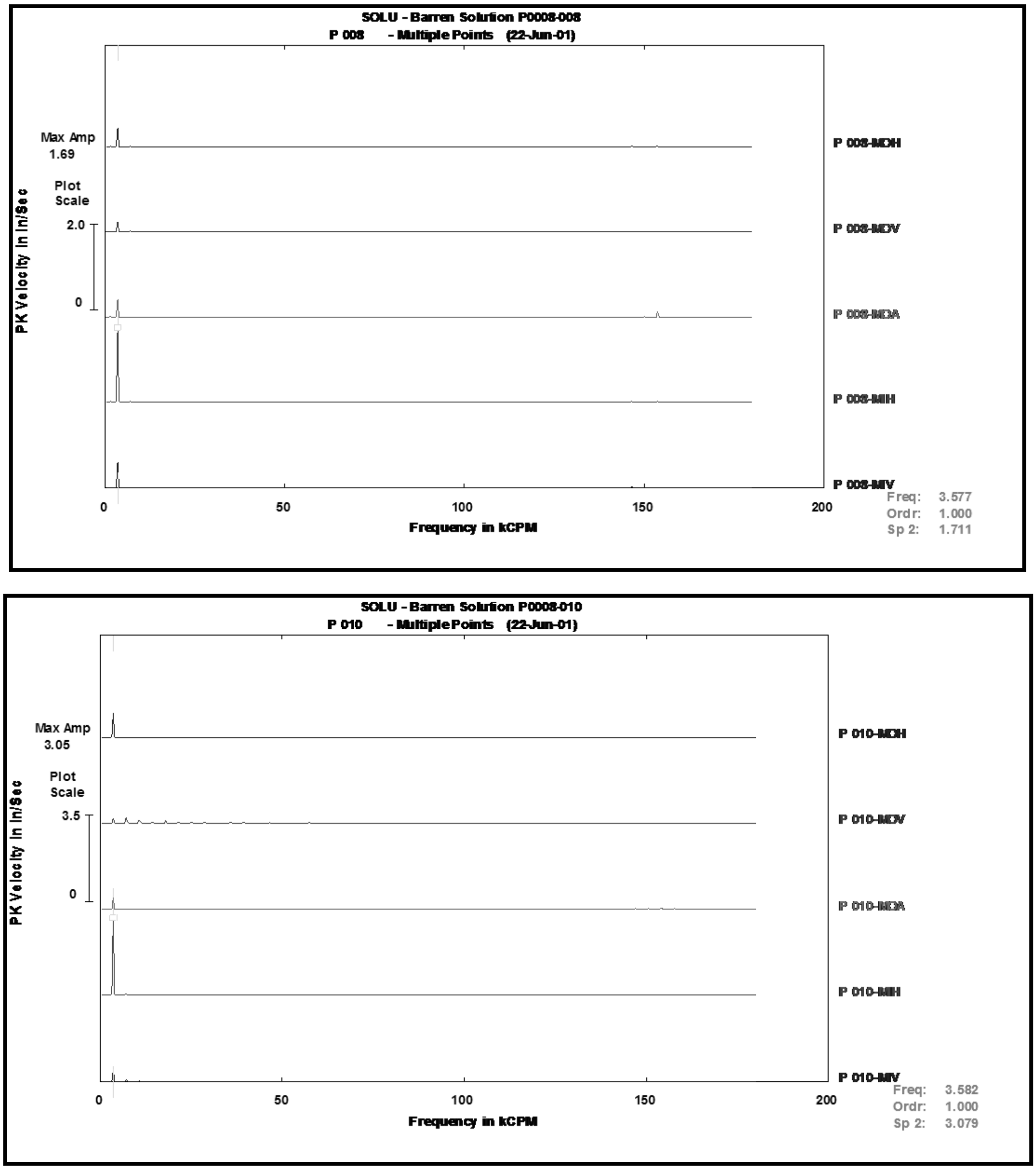


Figure 9-14

The highest reading is shown on #10 pump, MIH position below. A very high 1X turning speed peak at over 3 in/sec pk is noted.

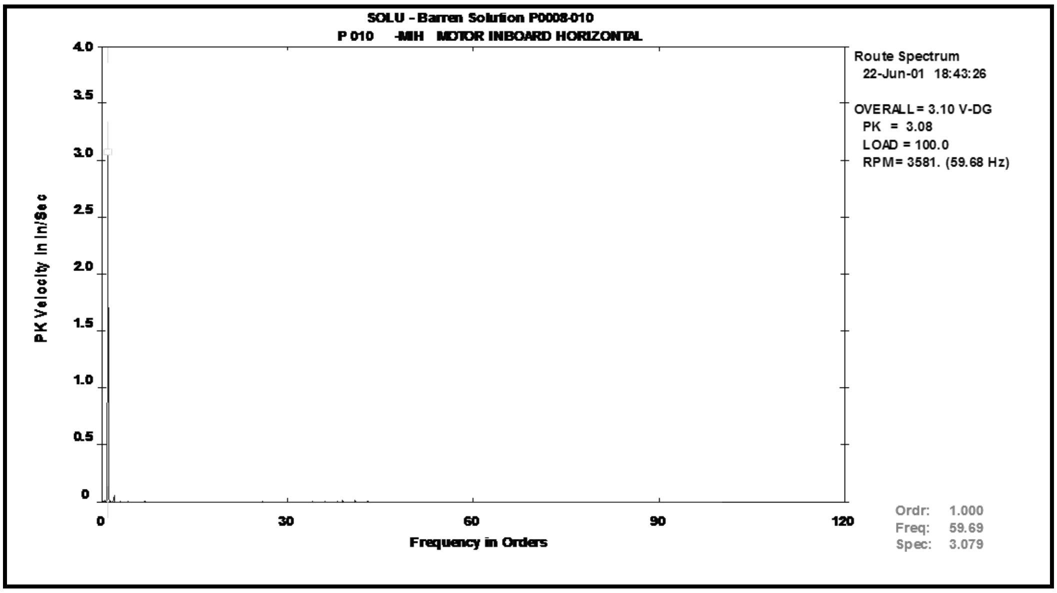


Figure 9-15

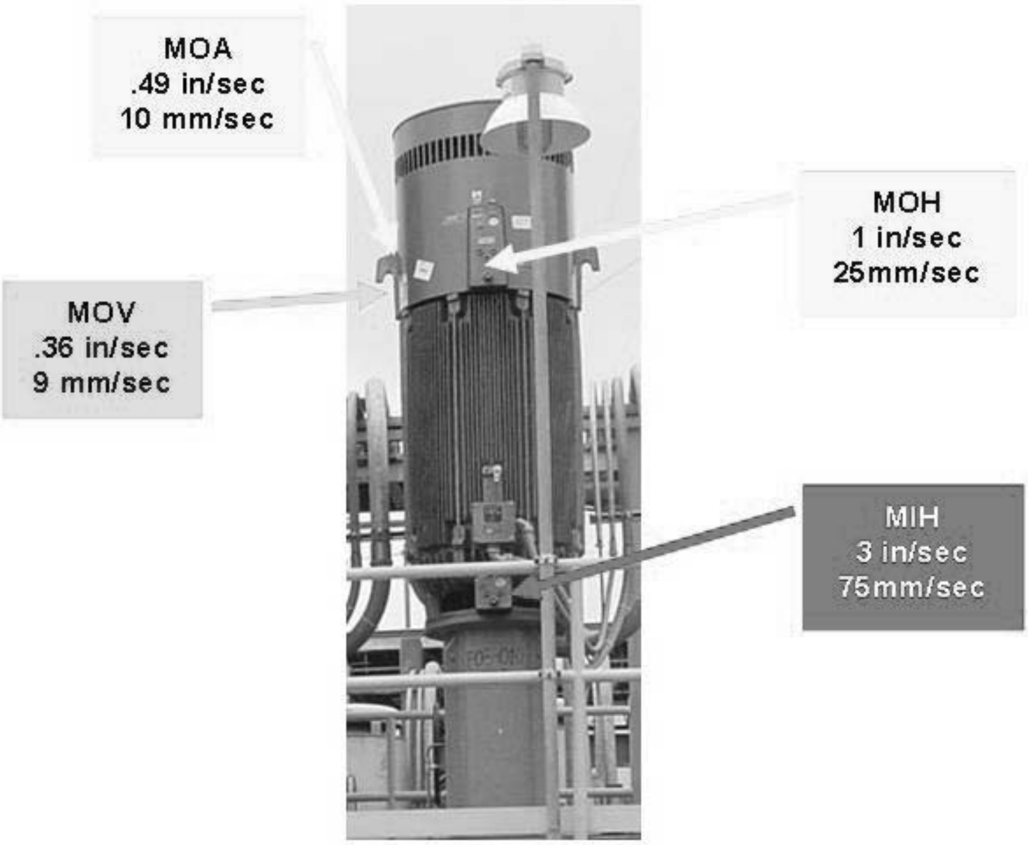


Figure 9-16 - #10 pump measurement locations and amplitudes

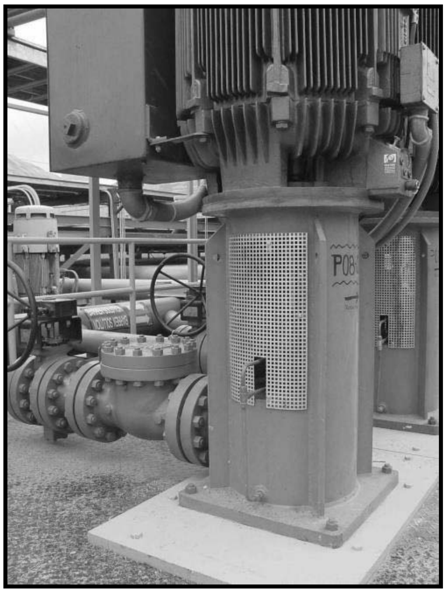


Figure 9-17

A negative averaging bump test was performed in order to better understand the natural frequencies. The results are shown below. Negative averaging is a two step process. First the machine data is recorded along with the impact excitation. In this case the excitation was provided via a rubber mallet (it was actually a modal hammer, but only used as a tool to excite the resonance). The second step involves recording the vibration without the impact excitation. The "valley" to the right of the cursor marker at 3510 CPM is where the running speed vibration was subtracted.

The first narrow peak at approximately 480 CPM is the first mode (cantilevered beam). This mode can be excited by flow turbulence or rubs and impacts. In this case, this mode was of no concern. The 2nd mode is shown as the group of peaks from 2400 CPM to 4400 CPM. The broad group of peaks suggests this mode is heavily damped.

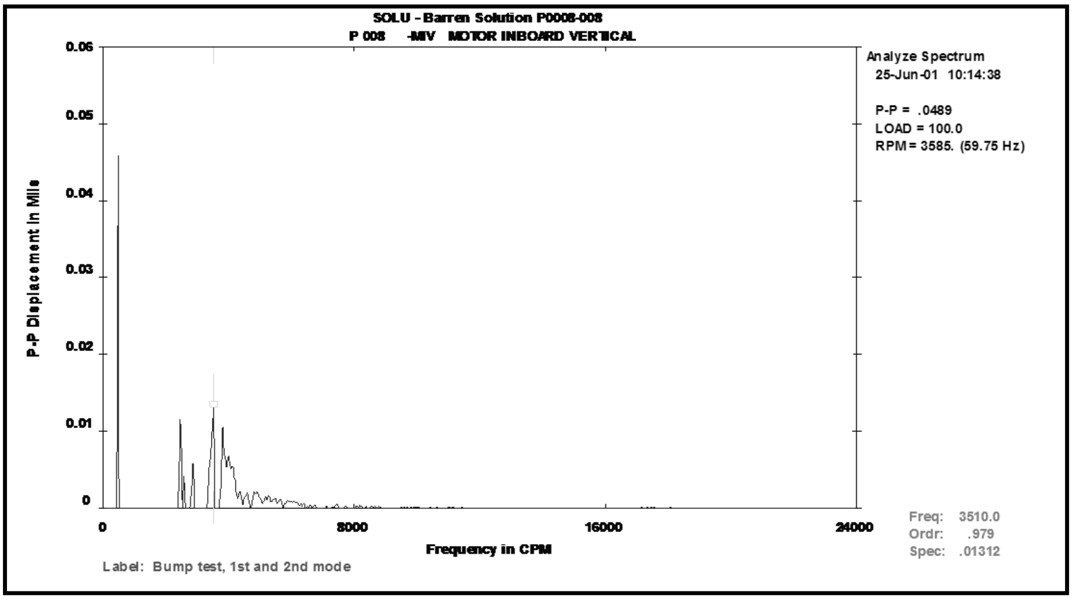


Figure 9-18

Phase readings were also taken from the pump and motor with a two-channel data collector. They supported the idea that a resonance problem may exist. With the top of the motor and pump moving 180° out-of-phase with the point at the bottom of motor, the classical "S" shape 2nd mode may be visualized. A full operating deflection shape analysis test was not performed.

An attempt was made to field balance the motor on the top balance ring. The amplitude and phase vectors did not respond to large trial weights. The coupling was checked for correct installation and machine was checked for misalignment and bent shaft. All were determined to be satisfactory by the pump manufacturer's representative. The base was evaluated for flexing or motion, however the results were considered normal.

Figure 9-19

I arrived at the same time as a representative of the pump manufacturer. He planned to implement design changes that were intended to stiffen the pump head. The photo below shows the pump after the changes were made.

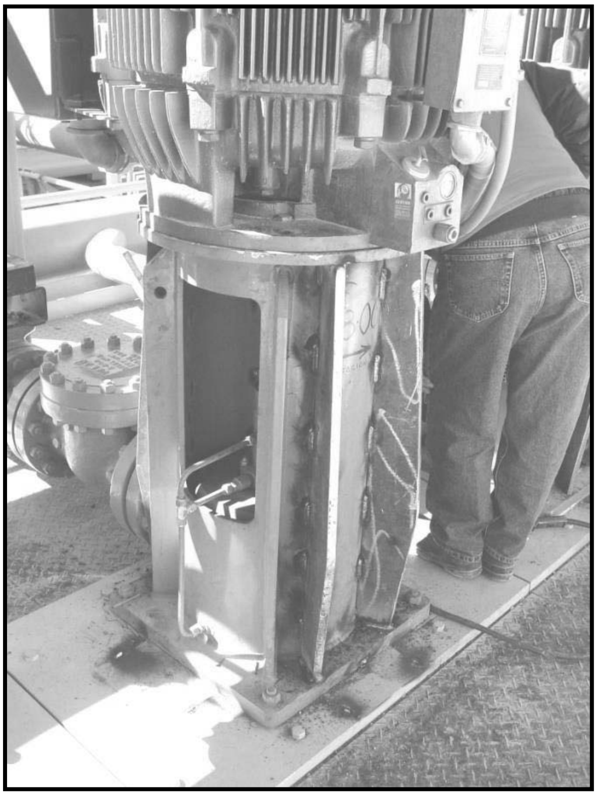


Figure 9-20

The plots below show vibration data taken before and after modifications were made to #8 pump. Although there was greater than 25% reduction in amplitude, the levels were still unacceptable for long term, reliable operation. They stiffeners were not very effective.

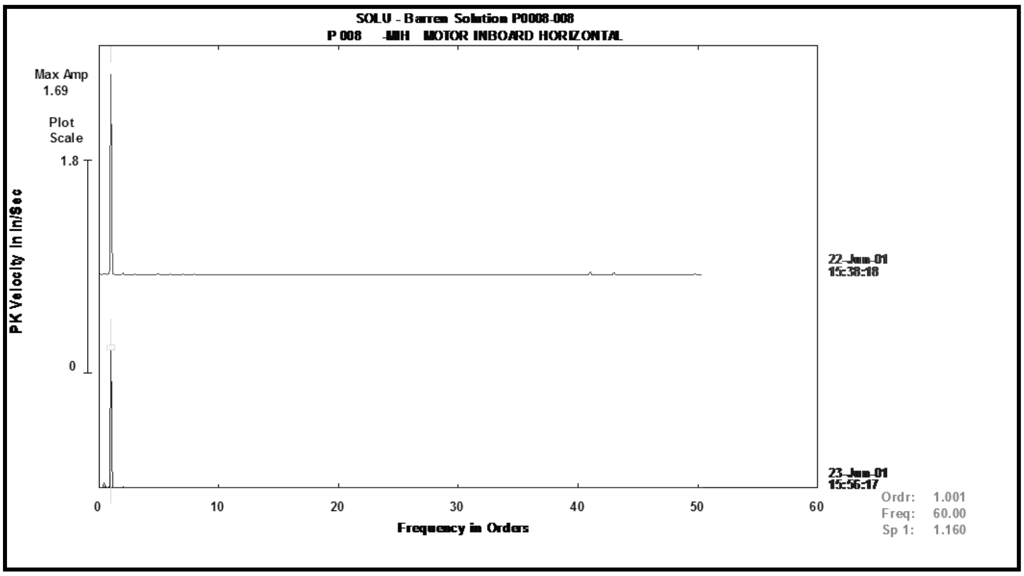


Figure 9-21

The forcing functions and the failure modes needed to be evaluated to solve the problem. The #10 pump motor had a failure during the testing process. The data shown below is a spectrum with 1X turning speed and harmonics. This may be associated with internal looseness or impacting.

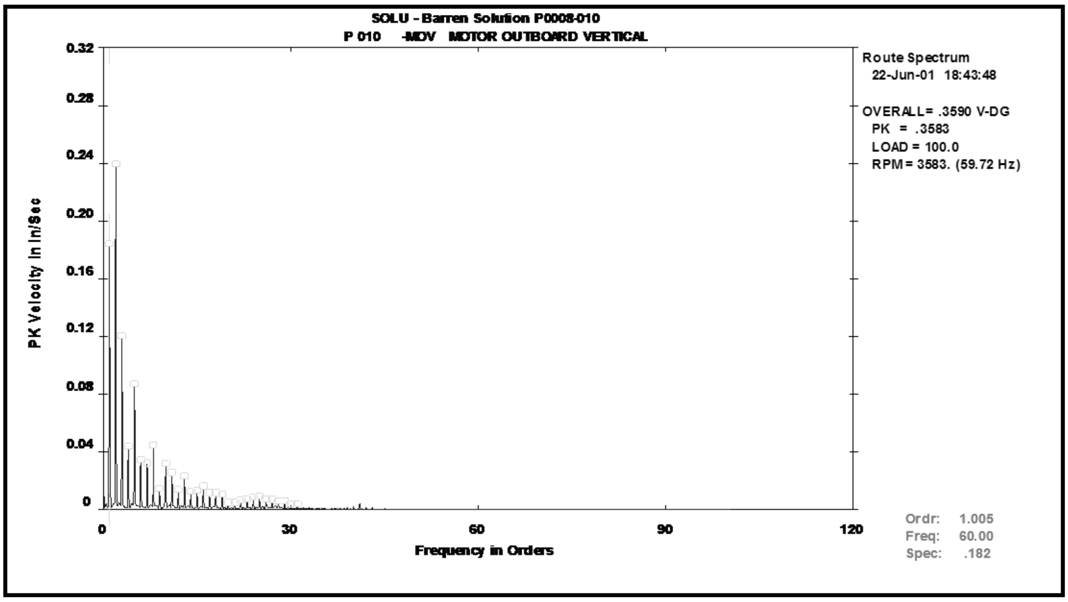


Figure 9-22

The time waveform below is very asymmetrical (+5.5 G-s to -1.92 G-s) and truncated. This clearly shows impacting and rubbing. Note the two large impacts in each revolution of the shaft. They are very repetitive. This motor failed shortly after these readings were taken and was removed to a facility for inspection and repair.

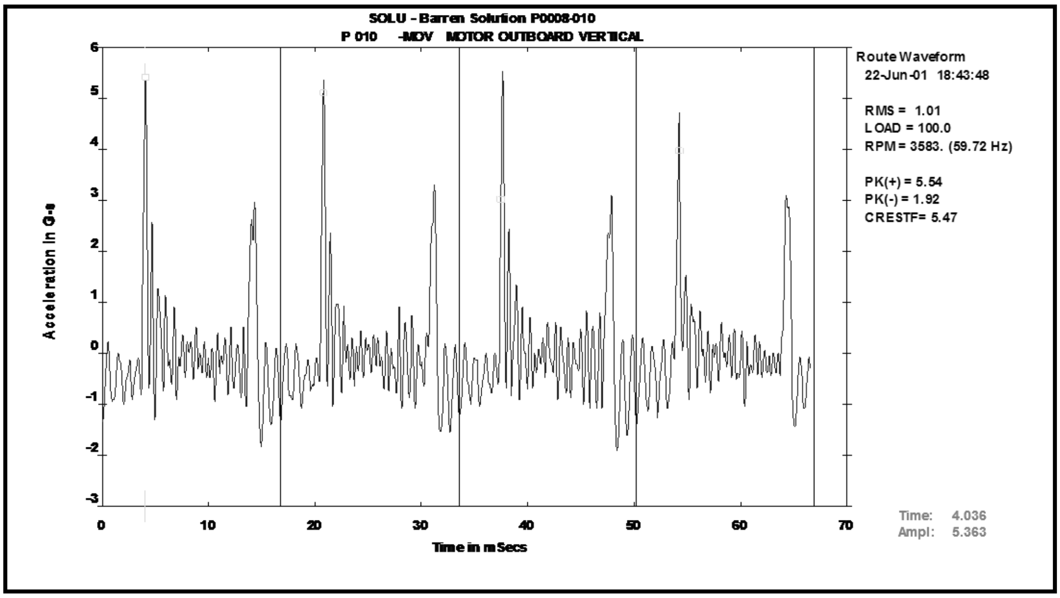
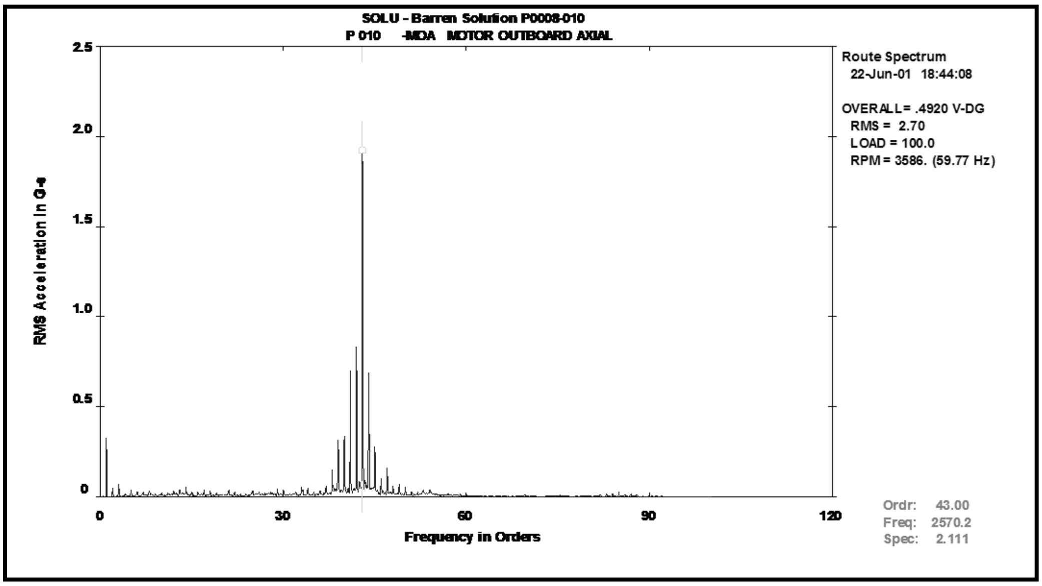


Figure 9-23

Another look at the spectral data in acceleration from the #10 pump motor shows a large synchronous peak in the area of rotor bar pass frequency (RBPF). The sidebands are spaced at 1X turning speed. This is a common indicator of dynamic eccentricity. The rotor may be out of round, or the bearing may be loose enough to allow the rotor to travel in an eccentric path.



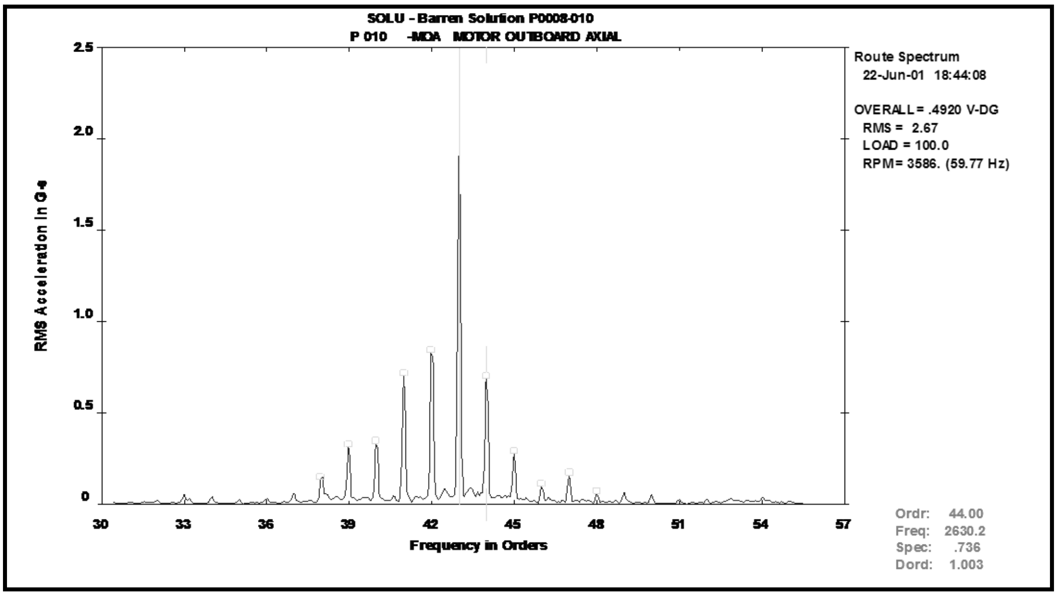


Figure 9-24

An orbit test was also performed. The orbit (derived from filtered acceleration data) reveals the motion of the bearing.

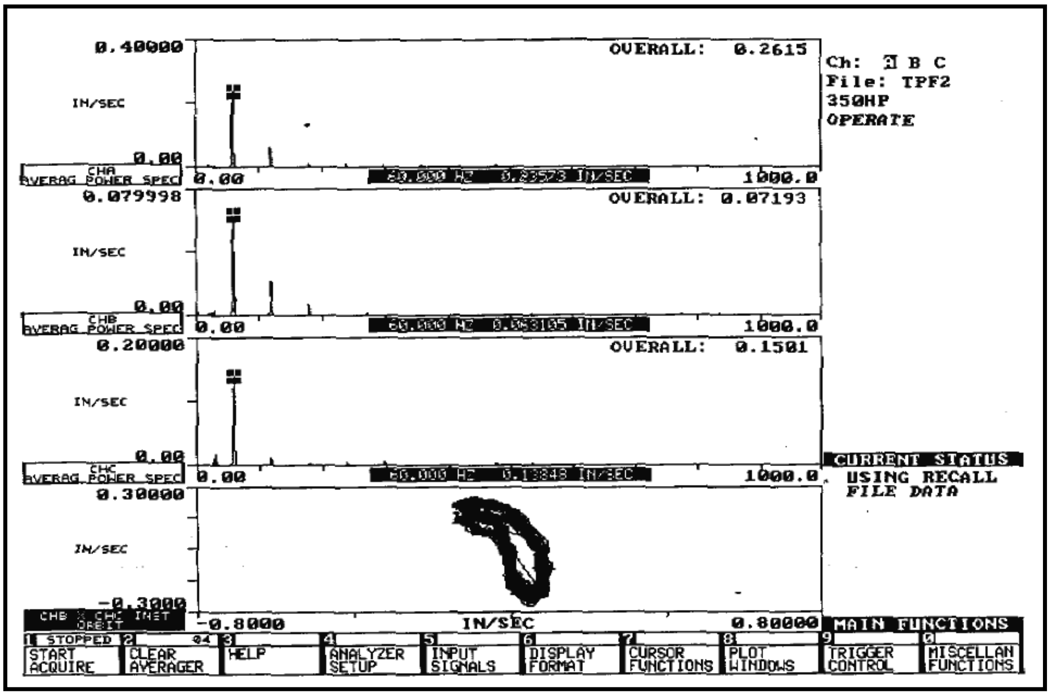


Figure 9-25

The top bearing carrier assembly for the motor is shown below.

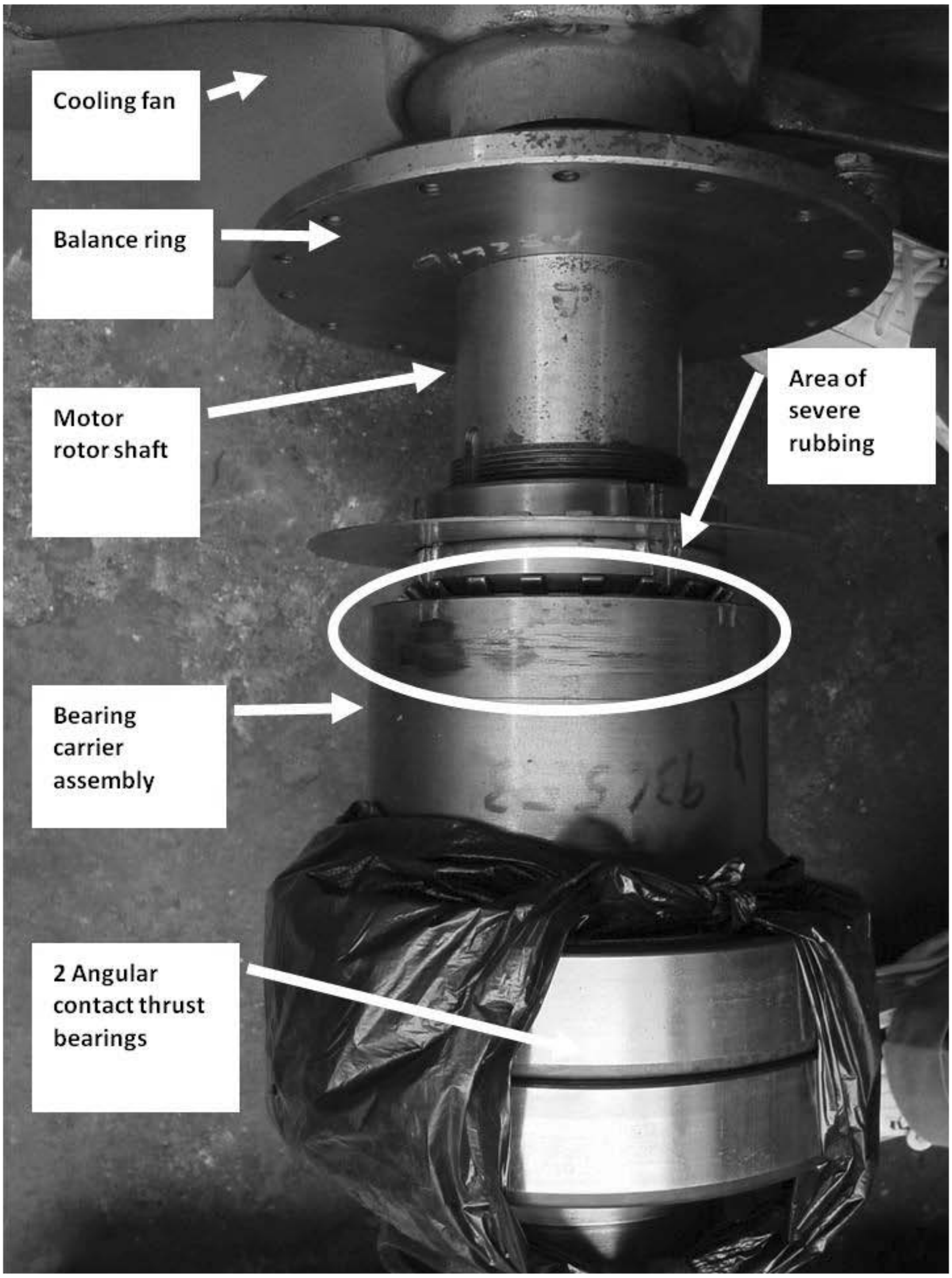


Figure 9-26

From the motor name plate it could be seen that the top bearings were 7226 BCB angular contact bearings. The damaged motor top cover plate is shown below. The seal area is severely galled from the bearing carrier assembly rubbing. This is normally a .010”-.015” clearance area, so the rotor was moving more than design tolerances.

Figure 9-27

It was suspected that the top bearings were not loaded sufficiently. An angular contact bearing must have sufficient axial load so the bearing does not allow excessive radial motion. The SKF Engineering Handbook Catalogue was consulted for minimum axial load requirements. Some rough estimates of static loads from the weight of the rotor and pump assembly were made. The static load was not enough to load the bearing. The pump develops head (pressure) and provides loading to the bearing when in operation.

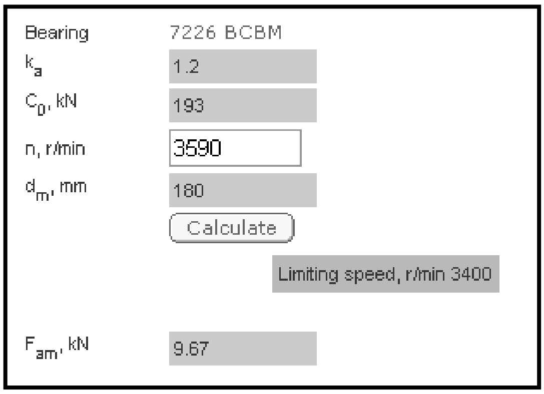


Figure 9-28

A typical pump curve is shown below. Pumps can produce output at points along their curve for a given impeller diameter, speed and horsepower. There is a point of operation known as the “Best Efficiency Point” (BEP). This is the point the pump designer will minimize flow disturbances and most efficiently impart velocity to the fluid. This velocity is converted to head (pressure) in the piping. Operation at points far away from BEP may introduce noise, cavitation, unbalanced dynamic pressures and vibration.

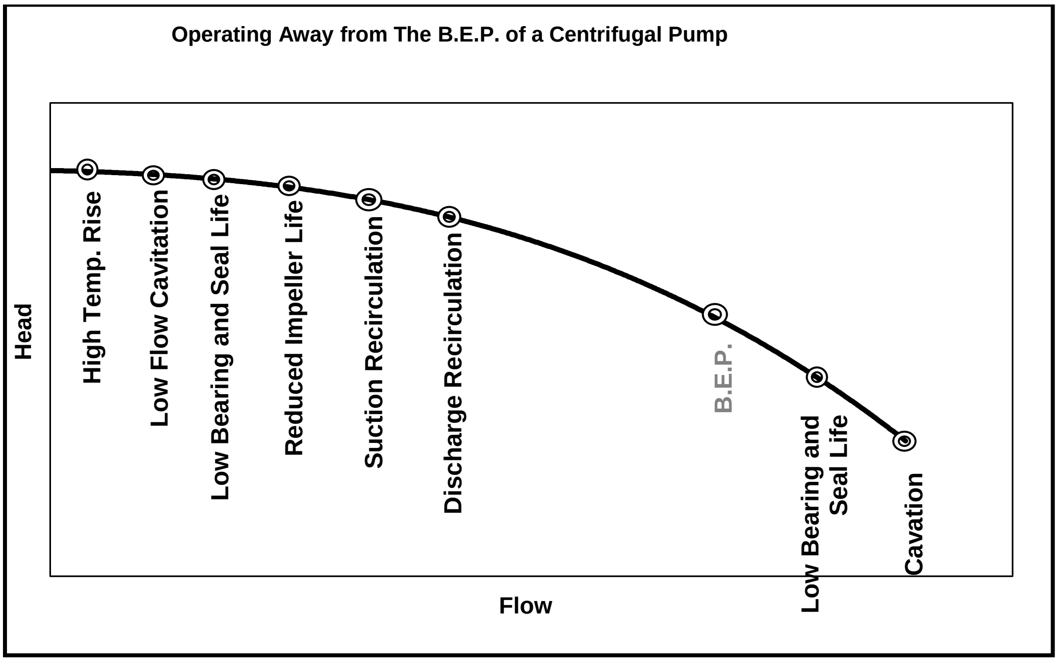


Figure 9-29

A gauge was installed in the discharge pipe to measure flow/pressure. It was determined the pump was operating to the very right side of the pump curve in a high flow, low head condition referred to as “run out”. In this condition the pump was not developing the head which would have loaded the top angular contact bearings sufficiently. The bearings were then able to move in the radial direction, and the rotor was moving in a gyroscope path at the top. The rotor would then impact and rub, exciting the 2nd mode shape.

To prove that too much flow was causing the excessive motion and vibration, the discharge valves were temporarily throttled 30%. The resulting vibration is shown before and after. The vibration has decreased from 1.8 in/sec to .3 in/sec. The pump is operating closer to BEP with less flow and more pressure. The additional head pressure is loading the pump impellers in the axial direction, which loads the angular contact bearings, reducing radial clearance.

Produced by Sumico Technologies

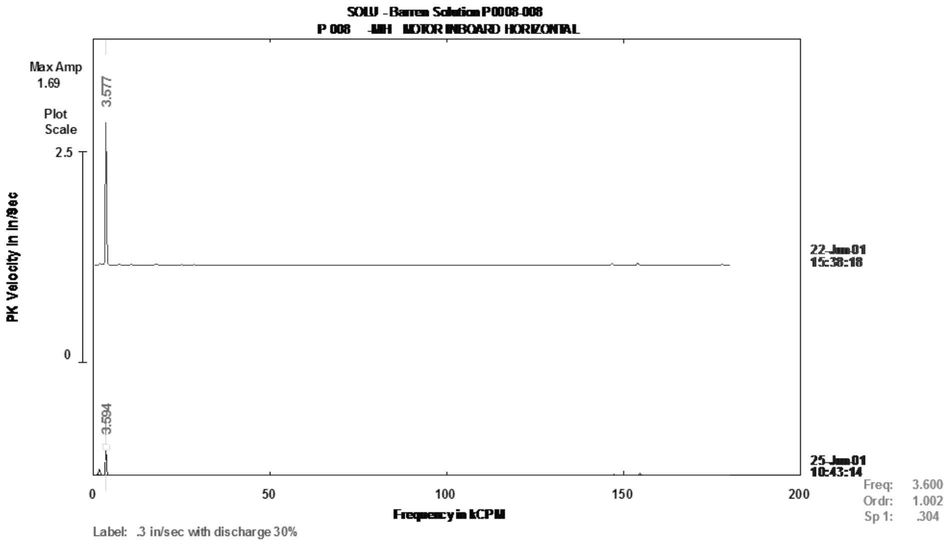


Figure 9-30

The other thing to note above is the running speed has increased by 17 RPM. By throttling the discharge valve, the pump has lower flow in GPM, which puts less load on the motor, so it runs faster.

The root cause was found to be in the system piping. The pump was designed for approximately four miles of piping. Because the mine was relatively new, the leach piles were not as large as they would eventually become, so all the piping was not installed yet.

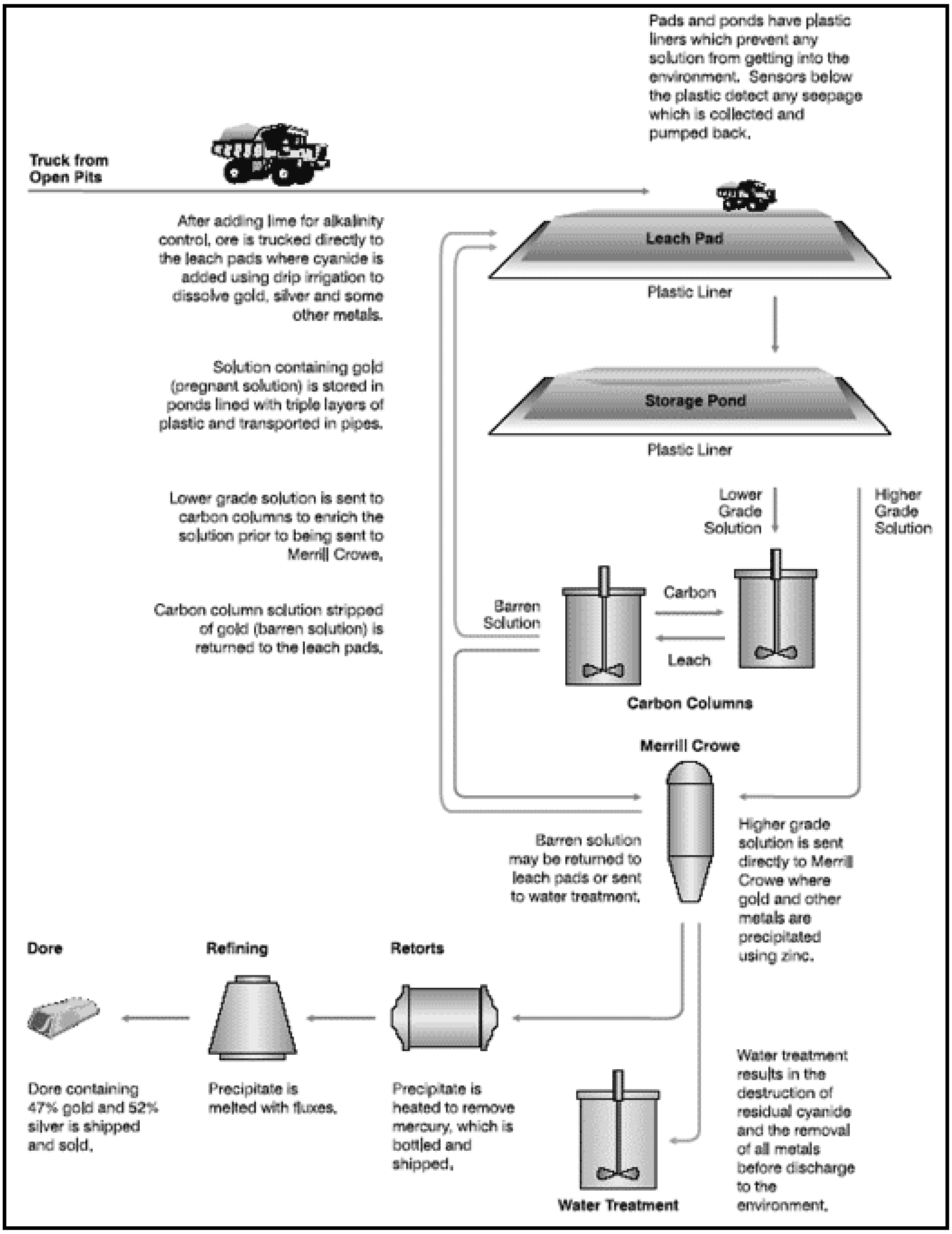


Figure 9-31

The fix was to remove a stage from the multi stage pump, and install a restrictor plate in the discharge piping to decrease flow and increase the head. The angular contact bearing configuration was changed from tandem (same direction) to face-to-face so the bearings would load each other, therefore assuring proper axial loading.

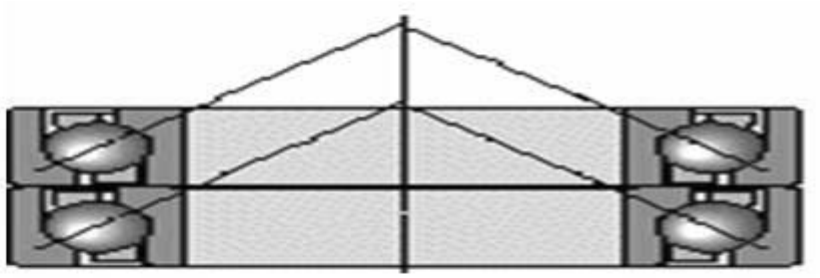


Figure 9-32 - Tandem bearing arrangement

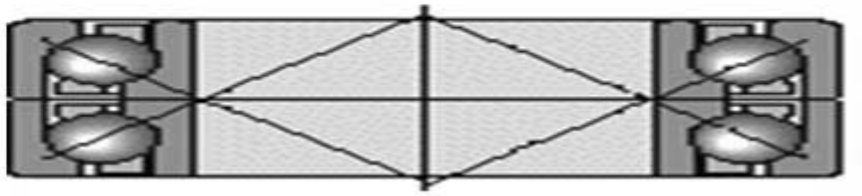


Figure 9-33 - Face-to-face bearing arrangement

The pumps were then evaluated again and the results are shown below.

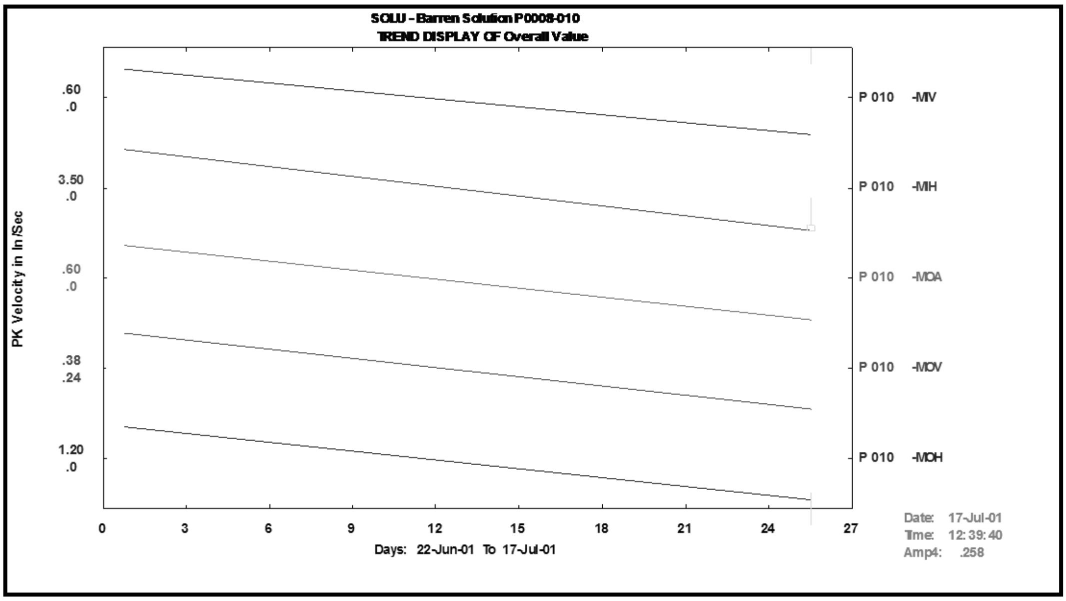


Figure 9-34

The vibration at 1X turning speed decreased from 3 in/sec to .25 in/sec.

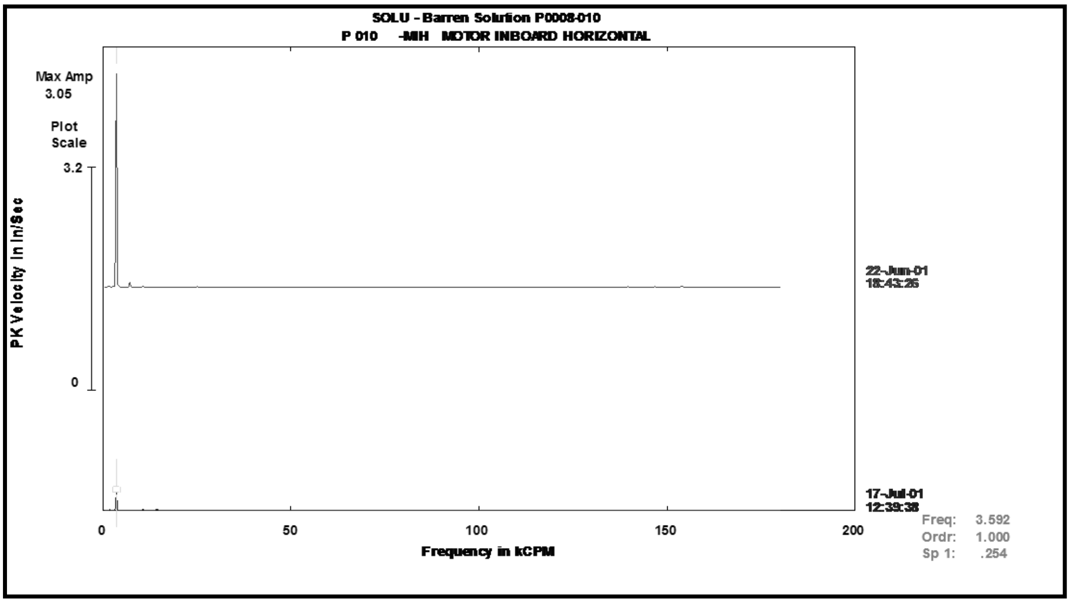


Figure 9-35

This case history shows the importance of pump system knowledge and how it relates to vibration. Careful study of all spectral data, waveform and phase were important, as was knowledge of bearing installation and operation. All performance data (flow, pressure) and pump curves may be needed.

Damping

At a natural frequency, vibration amplitude is governed by the system damping – stiffness and mass do not affect vibration level at the natural frequency. Stiffness (k) and mass (m) are multiplied by displacement and acceleration respectively. Displacement and acceleration are 180 degrees out of phase with each other. At a natural frequency the terms related to mass and stiffness cancel each other out, leaving only damping.

$$f(t) = kx + c_v \dot{x} + m\ddot{x}$$

Transmissibility (T) is a measure of the ability of a system either to amplify or to suppress an input vibration, equal to the ratio of the response amplitude of the system in steady-state forced vibration to the excitation amplitude; the ratio may be in forces, displacements, velocities, or accelerations, a ratio of output to input. Transmissibility (T) shows to what extent the vibration at frequency 'f' will be amplified or attenuated.

ζ (zeta) is the critical damping ratio or 'damping factor'. It is the ratio of the actual damping C_v to the critical damping C_c

f is the actual frequency, f_n is the natural frequency. At resonance, $f=f_n$

$$T = \sqrt{\frac{1+\left[2\zeta\left(\frac{f}{f_n}\right)^2\right]^2}{\left[1-\left(\frac{f}{f_n}\right)^2\right]^2+\left[2\zeta\frac{f}{f_n}\right]^2}} \qquad \zeta = \frac{c_v}{c_c}$$

Related to transmissibility is "Q"; the Quality or Maginification factor. When a system is heavilty damped, it will have a low Q, when lightly damped it will have a high Q.

$$Q = \frac{1}{\sqrt{\left[1-\left(\frac{f}{f_n}\right)^2\right]^2+\left[2\zeta\frac{f}{f_n}\right]^2}}$$

At resonance, f = fn and Q is related to the inverse of damping as can be seen in the equation below

$$Q = \frac{1}{2\zeta}$$

Q is also a measure of the bandwidth of the natural frequency as demonstrated in the equation below and image. In this equation, F_L and F_H correspond to the lower and upper half power points on the transmissibility curve. They are equal to 0.707 x the peak amplitude of the transmissibility curve.

$$Q = \frac{f_n}{f_H - f_L}$$

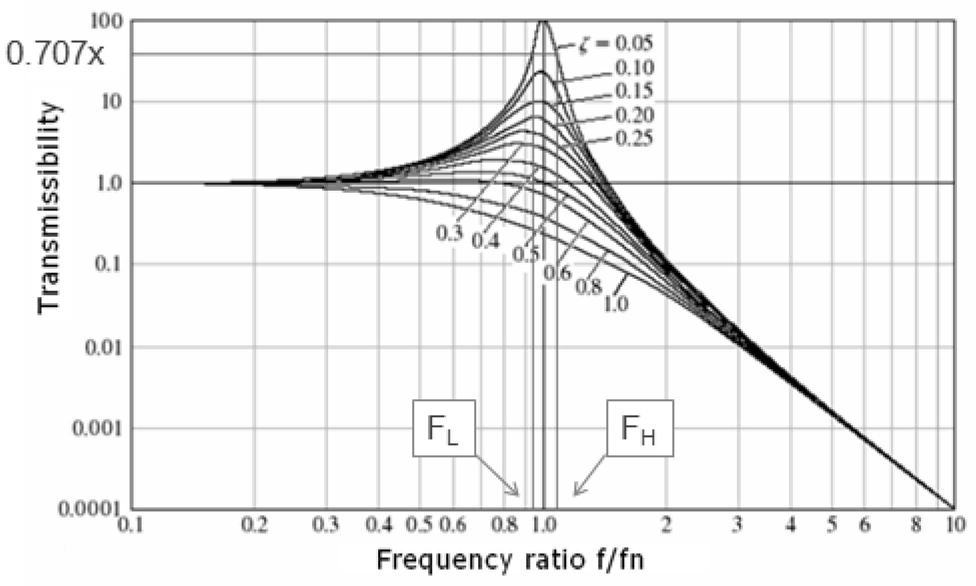


Figure 9-36

Based on the image above one can see that a high Q qould correspond to a sharp peak in the transmissibility curve whereas a low Q would be indicated by a wider rounder peak. It should also be pointed out that when one has measured the transmissibility from a run-up, coast down or modal analysis test, one can then calculate Q from the curve as shown above.

Adding damping

Damping is a property of a material, so if we have a motor or a piping system with a resonance problem and we determine that the only way to resolve the problem is by changing the damping, how does one go about doing that? The answer to this is to add damping materials to the object that we wish to dampen.

Structural damping reduces both impact-generated and steady-state noises at their source. It dissipates vibrational energy in the structure before it can build up and radiate as sound. Damping, however, suppresses only resonant motion.

Forced, nonresonant vibration is rarely attenuated by damping, although application of damping materials sometimes has thateffect because it increases the stiffnessand mass of a system.A damping treatment consists of any material (or combination of materials) applied to a component to increase its ability to dissipate mechanical energy.It is most often useful when applied to a structure that is forced to vibrate at or near its natural (resonant) frequencies, is acted on by forces made up of many frequency components, is subject to impacts or other transient forces, or transmits vibration to noise-radiating surfaces.

Although all materials exhibit a certain amount of damping, many (steel, aluminum, magnesium and glass) have so little internal damping that their resonant behavior makes them effective sound radiators. By bringing structures of these materials into intimate contactwith a highly damped, dynamically stiff material, it is possible to control these resonances.

Of the common damping materials in use, many are viscoelastic - they are capable of storing strain energy when deformed, while dissipating a portion of this energy through hysteresis.

Several types are available in sheet form. Some are adhesive in nature and others are enamel-like for use at high temperatures.

Two most common types:

- Free-layer or extensional damping
- Constrained-layer damping

Free-layer damping

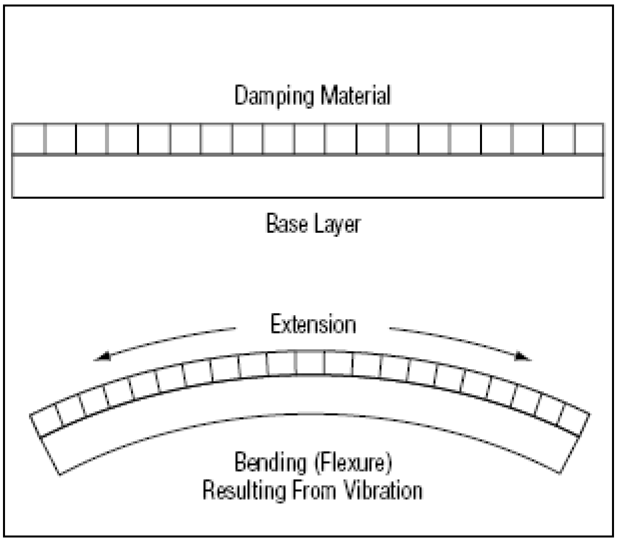


Figure 9-37

Free-layer or extensional damping is one of the simplest forms of material application. The material is simply attached with a strong bonding agent to the surface of a structure. Alternatively, the material may be troweled onto the surface, or the structure may be dipped into a vat of heat-liquefied material that hardens upon cooling. Energy is dissipated as a result of extension and compression of the damping material under flexural stress from the base structure. Damping increases with damping layer thickness. Changing the composition of a damping material may also alter its effectiveness.

Damping materials work to reduce the vibration in the material to which they are applied by dissipating the vibrating energy as heat, rather than radiating it as acoustic energy or noise. Damping materials are termed "viscoelastic," having both elastic and viscous properties. Essentially, the material is stretched when it is bonded to a vibrating surface.

There are two types of damping material, homogeneous or free-layer damping and constrained layer. Homogeneous/free-layer materials are generally vinyls which have platelet-type fillers in them. As the material to which they are applied vibrates, the platelets slide against one another and this friction between platelets converts the vibration energy into heat.

Free-layer damping material, made from stable vinyl and other polymers, work over an extremely wide frequency (50 to 5,000 Hz) and are very stable over a long period of time (10 to 20 years or more)

Constrained-layer damping

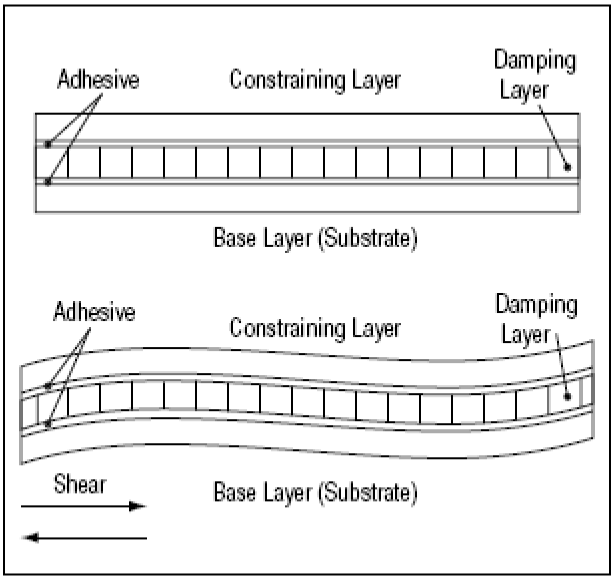


Figure 9-38

Constrained-layer damping (CLD) systems are usually used for very stiff structures. A “sandwich” is formed by laminating the base layer to the damping layer and adding a third constraining layer. When the system flexes during vibration, shear strains develop in the damping layer. Energy is lost through shear deformation, rather than extension, of the material.

The other type of damping is constrained layer. Here, the viscoelastic polymer is homogeneous (not filled) and is sandwiched between two plates. These are bonded together, usually with a structural epoxy adhesive. The ratio of the base thickness to the constraining plate thickness is between 1:1 and 4:1. Better damping is achieved at the 1:1 ratio.

Tuned Absorbers

Another way of dealing with a resonance problem is to utilize a tuned dynamic absorber. Remember that a resonance is occurring at a single natural frequency that can be described in terms of mass, stiffness and damping.

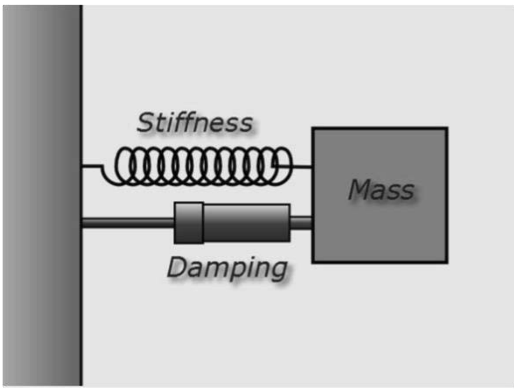


Figure 9-39

The image above describes a single degree of freedom (1-DOF) system with a single natural frequency – in this case, the one that we are concerned with. When constructing a tuned absorber, we are simply adding another degree of freedom to the system, and specifically, another mass spring system with the *same* natural frequency.

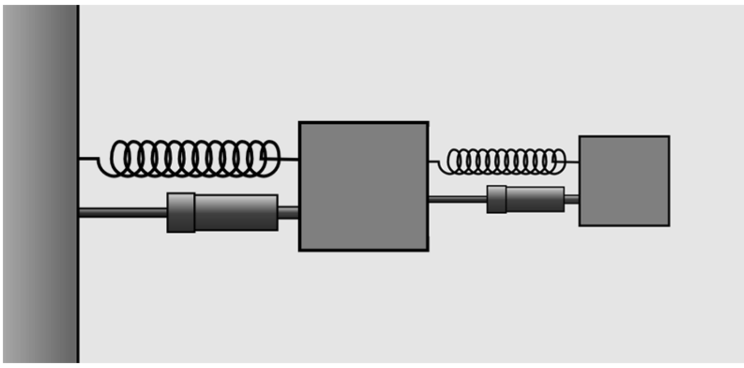


Figure 9-40

The 2 degree of freedom system (2-DOF) that results will have two natural frequencies, neither of which is identical to the natural frequency of the original mass spring system. One of the new natural frequencies will be higher in frequency than the original natural frequency and the other will be lower. This is sometimes referred to as “splitting the mode”

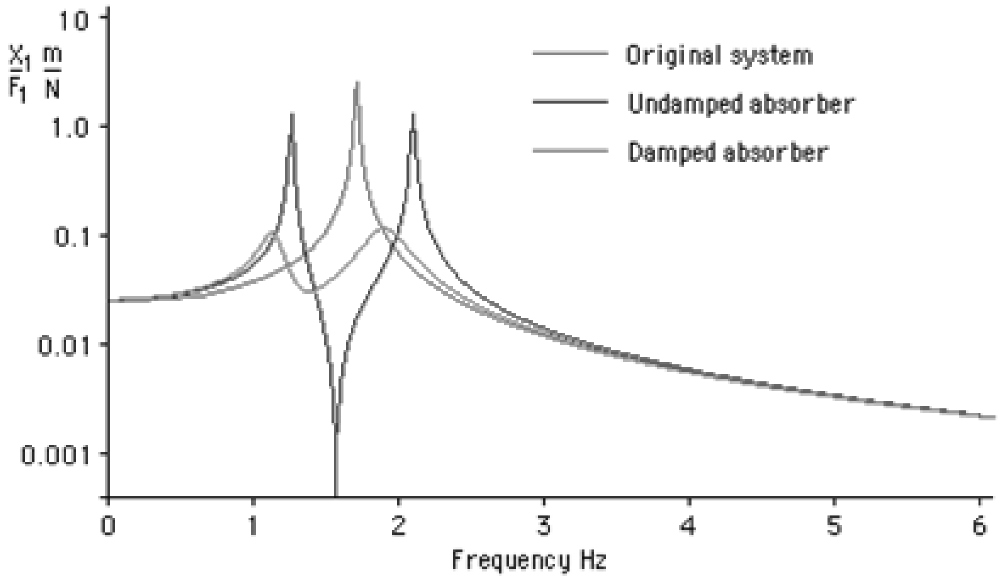


Figure 9-41

This concept is demonstrated in the image above where the orange line describes the natural frequency of the 1-DOF system – i.e. the natural frequency that is in resonance and that we wish to deal with. The blue curve shows the 2-DOF system that results when the absorber is attached to the original system.

The blue curve now shows a node at the point where the original system showed a natural frequency. Therefore, if our forcing frequency does not change then it will now coincide with this node and the vibration levels will be greatly reduced.

One problem with tuned absorbers results from the fact that new 2-DOF system has 2 natural frequencies where the original system only had only one. If the forcing frequency we are concerned with moves up or down in frequency, it may excite one of the two new natural frequencies (on the blue curve) and thus our resonance problem will not have been resolved. Therefore, tuned absorbers are only viable in systems where the forcing frequencies do not change. In the case of a machine this means the machine must run at a fixed speed.

Tuned absorbers may be as simple as a metal rod attached to the vibration object to create the second degree of freedom.

Tuned mass dampers

Another form of tuned absorber is called a tuned mass damper. The concept is the same as in the prior example however in this case the absorber also includes a viscoelastic damping component.

Figure 9-42

As you can see in the figure below, the blue curve resulting from the addition of the tuned mass damper also results in two natural frequencies, but these frequencies are more heavily damped. The damping will reduce the negative effects of the forcing frequency occasional coinciding with the new natural frequencies.

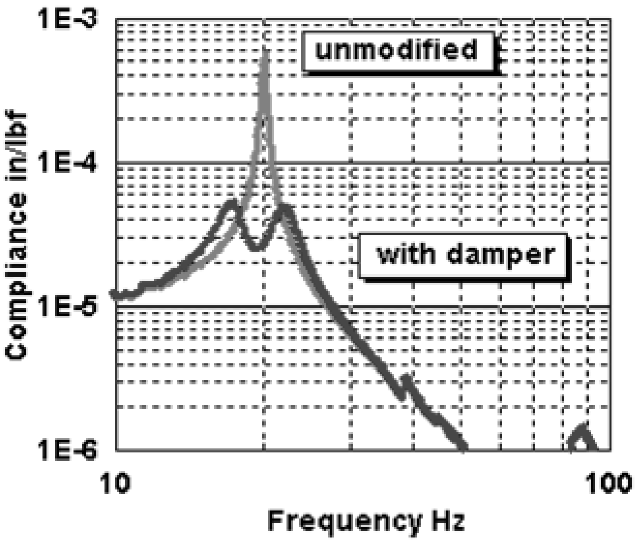


Figure 9-43

Another difference between the tuned mass damper and the tuned absorber that does not involve damping is that at the original natural frequency of the 1-DOF we now do not have a node. The amplification is greatly reduced at the original natural frequency but it is not zero.

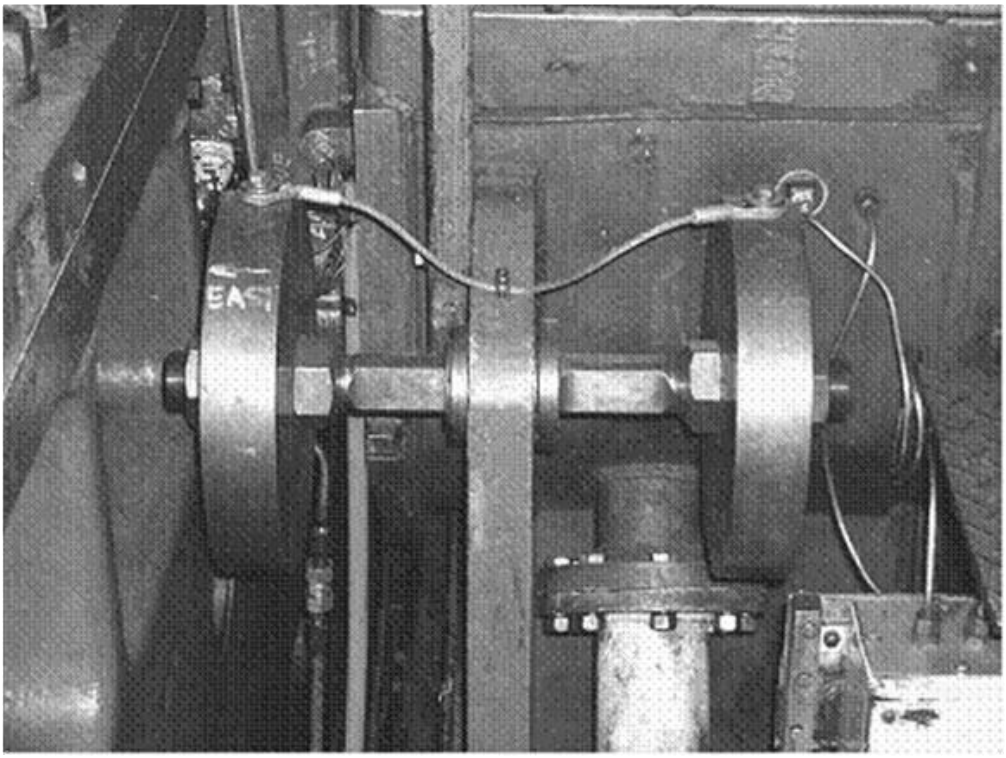


Figure 9-44

The image above is a tuned absorber in a power station in New Zealand. The absorber consists of two 30 kg weights on a 50 mm shaft tuned to 100 Hz.

The following example comes from Arsenal Football Club where the cheering and jumping of the crowds apparently excited a natural frequency of the stands.

Figure 9-45

One solution to the problem was to add stiffness in the form of columns to support the stands; however these would have blocked the view of the fans and were therefore deemed an unacceptable solution. Adding or removing mass to alter the natural frequency was also not acceptable.

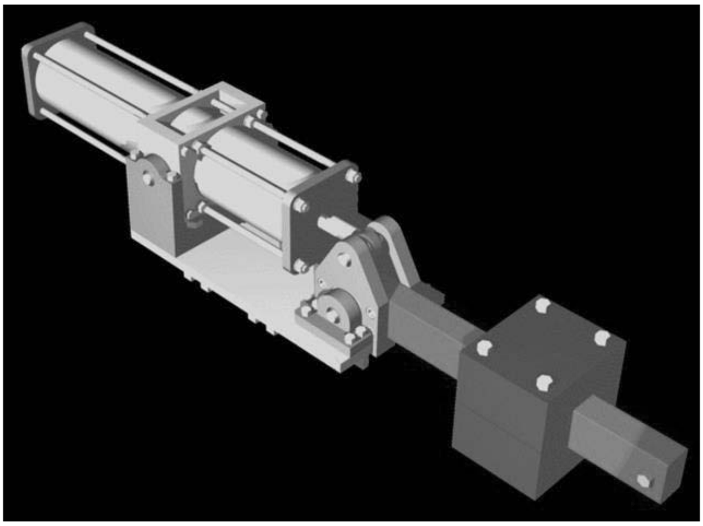
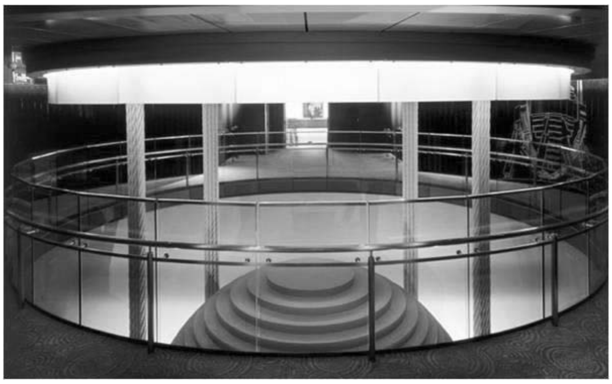


Figure 9-46

A tunable vibration absorber in the form of a rigid arm mounted on a spring and damper in a sealed cylinder. The position of the mass on the arm can be changed to alter its natural frequency. Once linked to the stands, the absorber significantly reduced the levels of vibration caused by the fans.

Another example of a large tuned absorber comes from Taipei 101



The new Taipei 101 holds the title of world's second tallest building (for now) at 1667 ft. Sitting just 660 ft. from a major fault line in Taiwan, the tower could be subjected to earthquakes, typhoons and fierce winds--major challenges to the rigidity of the building. The remedy for these potential seismic and atmospheric assaults is this 730-ton tuned mass damper (TMD). It acts like a giant pendulum to counteract the building's movement--reducing sway due to wind by 30 to 40 percent. Constructed by specialty engineering firm Motioneering, the damper was too heavy to be lifted by crane and had to be assembled on-site. Eight steel cables form a sling to support the ball, while eight viscous dampers act like shock absorbers when the sphere shifts. Able to move 5 ft. in any direction, the Taipei TMD is the world's largest and heaviest. This gold-colored orb is on view from restaurants, bars and observation decks between the 88th and 92nd stories. A bumper ring prevents the ball from swaying too far, should that much swaying ever need to occur.

Isolation

Resonance is a condition that exists when a forcing frequency coincides with a natural frequency of a structure. Another solution to resonance then is to simply not allow the forcing frequency and the structure to come in contact with each other in the first place. This is called vibration isolation and it can be used for many more reasons above and beyond dealing with resonance problems.

Vibration is a root cause of failure in machines and structures. It can cause “false brinelling” of the bearings in stand-by machines or even of bearings in a store room. Vibration can causes problems in a process or result in poor product quality. As an example, consider a machine that engraves miniscule circuits on circuit boards. If this machine or process were subjected to outside vibration it could decrease the accuracy of the engraving.

Vibration can also cause discomfort to people. Imagine going on a cruise for a relaxing vacation only to feel the thumping of the ships engines pounding into your head as you try to sleep or causing your room to shake!

Vibration is also related to noise and thus unwanted vibration is also a cause of unwanted noise.

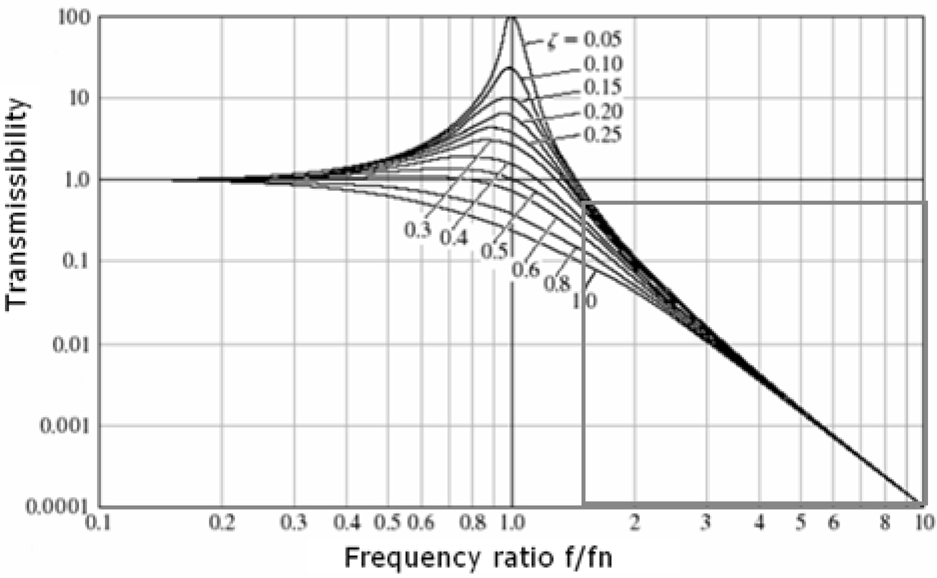


Figure 9-47

One way to employ isolation is to make use of the damping characteristics of a single degree of freedom system like a simple mass and spring. When the frequency f is higher than the natural frequency f_n the input vibration is attenuated. If the amplitude is attenuated enough one could say that it will not even pass through the isolator.

Don't forget that the mass and spring system in the figure above also includes a natural frequency and therefore one must be extremely careful to make sure that this natural frequency is not excited by any forcing frequencies. As an example, perhaps one wishes to prevent the 1x peak from a motor from traveling into the base plate. The isolator would be specified so that its natural frequency is below 1x – but what happens if there is a belt? One must be sure that the belt frequency is also above the natural frequency of the isolator to avoid a resonance problem.

Figure 9-48

When specifying isolators one must also know the weight of each foot of the machine and purchase appropriate isolators based on these weights.

Conclusion

Resonance can be a destructive condition. In order to avoid resonance, forcing frequencies should be kept 15% to 20% away from natural frequencies.

In order to deal with or correct a resonance problem one can attempt to alter the forcing frequency by changing machine speed or the characteristics of particular machine components. For example, one could exchange a 10 bladed fan for a 12 bladed fan in order to alter the blade pass frequency.

A second set of solutions involves changing the natural frequencies of the structure. Natural frequencies can be altered by changing mass and stiffness.

If neither the forcing frequency nor the natural frequency of the structure can be changed, and the resonant condition cannot be avoided, then another option is to add damping to reduce the amplification factor associated with the resonance.

Tuned vibration absorbers are also a possible solution in some cases.

Lastly, one can resolve a resonance problem by simply isolating the forcing frequency from the structure so the two never come in contact.

I hope you have gained an understanding of natural frequencies and resonances. When you look at vibration data in the future (and when you look at your machines), I hope you will be thinking about resonances and consider the possibility that resonances may be responsible for high vibration levels and machine/structural failure. If you do suspect a resonance, you now have a few simple methods for confirming the existence of the resonance, and to better understand how the machine is moving. And finally, I hope you can see that you do not need to be a professor or a civil engineer to come up with practical solutions to resonance problems.

A number of case studies were provided by Juergen Twrdek from the Wieland Werke AG plants in Ulm and Vöhringen, Germany

Thank you to Tony DeMatteo, Mobius Institute instructor, formerly of Emerson Process Management / CSI Division who wrote the paper "Operational Deflection Shape and Modal Analysis Testing To Solve Resonance Problems". That paper, and the associated data, can be downloaded from the Vibrant Technology Web site: http://www.vibetech.com/CaseStudies.htm

Chapter 10

Rolling Element Bearing Analysis

Objectives:

- Diagnose Inner Race defects in spectral data
- Describe the typical waveform characteristics associated with Inner Race defects
- Describe the spectral characteristics of the four bearing components
- Approximate the Inner and Outer Race defect frequencies knowing the number of balls.
- Describe the failure progression of typical rolling element bearings.

Rolling Element Bearings

In this chapter we are going to view the many vibration analysis techniques that can be applied to diagnosing rolling element bearing faults, and we will discuss the four stages of bearing wear. This chapter will cover spectrum analysis, time waveform analysis, enveloping (and demodulation) techniques, and additional high frequency analysis techniques: Shock Pulse, Spike Energy, PeakVue, and acoustic emission.

After completing this section you should understand how to apply these techniques so that you can confidently detect bearing faults from the earliest stage, and confidently tracks its condition through the four stages of wear.

Reliability

In this chapter we will learn how to detect faults and monitor their condition; however it is important to first review what should be done in order to extend the life of the bearing. As a vibration analyst you may consider that it is your job simply to detect the fault. That should be only part of the job description. If bearings were correctly specified, transported, stored, installed, lubricated and operated, then they would provide reliable service for many years. Sadly, almost all of the aforementioned factors are not taken care of – as a result bearings only reach 10% of their design life time.

When bearings fail we may experience downtime, secondary damage, additional labor costs (overtime), and even injury to staff. We wish to avoid these unnecessary expenses.

But the question is, what can you do about the problem? What are you willing to do about the problem?

There are basically two paths we can take – and we should take both paths. We can do what we can to ensure that we get the longest life out of the bearing, and we can monitor the bearing so that we can take the most appropriate action before it fails.

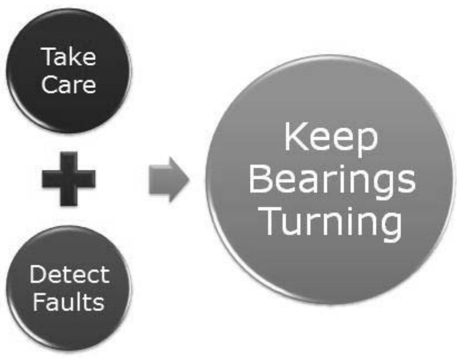


Figure 10-1

Poor design, poor installation practices, and poor maintenance leads to reduced bearing life.

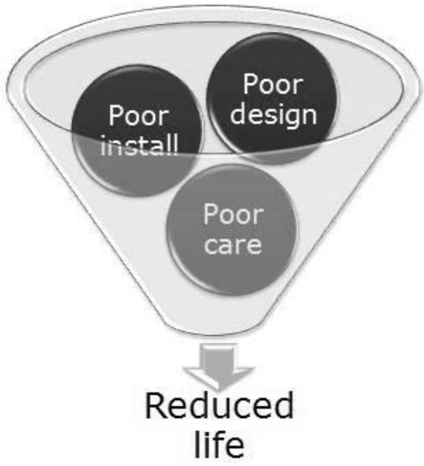


Figure 10-2

As stated earlier, and as generally agreed in the industry, less than 10% of bearings reach their design lifetime. The following graphic highlights the main reason why this is true.

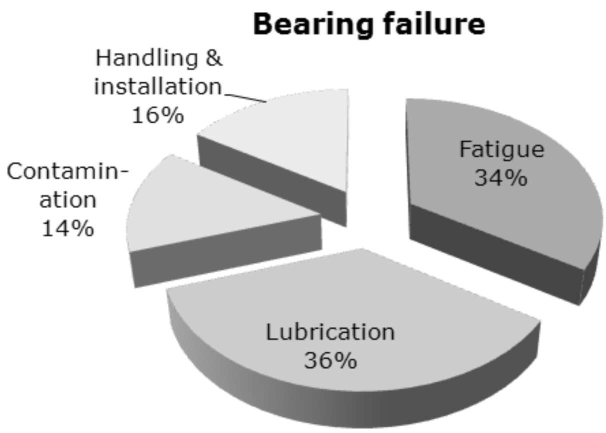


Figure 10-3

Fatigue: 34%

Normal or Pure Fatigue

- With ideal design, installation, lubrication, and contamination, a bearing can be selected for infinite fatigue life

Parasitic Loads

- These loads, above and beyond design loads, reduce the fatigue life of a bearing

Added loads decrease bearing life by a CUBED function (approximation, varies with bearing type)
A 10% load increase from misalignment reduces the calculated bearing life by one-third!

Lubrication: 36%

Is it really a lubrication failure?
Separate the lubricant type from the lubrication process
- Too Much
- Too Little
- Wrong Type

Evidence of Improper Diagnosis and Analysis
- Changing lubricant type to solve a lubrication problem
- Failure to understand the application and environment

Contamination: 14%

Bearing life theories are based on a contaminant-free bearing and lubricant
Sources of Contamination
Seal types
Poor Maintenance Practice
- Pressure washing
- Installation techniques

Often, the better the seal, the more sensitive to fits, tolerances, and shaft deflections
What are the true root causes of contamination?

Poor Handling & Installation: 16%

Perhaps the most avoidable of all bearing failure sources
- Fits and tolerances are critical
- Proper Tools and Skills are Essential

Storage, Transportation
- Before and after installation

Failure mechanisms are created before the shaft makes its first revolution

The Proactive Approach

- Correct bearing for the application
- Proper bearing installation techniques
- Proactive Skills for balance, assembly, alignment
- Lubrication regime
- Storage, shipping, handling
- Proper operation
- Training, Training, Training
- Time to do all of these jobs right

It is clear that if the bearings are properly transported, stored, specified, installed, and operated, and if they are correctly lubricated, then they would have the best chance at remaining in operation for the longest time possible.

Bearing Life

Bearing life is described as the L_{10} "Life Factor". The L_{10} life is the life expected due to normal wear by 90% of bearings. If one had a family of bearings operating under optimal operating conditions, the L_{10} life is the time it takes for 10% of the bearings to fail (90% still operating).

In this document we will discuss the four stages of bearing wear and we will quote the remaining L_{10} life available.

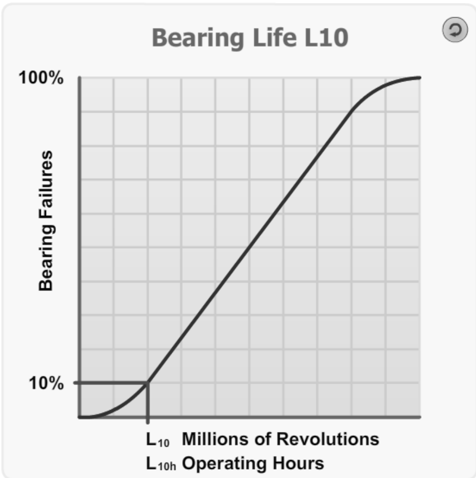


Figure 10-4

These equations show the relationship between the life of the bearing and the speed of operation and the load experienced by the bearing.

This equation shows that for a ball bearing, the life is inversely proportional to the speed of operation, and inversely proportional to the load to the power of three.

$$L_{10} \approx \frac{1}{\Delta RPM} \times \left[\frac{1}{\Delta LOAD}\right]^3$$

Interpreting this information reveals:

1. If we double the speed of operation, the life of the bearing is halved.
2. If we increase the load on the bearing by just 20%, the life of the bearing is halved.
3. If we double the load on the bearing, the life of the bearing is reduced to one-seventh of its design life time.

For bearings with roller bearings, the equation is slightly different, but the fact is still the same; speed and load reduces the life of bearings.

$$L_{10} \approx \frac{1}{\Delta RPM} \times \left[\frac{1}{\Delta LOAD}\right]^{3.33}$$

Condition monitoring

Condition monitoring programs have the potential to detect incipient bearing faults and to reveal the root causes of the bearing failures. Sadly, many programs do not take this extra step. Many programs can be described as follows: they collect data, analyze the vibration, then detect the faults (hopefully) and report their condition.

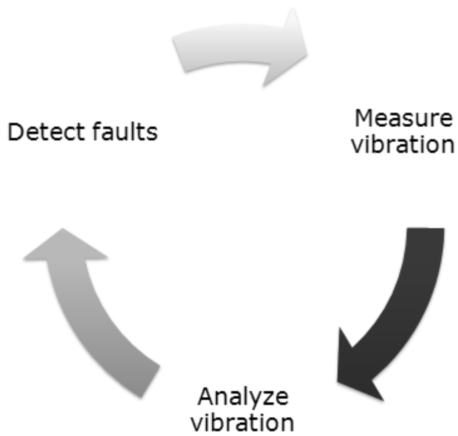
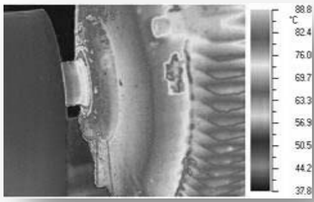


Figure 10-5

The focus is on detecting the bearing fault before it fails, recommending their replacement, and getting back to the measurement/analysis task. In many programs, conditions such as unbalance, misalignment, and resonance are not treated as seriously (often action is not taken when the conditions are reported). As vibration analysts, we can do more. We can employ additional technologies so that we have a complete picture of the health of the machine:

- Vibration analysis and acoustic emission
- Oil analysis and wear particle analysis
- Infrared analysis



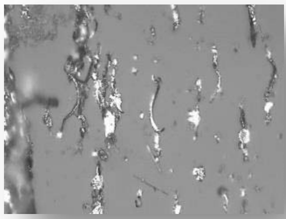


Figure 10-6

The Ideal Condition Monitoring Program

The ideal program would be well designed. Machines would be tested before they were accepted into stores after repair or purchase. Machines would be precision balanced and aligned. Resonance would be brought under control. Machines would be properly lubricated. When you diagnose faults and make your recommendation, not only would you provide a clear, actionable report, but you would verify your diagnosis and perform root cause failure analysis so that the fault would be less likely to occur again.

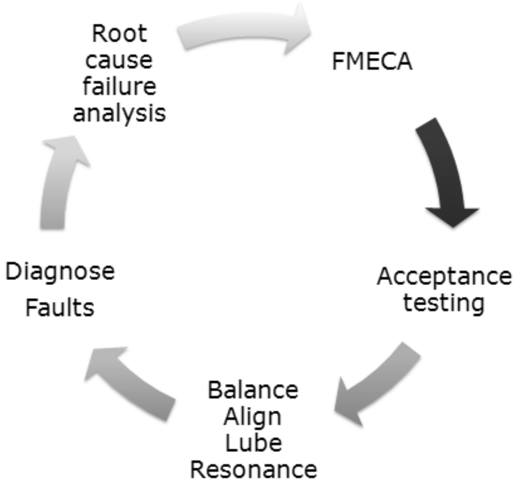


Figure 10-7

All of these steps are taken in the "world class" condition monitoring programs. The focus is on reliability; not just bearing fault detection.

Bearing fault conditions

This chapter is focused on bearing wear or fatigue; the process of gradual degradation of the bearing over time involving spalling, cracking and loss of metal. However there are additional fault conditions related to rolling element bearings that we should be aware of.

First we must be concerned with lubrication. Too much lubricant or too little lubricant will shorten the life of the bearings. There are ways to detect these issues, and it will be discussed shortly.

Second, we must be concerned with installation and fit errors. These topics are covered in greater detail elsewhere, however the following is a summary:

Cocked bearings

- The bearing may be cocked on the inner race and or the outer race. Spectrum analysis, backup up by phase analysis will allow you to detect these conditions.

Bearing loose on the shaft or loose in the housing

- Spectrum analysis and time waveform analysis can help you to detect excessive looseness, and you can attempt to diagnose of the bearing is turning in the housing, or turning on the shaft.

Third, we can detect a special condition experienced on DC motors and AC motors that are driven by variable frequency controllers; fluting. Fluting occurs when current flows through the bearing and causes arcing to occur between the rolling elements and the raceways. It will leave a characteristic pattern on the inner and outer race of the bearing. We will discuss this in the next section.

Bearings and lubrication

The lubricant between the rolling elements and the raceway is under tremendous pressure, especially in the load zone. In fact, the lubricant acts more like a solid than a liquid at this interface. The pressure is so great that there is deformation in both the rolling element and the metal of the inner and outer race. It is this constant, repeating pressure that ultimately causes the metal to fail.

Let's focus for a moment on the interface between the rolling element and the race.

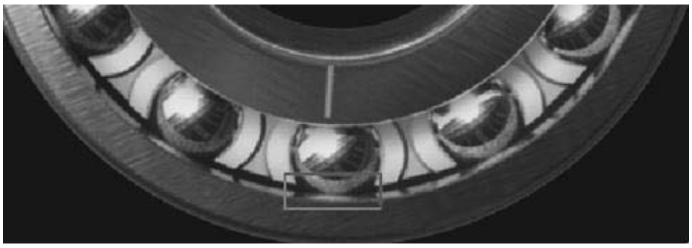


Figure 10-8

If we were to look at the surface of the bearing and race under a powerful microscope, we would see that the bearing is not perfectly smooth.

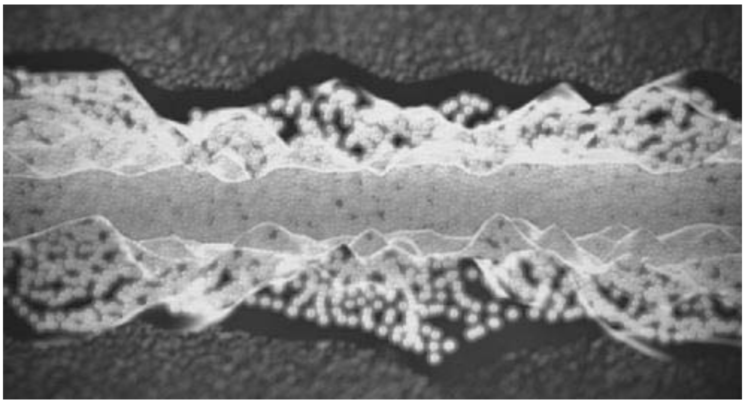


Figure 10-9

There are high-points and low-points on each surface. The lubricant's job is to keep those surfaces apart. But we are talking about very close tolerances. The gap between the two surfaces is very small: 0.5 µ m or 20 µ inches. The diameter of a human hair is 40 times as large as the gap!

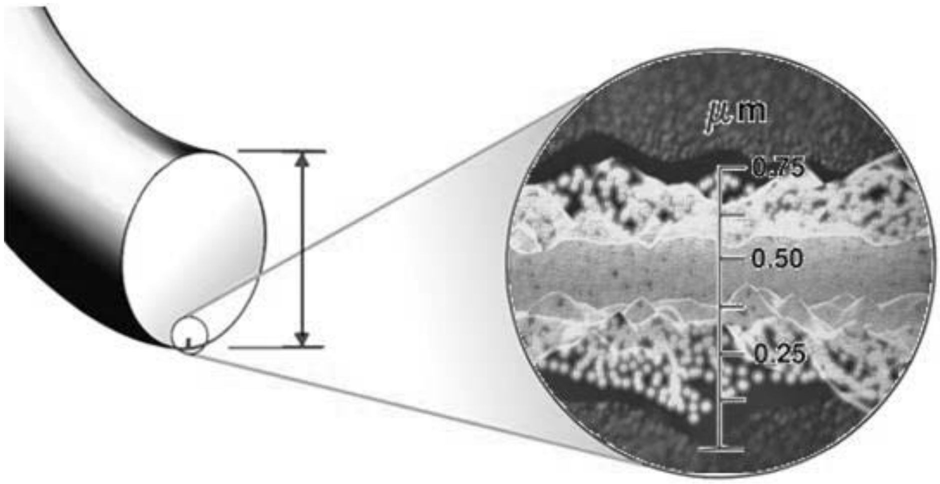


Figure 10-10

If the bearing is poorly lubricated, those surfaces will come closer together, and the high-points from each surface will come into contact. This contact is random, but the forces involved will both generate noise, and increase the level of wear.

Lubrication is very important. The volume of lubricant must be correct and it must have the properties intended for the application - the viscosity and other properties are designed to keep the surfaces apart and thus reduce friction and wear, and of course it should not be contaminated.

Installation errors

If the bearing is cocked on the shaft, we expect to see increased amplitude at 1X, 2X and/or 3X. The key indicator is phase – take measurements on all four sides of the shaft. If the bearing is cocked on the shaft, we will see a 180 degree phase shift left to right and top to bottom – it is rotating. If the bearing is cocked in the housing, there will be a 180 phase change from one side to the other, however you will need to find the axis of the angle of the bearing. If it was cocked vertically (i.e. the top pushed in and the bottom pulled out) then the 180 degree phase difference will be from top to bottom.

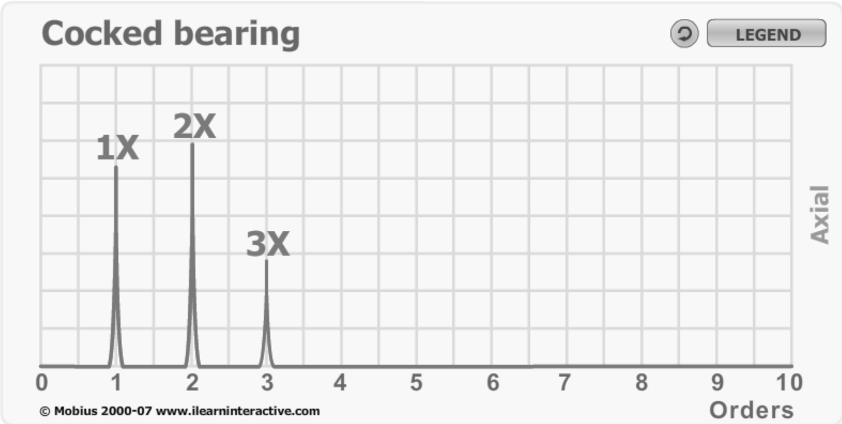


Figure 10-11

Clearance

Excessive clearance can lead to looseness if there is a driving force and to skidding if there is low load or an incorrect lubricant. If the bearing is skidding, there will not be a periodic frequency; instead high frequency noise is generated. This is because there is a rubbing rather than an impacting.

If bearings are skidding they can heat up and even catch on fire as you will see in the following example.

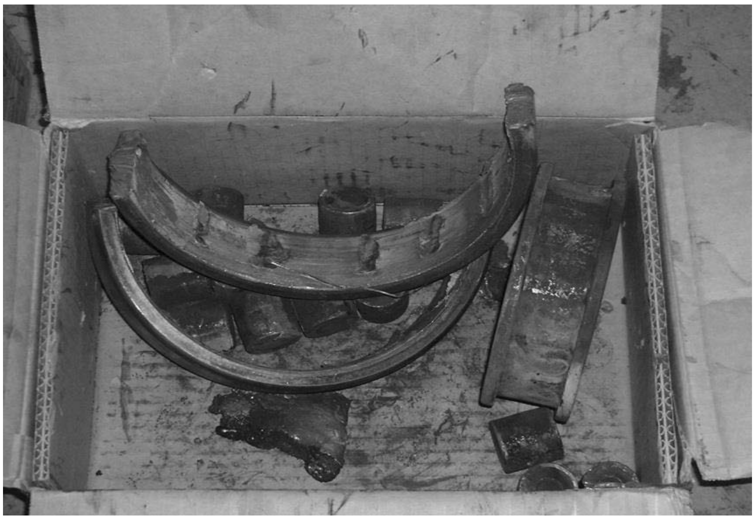


Figure 10-12

The figure above shows the burnt up bearing from a condensate extraction pump motor. The motor had run for 24 years without problems. Recently the motor was rewound, the rotor was balanced, new bearings were installed (cylindrical roller DE) and the unit was laser aligned. Three months later, the motor DE bearing over heated and the motor stopped. Another overhaul was completed, the rotor was balanced and aligned again and new bearings were installed.

Six months later a loud explosion rocked the whole station. When someone went to investigate they found the DE bearing on fire and they shut the machine down. Both coolers had been blown off the motor. One of them flew 3 meters away and just missed a hydrogen carrying pipe. The whole motor was scrapped.

Figure 10-13

The cause

The roller bearings in the motor require 30% of their design load to operate correctly. When the load is removed, the rollers skid and cause heating. When the machine was balanced and aligned, the loads were greatly reduced, the bearing skidded causing the grease to vaporize and finally ignite. In order to resolve the problem a new motor was installed with deep groove bearings on the drive end.

Rotating looseness

Rotating looseness may be the result of the shaft loose in the bearing or the bearing loose in the housing. Either way, rotating looseness results in a series of high 1x harmonics in the spectrum, often out to 10x. This may be accompanied by a raised noise floor as the result of random vibration or random impacting. If phase measurements are collected you will notice that the phase readings do not remain steady. This is because the phase reference on the shaft is not remaining steady because it is loose. In any case, an unstable phase reading is another indication of rotating looseness.

Bearing loose in housing or slipping on shaft

A bearing loose in its housing may result in a high 4x peak and a bearing slipping on the shaft may result in a high 3x peak.

Excessive voltage or current

Current flow or excessive voltage (especially where there is looseness or excessive clearance) will damage a bearing. Current may flow from the inner race, through the rolling elements to the outer race. Pitting can occur when there is a large voltage differential between the races – sparks/arcs between the raceway and rolling elements.

Surface can appear "milky" or have a rippled or "washboard" effect. The following is a classic example.

Figure 10-14

High frequency vibration is generated – a series of BPFO, BPFI or BSF sidebands will appear in the spectrum between 100,000 & 180,000 CPM. The following data was taken from the bearing above.

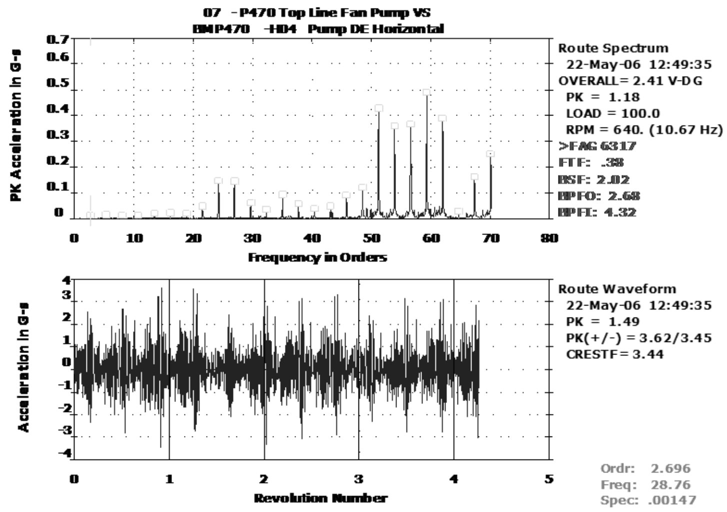


Figure 10-15

If we look at the history of the data you can see the family of peaks grow over a period of time. When the bearing was replaced, the spectra no longer displayed the peaks.

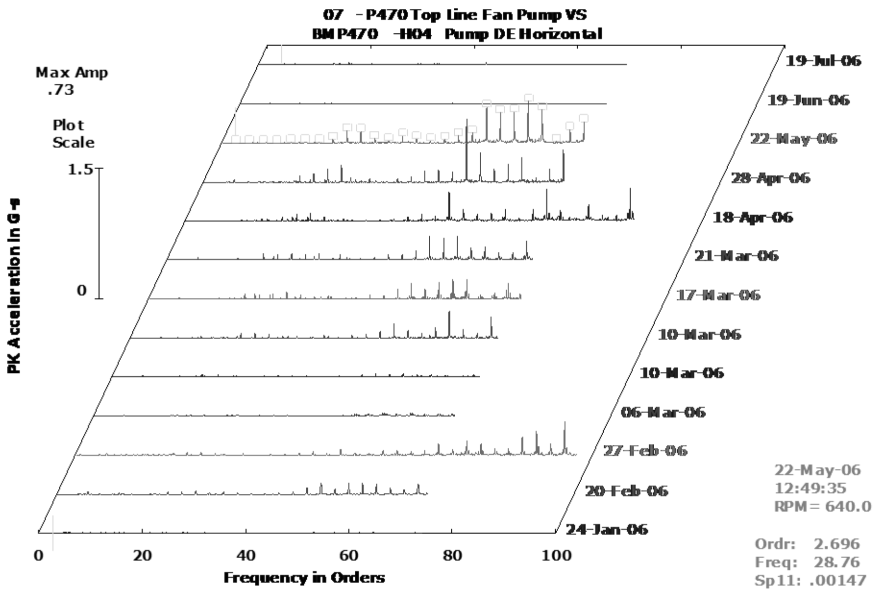


Figure 10-16

Bearing geometry and vibration

There are four key frequencies generated by rolling element bearings:

1. The Ball Pass Frequency – Inner race (BPFI). This is the rate at which a ball or roller passes a point on the inner race of the bearing. If there was damage on the inner race, we would expect to observe a periodic "pulse" of vibration at this rate.

2. The Ball Pass Frequency – Outer race (BPFO). This is the rate at which a ball or roller passes a point on the outer race of the bearing. If there was damage on the outer race, we would expect to observe a periodic "pulse" of vibration at this rate.

3. The Ball Spin Frequency (BSF). This is the angular frequency of the rolling elements. If there was a point of damage on one of the rolling elements, we would expect to observe a periodic "pulse" of vibration at this rate. In fact, because that point of damage may strike the inner race and outer race for each rotation, we may observe higher levels of vibration at twice this frequency.

4. The Fundamental Train Frequency (FTF). This frequency, also called the cage frequency, is the rotation rate of the bearing cage. It is the time it takes for a rolling element to complete one trip around the bearing. We do not always measure vibration at this frequency, but it is observed in other ways (we see sidebands of the FTF).

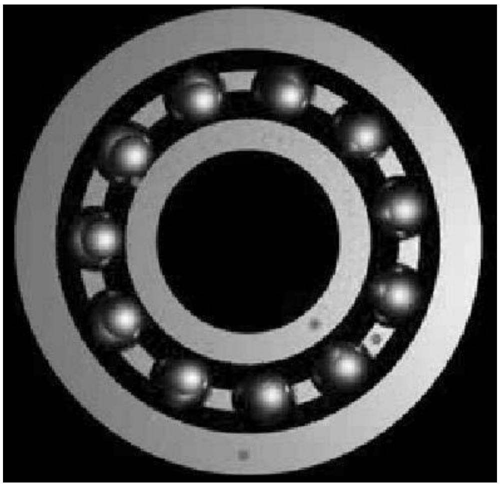


Figure 10-17

All of the following equations are based on the following dimensions.

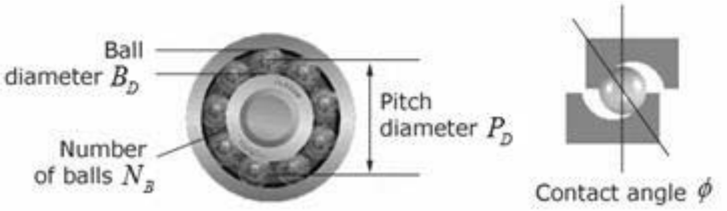


Figure 10-18

Fundamental train frequency (FTF)

The Fundamental Train Frequency, or cage frequency, can be calculated as shown below. The frequency is calculated differently depending upon whether the inner race is rotating (the most common situation) or if the outer race is rotating (for example on some conveyors and on wheel bearings).

Inner race rotating:

$$FTF = \frac{1}{2}\left[1 - \left(\frac{B_D}{P_D}\right) Cos(\emptyset)\right]$$

$$FTF \approx \frac{1}{2} - \left(\frac{1.2}{N_B}\right)$$

The FTF is commonly between 0.33X and 0.48X – however with certain thrust bearings, the contact angle is 90 degrees, and therefore the FTF will be 0.5X.

Outer race rotating:

$$FTF = \frac{1}{2}\left[1 + \left(\frac{B_D}{P_D}\right) Cos(\emptyset)\right]$$

$$FTF \approx \frac{1}{2} + \left(\frac{1.2}{N_B}\right)$$

The FTF is commonly between 0.52X and 0.67X (or 0.5X for the thrust bearings mentioned above).

The Ball Spin Frequency (BSF)

The ball spin frequency (BSF) can be calculated as shown below. The ball spin frequency (the angular frequency of the rolling elements) can be calculated the same way regardless of whether the inner race is rotating or the outer race is rotating.

$$BSF = \frac{P_D}{2B_D}\left[1 - \left(\frac{B_D}{P_D}\right)^2 Cos(\emptyset)^2\right]$$

This frequency will be non-synchronous (non-integer), and it will typically be in the range 1.9X – 5.9X.

As mentioned previously, the calculation computes the actual rotating frequency of the rolling elements; i.e. the number of times the rolling elements rotate per second (or minute). It is common to observe higher vibration amplitudes at twice these frequencies because a spall on the rolling element may come in contact with both the inner race and inner race per rotation – thus there are two events per rotation of the rolling element.

(In fact, these calculations determine the ratio of the ball spin frequency to the turning speed of the shaft – to compute the BSF in Hz or CPM it must be multiplied by the turning speed of the shaft.)

This calculation can be simplified as follows to provide an estimate of the frequency when only the number of rolling elements is known.

$$BSF \approx \frac{1}{2}\left[\frac{N_B}{2} - \frac{1.2}{N_B}\right]$$

This is important to know in the event that you observe a peak (or peak with harmonics) that you suspect may be a bearing defect frequency. If you can estimate the number of rolling elements you can repeat the calculation ± 2 rolling elements and see if one of the calculations comes close to the frequency you observe. We will discuss this "reverse engineering" approach in greater detail shortly.

Ball Pass Frequency Inner race (BPFI)

The following formula can be used to calculate the Ball Pass Frequency Inner race (BPFI). This frequency is independent of whether the inner race is rotating or the outer race is rotating.

$$BPFI = \frac{N_B}{2}\left[1 + \left(\frac{B_D}{P_D}\right) Cos(\emptyset)\right]$$

This frequency will be non-synchronous (non-integer), and it will typically be in the range 6.3X-13.8X.

Note that this formula can be simplified to a much simpler calculation. If you know the number of rolling elements you can estimate the BPFI.

$$BPFI \approx \frac{N_B}{2} + 1.2$$

Ball Pass Frequency Outer race (BPFO)

And finally, we can also compute the Ball Pass Frequency Outer race. Once again the calculation is independent of whether the inner race or outer race is rotating.

$$BPFO = \frac{N_B}{2}\left[1 - \left(\frac{B_D}{P_D}\right) Cos(\emptyset)\right]$$

This frequency will be non-synchronous (non-integer), and it will typically be in the range 4.2X-11.3X

This frequency can also be estimated as follows:

$$BPFO \approx \frac{N_B}{2} - 1.2$$

This will be the most commonly observed defect in rolling element bearings (on horizontal machines). The point at the bottom of the outer race is constantly taking the full load of the machine's mass and dynamic forces.

Bearing defect frequency tips

Please note that the defect frequencies are not exact. Due to slippage, skidding, wear, and imperfect bearing details (i.e. the dimensions may not be perfectly accurate), the frequencies may be off by a small amount.

When we do not have the part number of the bearing, or we suspect that the information we have is incorrect, then we can still estimate the defect frequencies. There are a few "rules of thumb" that you may find useful to know.

Rule of thumb #1

First, if you add the BPFI to BPFO then you can compute the number of rolling elements. For example, if BPFO is 3.2X and BPFI is 4.8X then we know that the bearing has 8 rolling elements. This is particularly useful to know because if we observe two series of peaks that we suspect are bearing defect frequencies, one BPFI and the other BPFO, then we can add them together to see if they add up to a whole number. If they do then your confidence about the origin of the vibration will increase.

Rule of thumb #2

Second, if you know how many rolling elements are in the bearing then you can estimate the defect frequencies as described in the earlier formulas, or more simply as follows:

$$BPFO \sim 0.4 \times \#Balls$$

$$BPFI \sim 0.6 \times \#Balls$$

This generalization should only be applied for bearings with 8 to 12 balls/rollers.

Rule of thumb #3

If we do know the bearing part number (and a bearing database) we may still find that there are no peaks in the spectrum that line up with the frequencies in the table. This may mean that the bearing is OK, or the bearing information is incorrect. If are peaks at frequencies that are "close by" that appear to be bearing defect frequencies (there are harmonics, they are non-synchronous, and there are FTF or 1X sidebands) then it would be useful if we could confirm that they are in fact defect frequencies. Well, there is a way to do just that:

$$BPFO\left(\frac{\#B}{R_{N\pm1,2,3..n}}\right) = BPFO \pm FTF \times n$$

Based on the defect frequency equations provided earlier, we can learn that the difference between the BPFO from a bearing with 13 rolling elements and the "same bearing" with 14 rolling elements is equal to the cage rate (FTF) of the bearing.

For example, if we look in this table of bearing information for a 22312 bearing, we can calculate the difference between the defect frequencies when different numbers of rolling elements are used. If we believed that the bearing had 13 elements however the peaks in the spectrum did not line up, we can add (or subtract) FTF (0.409) to the quoted frequencies to estimate the BPFO, BPFI and BSF for bearings with 12, 14, 15 and 16 rolling elements. We can perform the calculations and compare them to the results provided in the table. As you can see below, they correlate quite well.

Example using 22312 with 13 rolling elements. Looking at the table below we can see that the FTF is 0.409. We can now look at the other lines in the table for different numbers of rolling elements. We can compute the estimate and compare to the table.

- n=14 » BPFO = 5.312 + 1 x 0.409 = 5.721
- n=15 » BPFO = 5.312 + 2 x 0.409 = 6.13
- n=16 » BPFO = 5.312 + 3 x 0.409 = 6.539

BEARING TYPE	#B/R	FTF	BSF	BPFO	BPFI
22312	15	.413	2.697	6.192	8.808
22312	16	.411	2.665	6.584	9.416
22312	16	.411	2.665	6.584	9.416
22312	16	.411	2.665	6.584	9.416
22312	16	.411	2.665	6.584	9.416
22312	16	.411	2.665	6.584	9.416
22312	14	.409	2.570	5.721	8.279
22312	14	.409	2.570	5.721	8.279
22312	13	.409	2.570	5.312	7.688
22312	14	.402	2.385	5.626	8.375
22312	16	.411	2.665	6.584	9.416

Figure 10-19

Rule of thumb #4

It is also helpful to know that bearings of the same dimension series have very similar defect frequencies. The second the third digit of the bearing number represent the width and diameter of the bearing. So if you are unsure of the exact bearing number, but believe you know the dimension series, then you can look up the bearing database to get an estimate of the defect frequencies.

This is not an exact science, but it can help a great deal if you are not confident that the changes in the spectrum you are observing are related to a bearing with increasing wear.

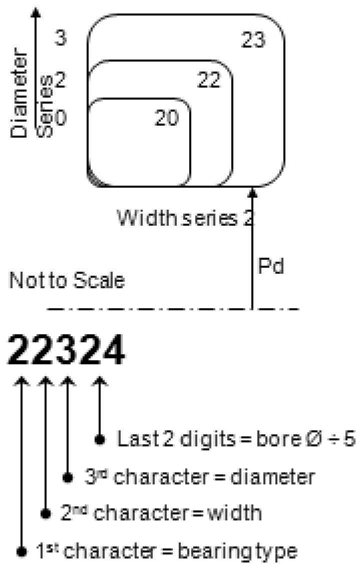


Figure 10-20

To aid in this process we can recognize that certain types/series of bearings are commonly used in the same applications. The first digit of the bearing relates to the type of bearing, as shown below.

Figure 10-21

As a guide only, the following table shows the type of bearing commonly used in the listed applications.

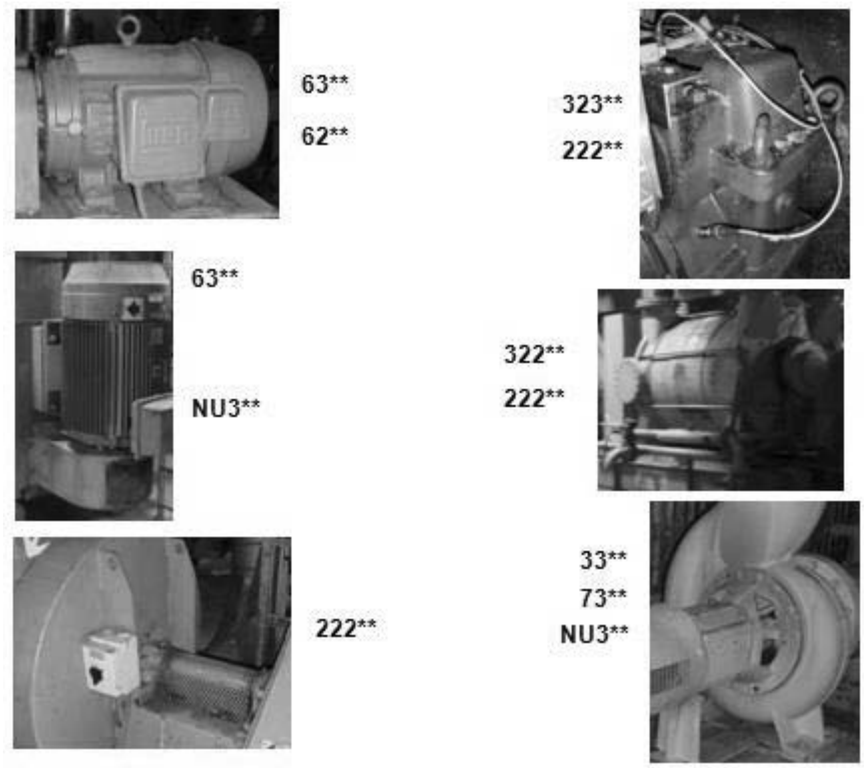


Figure 10-22

Again, it must be stressed that this is not an exact guide; instead it is provided to help you make an "educated guess" about the approximate defect frequencies. If you know the bearing type and/or application and/or number of rolling elements, and you have a bearing database, then this information should help you be more successful detecting bearing fault conditions.

Defect frequencies are non-synchronous

There is another key observation we can make; all of the defect frequencies are non-synchronous – that is, they are non-integer multiples of the shaft turning frequency. We will soon learn in detail that bearing faults (wear, spalls, cracks) will generate harmonics in the spectrum, and there will often be 1X sidebands (inner race fault), FTF (<0.5X) sidebands (ball/roller damage), or no sidebands (outer race fault).

Therefore we can detect bearing faults without knowing the bearing number with a fair degree of confidence. Coupled with the rules of thumb that we have just covered, we can be quite confident of detecting faults without a great deal of supporting information.

Vibration – The complete picture

It is important to understand the nature of the vibration generated by rolling element bearings. We just mentioned the high frequency vibration that is generated, but as we will learn in greater detail, we can break this vibration up into four components:

1. Friction: There will always be some friction between surfaces, regardless of how good the lubricant may be. As the lubricant degrades (too much, too little, lubricant additive pack degradation, or contamination), friction will increase. Frictional vibration is non-periodic. Friction generates broadband noise – vibration at a wide range of frequencies.

2. Stress wave: If we have metal-to-metal contact we will generate very high frequency, very short duration pulses. These pulses, or shock waves, may be random in nature or periodic – it depends upon the reason for the metal-to-metal contact.

3. Periodic vibration: When an even occurs periodically, at a sufficient amplitude to be detected, the vibration will be present in the time record and a peak will appear in the spectrum. Common examples are the vibration related to unbalance or the vane pass rate, however when there is damage on the inner race of a bearing, for example, a periodic source of vibration will also be generated.

4. Resonance: When surfaces impact they generate vibration, *and* the structure and machine components (and the attached accelerometer) will resonate as a result. The natural frequency may be quite high when dealing with vibration from bearings. We can detect the vibration from those resonances because the vibration is amplified and because the frequencies are distinct from the much higher amplitude vibration that masks the lower frequency vibration generated by the bearing.

This vibration can be detected using a number of techniques, summarized below:

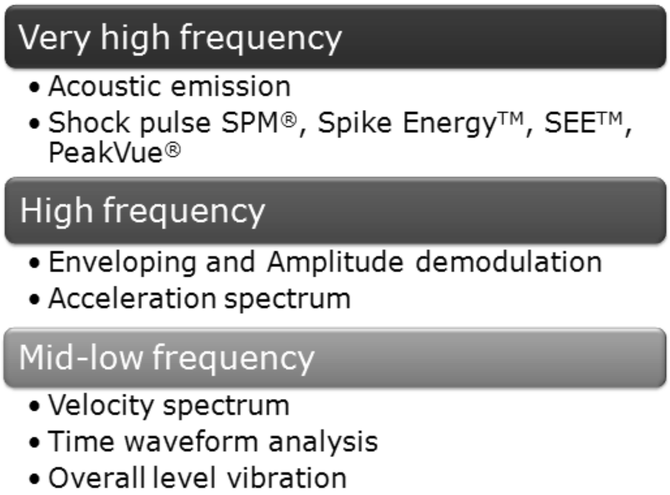


Figure 10-23

The vibration generated due to poor lubrication or minor bearing damage is very high in frequency, but low in amplitude. The frequency generated will be over 20 kHz in the earliest stage – some systems look at much higher frequencies to detect the first signs of bearing wear. We need techniques to remove the low frequency, high amplitude vibration in order to detect this source of vibration. Techniques such as acoustic emission, Shock Pulse, Spike Energy, and PeakVue are designed to do just that.

Enveloping and demodulation are also designed to monitor high frequency vibration and detect first stage bearing wear, however some systems are not designed to monitor such high frequencies, and many systems capable of detecting these high frequencies are not set up to do so.

As the fault develops the vibration generated is lower in frequency. The high frequencies are still generated, however now lower frequency vibration (over 2 kHz) is generated. Time waveform analysis and spectra (in units of acceleration) can be used to detect the fault and enveloping/demodulation can be used.

With increased wear comes higher amplitude, lower frequency vibration. Now the conventional velocity spectra will display signs of bearing wear.

And finally, as the fault develops further, and damage becomes more severe, low frequency noise is generated, and some of the characteristic patterns we expect to see begin to change to vibration patterns normally associated with looseness. At this stage we may not measure the high frequency vibration; there is too much damage.

This process, and the analysis techniques, will be discussed in greater detail in this chapter.

Stage One Bearing Faults

The following points summarize the characteristics observed when a bearing is in Stage One:

Sub-surface damage only
- Friction and minor impacts

Very high frequency vibration
- Friction: greater than 20 kHz in earliest stage
- You can't hear it (without assistance)
- 'Noise' due to inadequate lubrication
- Very low levels
- *Very* short duration impacts

'Stress waves' or 'Shock pulses'
- 1 kHz » 15 kHz

Traditional vibration analysis techniques are inadequate at this stage.

The following graphic provides a summary of the change you will see in the spectrum. Note that the velocity spectrum will not show any sign of the fault condition.

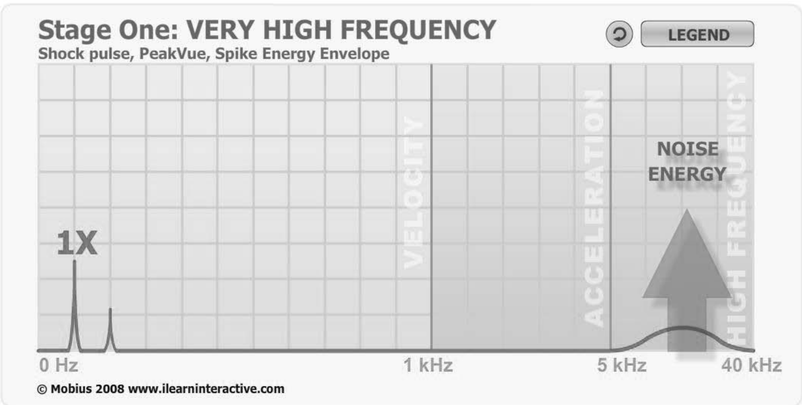


Figure 10-24

The effectiveness of each bearing detection technique is summarized below:

HFD	• Yes: Effective methods • Trend will continue upwards
Envelope	• Earliest signs of faults • Noise floor may rise
Spectrum	• *Maybe:* Acceleration spectrum • Not: Velocity spectrum
Time Waveform	• No: Vibration in noise floor • Effective for low speed machines

Figure 10-25

The following table summarizes the possible actions you could take at this stage:

Damage	• Very difficult to see damage • Sub-surface damage
Life	• 10-20% of L_{10} life
Action	• Lubricate correctly • Continue monitoring

Figure 10-26

Stage Two Bearing Faults

The following points summarize the characteristics observed when a bearing is in Stage Two:

Sub-surface damage only
- Friction and minor impacts

Very high frequency vibration continues to increase in amplitude.

Vibration significant enough to excite 'resonances'.
- Structural, bearing, and sensor

Envelope (demodulation) spectrum should show signs
- Defect frequencies present in spectrum
- Velocity spectrum still won't indicate fault. Acceleration spectrum should indicate fault.

Figure 10-27

The following graphic provides a summary of the change you will see in the spectrum. Note that the velocity spectrum will not show any sign of the fault condition.

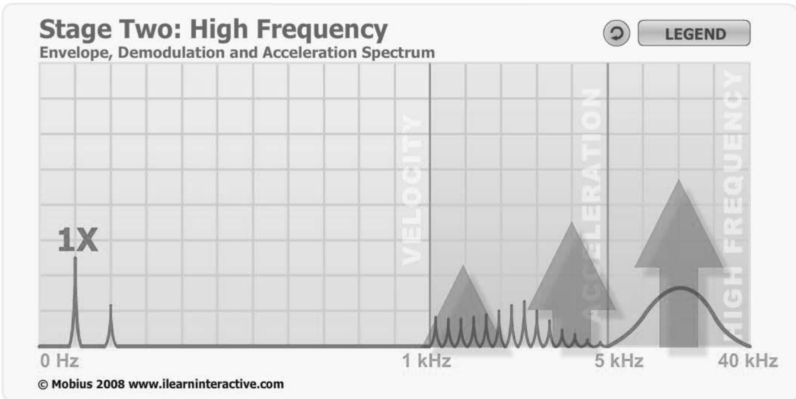


Figure 10-28

The effectiveness of each bearing detection technique is summarized below:

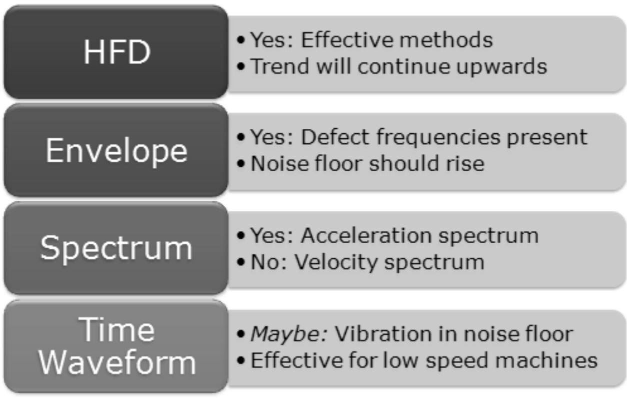


Figure 10-29

The following table summarizes the possible actions you could take at this stage:

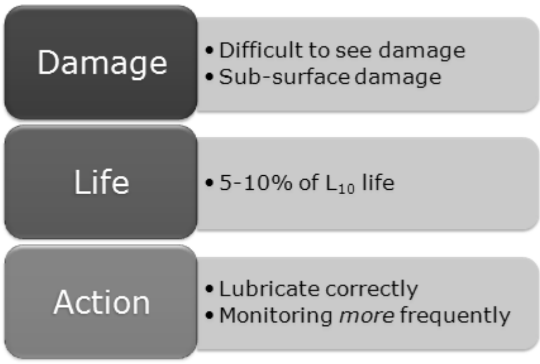


Figure 10-30

Before we discuss stage three and four bearing faults, we will discuss the high frequency vibration analysis techniques.

High frequency vibration analysis techniques

As the fault develops in a bearing, we will see the stress wave, and if the forces are great enough, we will also see resonance of one of more of the bearing components. We will also see excitation of the accelerometer itself.

We will now explore this process in greater detail.

Stress waves (shock pulses)

Stress waves are produced in rotating machinery when metal-to-metal contact is made in rolling element bearings and gears from faults such as impacting, fatigue, friction, cracking, scuffing and abrasive wear. When no faults are present, lubrication buffers metal surfaces from contact. When the surfaces become damaged or the lubrication runs out, metal contacts metal and stress waves are produced.

Figure 10-31

Contact between two surfaces produces a shock or pressure wave that radiates away from the point of contact at very high speed. This pressure wave is called a Stress Wave. After the stress wave passes, secondary vibration may occur from the recoil and resonance caused by the impact. The duration of the stress wave is very short compared to the secondary vibration.

To visualize a stress wave, think about dropping a pebble in a pond. The rippled wave moves outward in all directions from where the pebble entered the water.

The waves radiate away from the fault area at the speed of sound in metal at 5,000-10,000 feet/second. As the stress waves move through the machine, natural frequencies are excited resulting in the occurrence of secondary vibration (resonance). The stress wave's energy diminishes quickly. In order to be effectively measured, a direct transmission path through steel is required. The waves do not travel through air gaps and are attenuated through interfaces such as the ball to the race, the race to the housing and the housing to the sensor.

We can utilize this information, however we must overcome a number of challenges. We will now discuss the challenges and the solutions.

Challenge one: Low amplitude

A vibration analyzer has a dynamic range of about 90 dB and a typical accelerometer has a slightly better dynamic range. This means that for two spectral peaks to be visible in the same spectrum the amplitude difference between the peaks must be less than 90 dB. If it is greater than 90 dB, the smaller of the two peaks will fall below the noise floor and be invisible. It is to be expected that the vibration signal will contain low frequency, high level energy related to unbalance, misalignment and other faults. If the signal also contains high frequency, low amplitude bearing fault energy, it is very likely that the high frequencies will not be seen. The low frequencies dominate the signal causing the high frequencies to fall below the noise floor of the measurement.

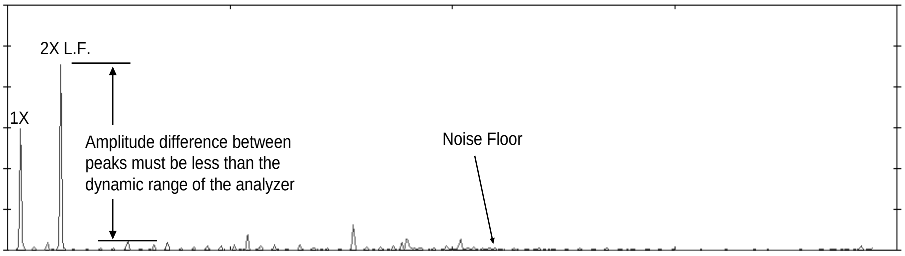


Figure 10-32

Challenge two: Short duration

A stress wave is a very short duration spike or impact lasting several microseconds to a few milliseconds. The duration of the event depends on the type of event *(impacting, friction, fatigue cracking)*, the location of the accelerometer to the fault area and the severity of the fault.

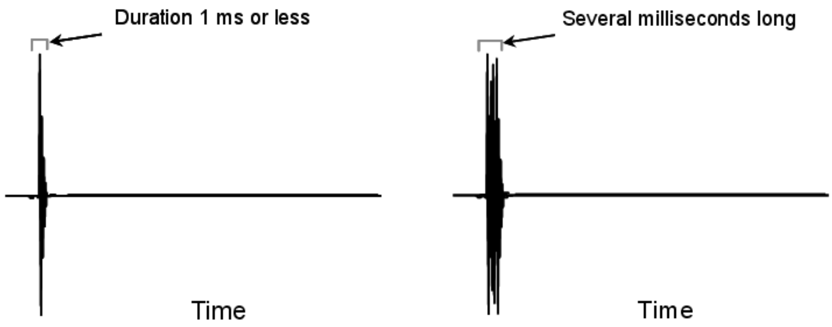


Figure 10-33

Bearing and gearbox faults are mostly periodic and therefore produce reoccurring stress waves. Stress wave magnitudes are very small in the early stage of bearing failure. The signal may appear stronger at one instant and weaker at another due to slight changes in load or lubrication film thickness.

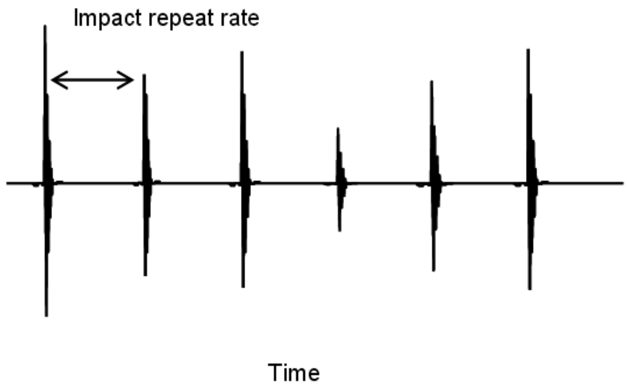


Figure 10-34

If the signal is sampled too slowly, some of the stress wave events fall between samples and are lost.

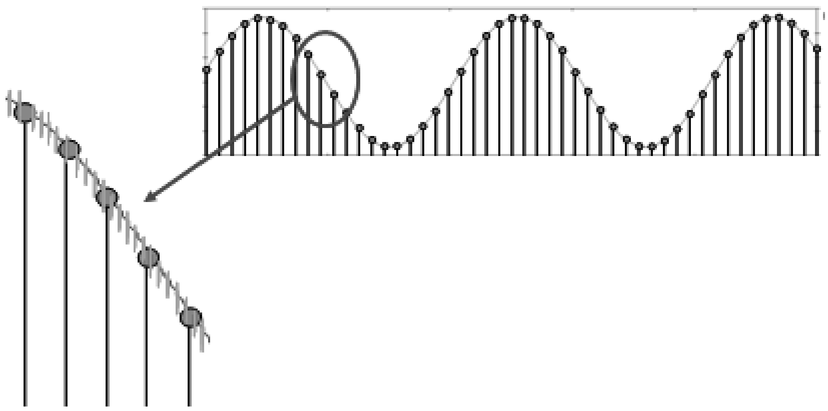


Figure 10-35

Stress waves produce energy in the 1,000–50,000 Hertz frequency range. Different faults types produce stress wave energy which is concentrated in different parts of the spectrum. Stress wave energy from friction and fatiguing is independent of machine mass and is concentrated higher in the frequency range (above 10,000 Hz.). The stress wave energy from impacting faults depends on machine mass. For most bearings with impacting faults, the stress wave energy is concentrated in the 800–10,000 Hertz range.

Challenge three: Measurement

A general purpose, 100 mv/g accelerometer has a frequency response from near zero Hertz (DC) to slightly above the mounted natural frequency. If stud mounted to a smooth, flat surface, a general purpose accelerometer might have a mounted natural frequency of 14,000 Hertz. The amplitude response of the accelerometer is flat from about zero Hertz out to the point where it approaches its mounted natural frequency. Signal amplification results around the mounted natural frequency and signal attenuation occurs above the mounted natural frequency. The response curve around the mounted natural frequency of the accelerometer is called the amplification curve. The frequency response of the accelerometer decreases when the accelerometer mounting technique is changed. Magnets and mounting adapters add mass and are less stiff than the stud mount. Increased mass and decreased stiffness both lower the frequency response of the accelerometer.

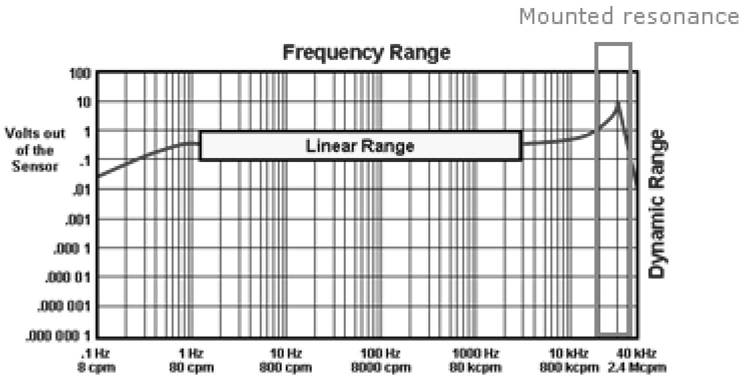


Figure 10-36

Stress wave measurements should be made as close to the bearing as possible in the load zone of the machine. Stay away from the *unloaded zone* of the bearing. On a horizontal direct coupled machine, gravity is pulling down on the machine. The most loaded zone on the bearing is down and the least loaded zone is straight up. It's not good practice to take a stress wave measurement at the top of the bearing housing where all of the clearance is located. A horizontal measurement is usually a good choice on a direct coupled horizontal machine. For angular contact bearings, the axial direction is a good choice for stress wave measurement.

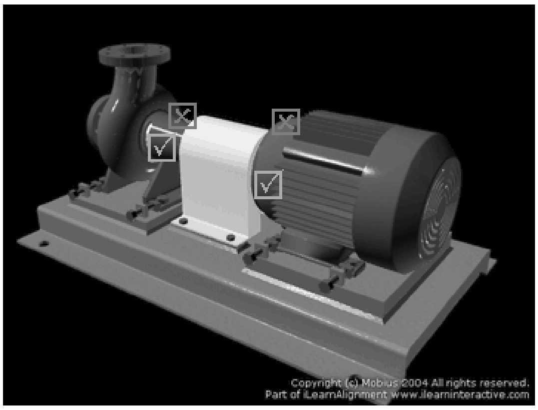


Figure 10-37

The amount of stress wave energy that actually makes it to the sensor is determined by many factors such as the type and severity of the defect, distance between the fault and the sensor, transmission path, speed of the machine, sensor type and sensor mounting.

Every high frequency bearing fault detection product provides specific instructions related to the sensor type and mounting. Some products require the use of a specific sensor and mounting method. Other products place fewer requirements on the sensor type and mounting.

Compromise is an accepted part of route based vibration measurements. Use common sense and take care how and where the accelerometer is mounted. Put the accelerometer in a place that has good signal transmission to the bearing. Clean and smooth the surface if needed.

Solutions: Four different approaches

There are basically four different approaches taken by different vendors in our industry:

- Use the accelerometer to amplify the high frequency vibration.
- Use high speed data acquisition to capture the stress waves.
- Use demodulation techniques to capture high frequency vibration.
- Monitor the machine with acoustic emission (airborne ultrasound) instruments.

Airborne Ultrasound (Acoustic Emission)

When there is insufficient lubrication, the bearing will generate high frequency "noise" as the microscopically rough surfaces of the bearing begin to make contact. As the lubricant is depleted, so the noise level will increase. Unfortunately we cannot hear this sound; it is well above 20 kHz – the limit of our ears.

The term "ultrasound" means "above our hearing range". For most adults, the limit is more like 16 kHz.

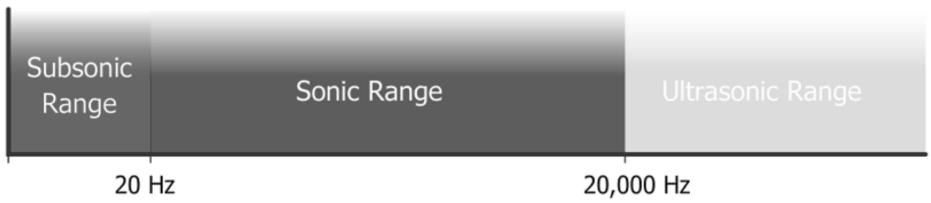


Figure 10-38

Products exist that are able take the high frequency sounds and convert them, through a process called heterodyning, to a frequency that we can hear.

Figure 10-39

The measurement process is quite simple. The technician places the end of the probe on the bearing and listens to the sound generated.

A low, muffled smooth noise, and stable readings, indicates a normal functioning bearing.
A high-pitched sound or a rushing sound indicates a lack of lubrication, an overload, or a rotating speed beyond the specification of the bearing.
A crackling noise, associated with unstable (or random) readings, indicates faults, wear or loose metal particles in the lubricant.

Many meters also provide a dB display so that the noise level can be compared to previous readings and trended.

A NASA study provided to us by UE Systems has shown the following relationships. Note that a baseline dB reading is collected while the bearing is being greased. Future dB readings are compared to this baseline.

- 8 dB increase indicates pre-failure lack of lubrication
- 12 dB increase indicates beginning of failure mode
- 16 dB increase indicates advanced failure condition
- 35-50 dB increase warns of catastrophic failure

This technology is used in other areas of condition monitoring; leak detection, detection of electrical faults, and others.

Some people also use this technology whilst greasing their bearings. While slowly adding grease they listen to make sure the grease actually reaches the bearing; the lines may be blocked and the grease could be going into the motor.

Figure 10-40

Demodulation/enveloping

Demodulation and enveloping are two names that refer to the same technique. These techniques take the high frequency vibration that we observe as bearings fail and make it available as low frequency vibration that we can analyze.

In order to understand what is happening, consider ball or roller impacting a spall in the outer race. As we have seen, the roller will pass or hit the spall at a certain rate as dictated by the bearing fault frequency calculations. This gives us the repetition rate of the impacting and this is what we look for in the normal velocity spectrum. Perhaps there are 3.4 hits impacts per revolution of the shaft; the frequency we look for is 3.4x.

Now consider that each impact actually generates a high frequency vibration or sound. You can think of it as the sound the bearing makes when the roller hits the spall. This sound will also happen repetitively each time the spall is hit by a ball. In brief, what demodulation techniques do is listen for this high frequency "sound" occurring and then using a bunch of filters, the technique tells us the repetition rate of the impacting – which is the bearing frequency.

There are a number of steps in the process.

Step one: High-pass or Band-pass Filter

We are working with a vibration signal that is dominated by low frequency, high amplitude vibration that comes from the unbalance, misalignment and other conditions. The vibration signal also contains a very high frequency, very low amplitude signal. It is very difficult to detect this signal without special processing. The following is an example, however the bearing vibration is GREATLY amplified.

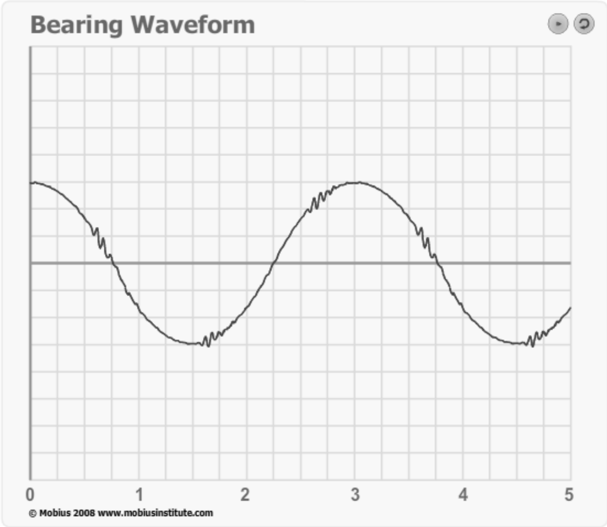


Figure 10-41

Now we will exaggerate the vibration amplitude still further.

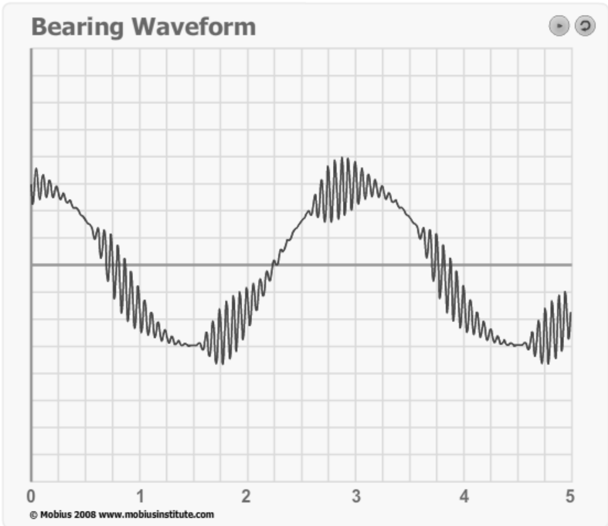


Figure 10-42

Once we filter the vibration signal we will have only the vibration that existed within the filter characteristics. The intent is for the bearing ringing to be captured. We must be careful to set the filter values so that we do not also capture vibration components from the motor or gearbox, etc.

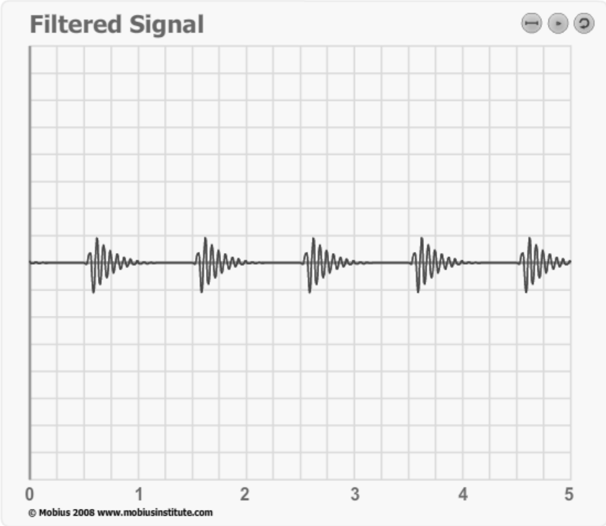


Figure 10-43

What we have now is a signal that shows only the periodic ringing of the bearing (or accelerometer, or of some other metallic component). We do not care about the frequency of the resonance (i.e. the high frequency component. All we want to know is whether we have ringing at all, what the amplitude is, and how often the ringing occurs – that should correspond to the bearing defect frequency. In this example we are simulating the vibration from a defect on the outer race.

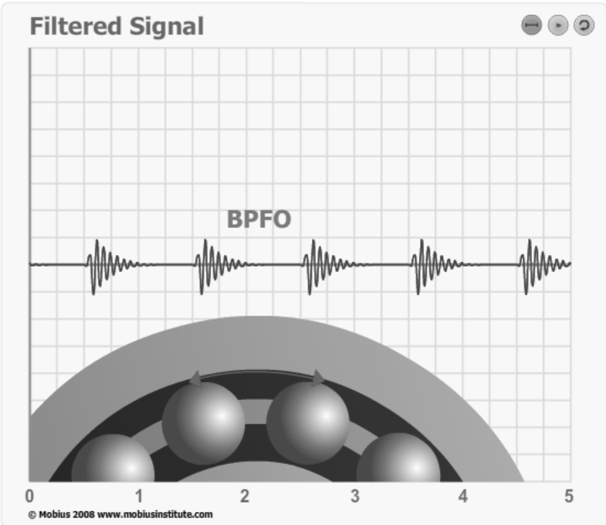


Figure 10-44

Step two: Rectify (or Envelope)

The task at hand is to capture the ringing transients and create a low frequency signal that we can view in a time waveform or spectrum. We can do this by "rectifying" the signal. (Later we will discuss the Spike Energy and PeakVue techniques that do this a little differently).

If we pass the signal through a "full-wave rectifying" in the analog domain, or perform the same task digitally, the negative-going vibration is folded over and made positive.

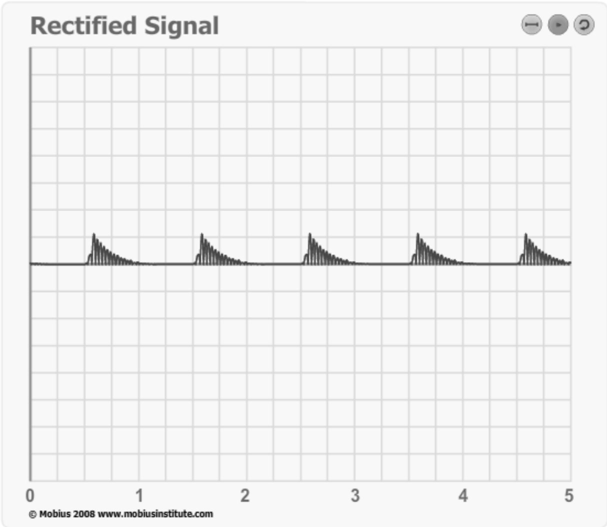


Figure 10-45

Step Three: Low pass filter

Now that we have the rectified signal, we can pass it through a low-pass filter. This has the effect of removing the high-frequency ringing. However, because we rectified the signal, we will be left with a waveform that is the same basic shape as the pulses (it looks like an envelope of the data).

The low pass filter may just be the anti-aliasing filters used to generate the spectrum. You must select an Fmax for the envelope/demod spectrum.

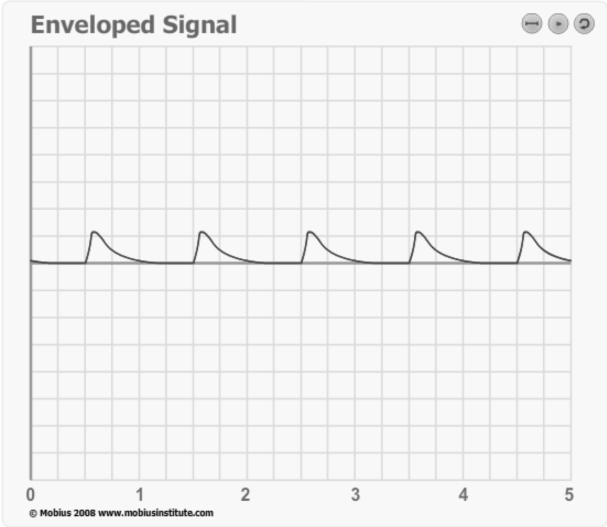


Figure 10-46

Setting up the measurement

The enveloped signal can be routinely collected and should be included in normal route measurements. Both the demodulated time waveform and spectrum can be analyzed. Remember that although internally, the demodulation technique is looking at the high frequency sounds made by the impacting in the bearing, what it returns to you is the repetition rate of the impacting, which is the bearing defect frequency and is relatively low. Therefore, when setting up the measurement, the demodulated spectrum should be configured with an Fmax setting equal to about 3.5x the calculated inner race defect frequency for the bearing. If you do not know the inner race defect frequency and the bearing has between 8 and 12 balls or rollers then you can make the Fmax 15x to 20x.

Step four: Analyze it

We can now analyze the waveform or spectrum. If we see peaks separated by BPFO, BPFI or BSF, then we can be confident that there is bearing wear – a defect exists.

We can use multiple band-pass or high-pass filter settings to focus on different frequency spans in order to assess the severity of the fault; the vibration appears in the higher frequencies first.

We can also study the demodulated (enveloped) spectrum and compare the height of the defect frequency peaks to the height of the noise floor.

If the bearing is in good condition, then the demod spectrum will be all noise (assuming that nothing else is generating impacts: gear wear, looseness, etc.).

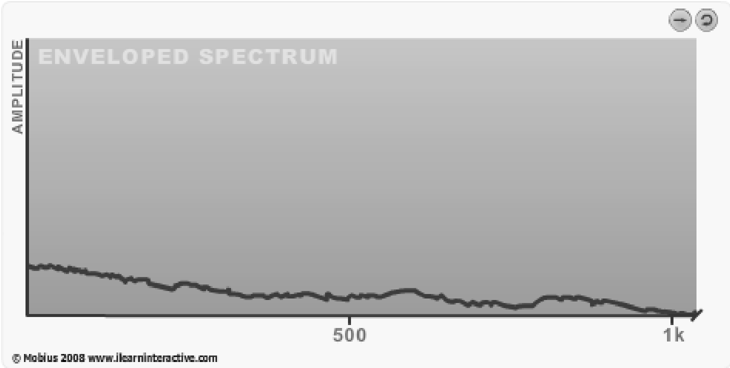


Figure 10-47

Once the fault begins to develop, the peaks will begin to grow. With most products on the market, it is not possible to state what the levels should be in order to assess the condition of the bearing. Instead it is a case of watching what frequencies are present and to compare the height of the peaks to the height of the noise floor.

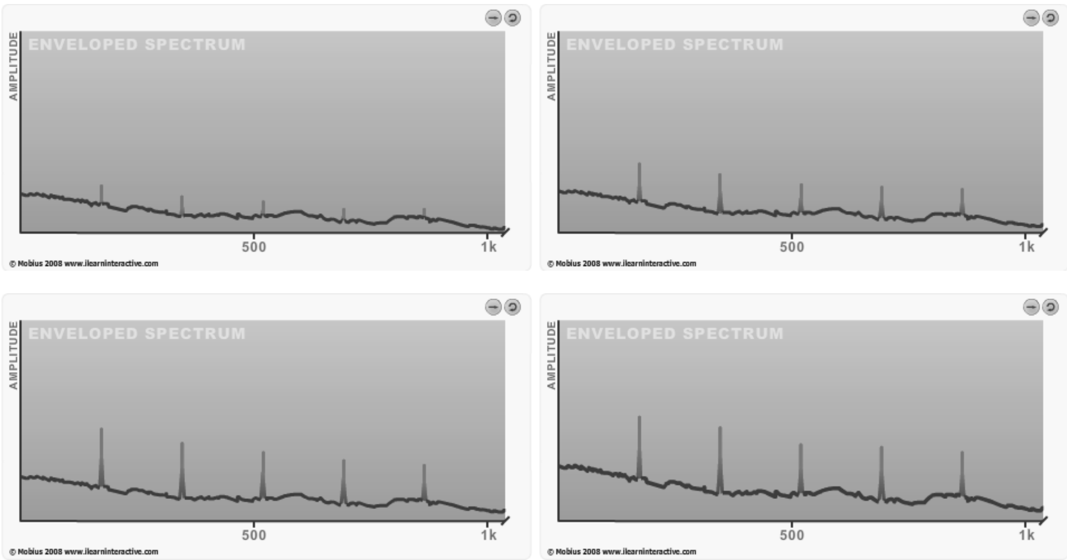


Figure 10-48

Once the noise floor begins to lift and swallow the peaks, so know that the bearing is in stage four and will fail soon.

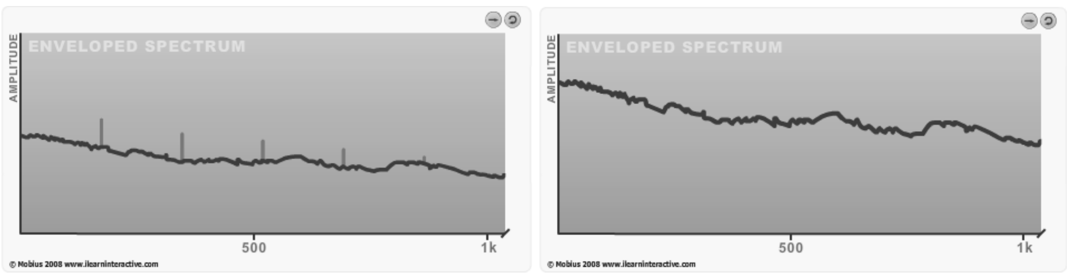


Figure 10-49

Filter settings

One of the keys to being successful with enveloping is the selection of the filter settings. It is important that the filter band does not contain forced frequencies from the machine. If the gear mesh frequencies (or harmonics), or blade pass frequencies, or rotor bar pass frequencies, happen to fall within the filter range, they will adversely affect the results.

It is also important that the filter band contains a natural frequency from the machine or accelerometer. If there is not a resonance in the band, we might have the vibration necessary after we filter the data. There are two ways to approach this issue:

1. You can simply select multiple filter bands supported by your analyzer and then either:

    a. See which gives you the best results, or

    b. Use the bands to provide a warning of the progress of the bearing defect. Because the fault will appear at the higher frequencies first, we may see the demod spectrum taken from the higher frequency bands before the lower-frequency bands. You can be assured of detecting the fault, and you get the bonus of learning more about the severity of the fault.

2. You can study a high frequency (the highest Fmax supported by your analyzer) in dB format and look for resonances.

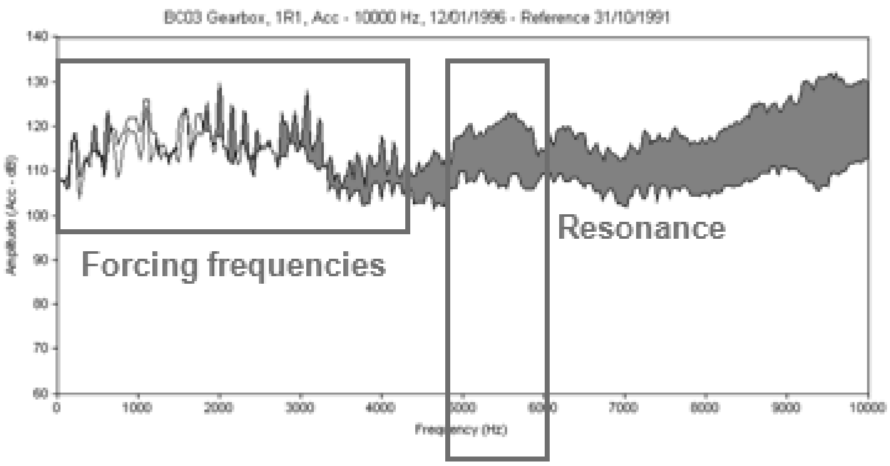


Figure 10-50

Sampling rate

The second issue is whether the technique can capture the stress waves that we are trying to detect. If the amplitude is high enough, and the resonance is strong enough, then these methods will work.

But what if the duration of the stress wave is very short? We have previously discussed the potential problem; the sample rate of the system could be too slow to actually detect the short-duration pulses.

Two products have electronics designed to specifically address this issue: PeakVue and Spike Energy. The shock pulse method is designed to specifically excite the resonance of their calibrate sensor which they monitor.

The Shock Pulse Method

The shock pulse method was developed by the SPM Instrument company in Sweden. The shock pulse method is designed to perform at least three functions:

1. Provide an indication of the condition of the lubricant.

2. Provide an indication of the condition of the bearing.

3. Allow the analyst to view spectra that reveals more detailed information about the condition of the bearing.

Lubrication

To recap the section that discussed lubrication earlier in the chapter, the lubricant in the bearing is designed to keep two surfaces apart. We should focus on the interface between the rolling element (ball) and the outer race. As the balls roll along the outer race the pressure on the ball and outer race are tremendous. The lubricant must keep the surfaces apart and reduce friction. At the point where the load is concentrated, the lubricant acts like a solid and there is actually deformation in the ball and outer race. It is this constant load that ultimately causes the metal to fail and spalls to develop.

If we look at these surfaces under a microscope we would see that they are actually quite rough. The following graphic compares the gap between the ball and the outer race to a human hair. Clearly we are talking about a very small gap. The oil is thick enough to keep these rough surfaces apart. Having said that, there will still be occasional contact, and the stress in the surface of the ball and outer race will cause vibration.

Therefore, at this point, we expect random noise to be generated. There is nothing to make the vibration periodic – no damage as yet.

The shock pulse method will report this “noise” as the “carpet” level as a dB value. The vibration is random, however the impacts do generate stress waves, and they do excite the resonance in their calibrated sensor.

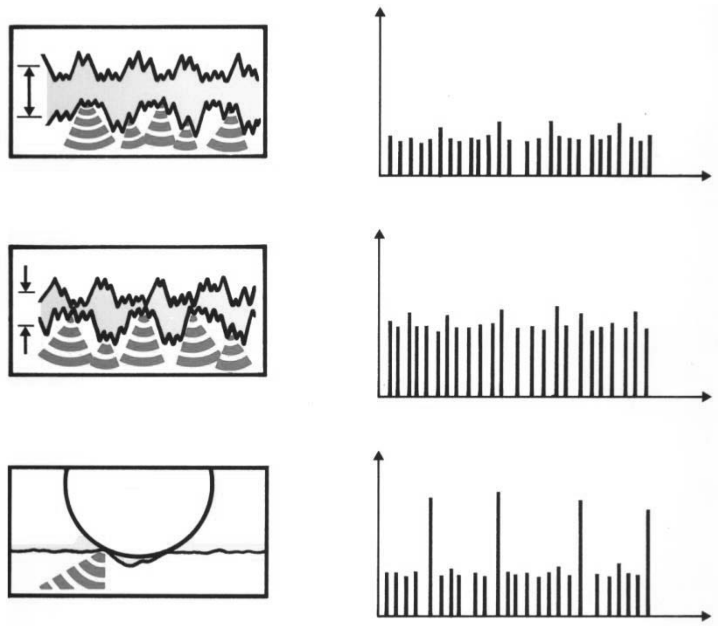


Figure 10-51 - [Graphic from SPM Instrument Company]

If there is insufficient lubricant, these surfaces will come closer together. Now more and more of these "peaks" in the surface roughness will come into contact. The level of noise will increase. We may hear the increase in noise with acoustic emission products, and the shock pulse method will show an increase in the carpet level, but more importantly, a greater increase in activity above the carpet. Thus a trend of the carpet versus the exceptions provides an indication of the state of the lubricant. This is called dBc and dBm.

It would be possible at this point to lubricate the bearing and see if the readings return to their normal level. If they do but only temporarily, then you will know that damage has occurred. If the sensor has been mounted correctly, the values provided by the shock pulse method do provide a good indication of the state of the lubricant.

If no action is taken (i.e. the bearing is not lubricated), then the random noise will increase. There will be more exceptions above the carpet – but again, they will be random; there is nothing periodic about the vibration - until the bearing begins to fail.

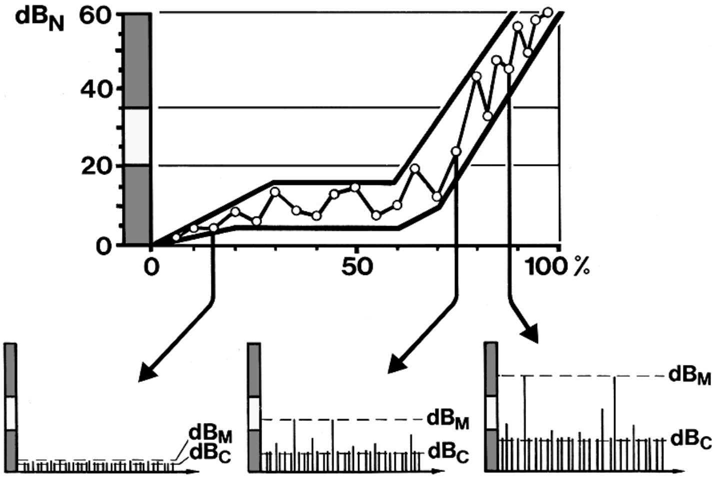


Figure 10-52 - [Graphic from SPM Instrument Company]

Whether it is because of poor lubrication or because of prolonged operation under load, or because of excessive load (misalignment, etc.), a subsurface defect may develop. With time the defect will develop further and metal will break away from the surface. Now we will get stress waves due to metal-to-metal contact, and now they are periodic. The time between each contact is the same as the defect frequencies described earlier. If the damage is on the outer race, the time between events (either compression in the subsurface defect, or metal-to-metal contact) will be the inverse of the BPFO frequency (i.e. the period).

The shock pulse method is designed to detect the increase in the exceptions above the carpet level, and it can provide a spectrum of the signal, which is similar to the enveloping technique.

The SPM sensor

SPM manufactures its own sensor. The transducer is designed to resonate at 32 kHz. The resonance *amplifies* the shock pulse and other vibration in that frequency range. In order to get the best results, the transducer is mounted in the load zone. It must also be mounted as directed; it is not mounted magnetically.

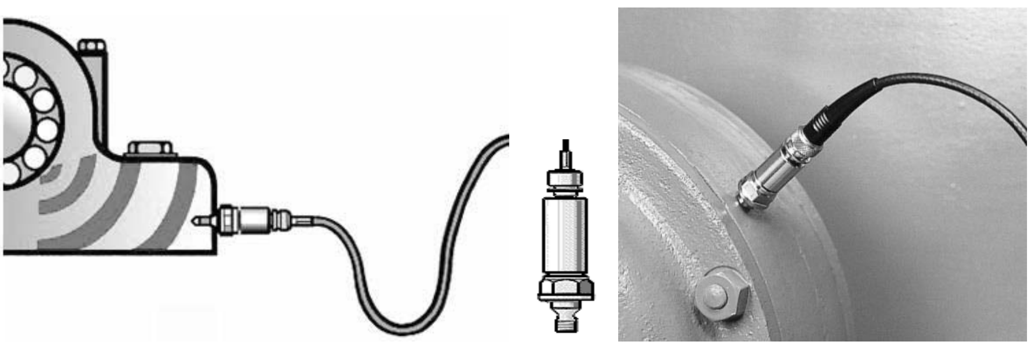


Figure 10-53

The PeakVue method

The PeakVue method is available from the CSi division of the Emerson company. It was specifically designed to overcome the potential loss of short duration stress waves, and to provide a trendable vibration waveform and spectrum.

The issue of stress wave duration was covered earlier in the manual.

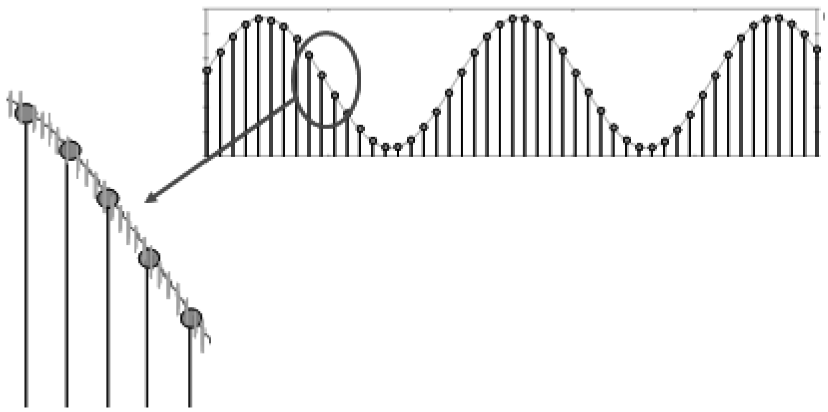


Figure 10-54

For example, if you select an Fmax of 1000 Hz the sample rate is 1000 x 2.56 = 2,560 samples per second (i.e. there is 0.39 ms between samples). For the PeakVue method, the signal is sampled at 102,400 samples per second (~0.01 ms between samples). The sample rate is 40 times faster!

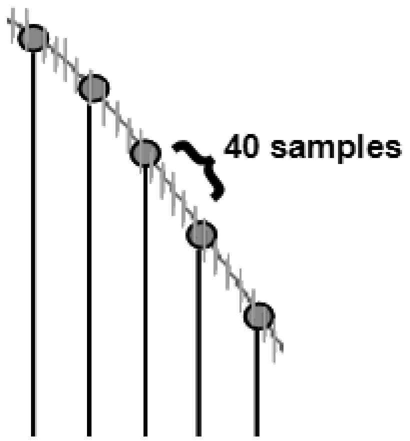


Figure 10-55

PeakVue processing creates a PeakVue time waveform comprised of stress wave absolute peak values. The PeakVue waveform is used for three tasks:

1. The magnitude of the PeakVue waveform is used for severity assessment,

2. Compute the PeakVue spectrum where periodic faults are identified and

3. Compute the auto-correlation function coefficient which is used as a diagnostic aid.

A PeakVue spectrum and waveform are shown below.

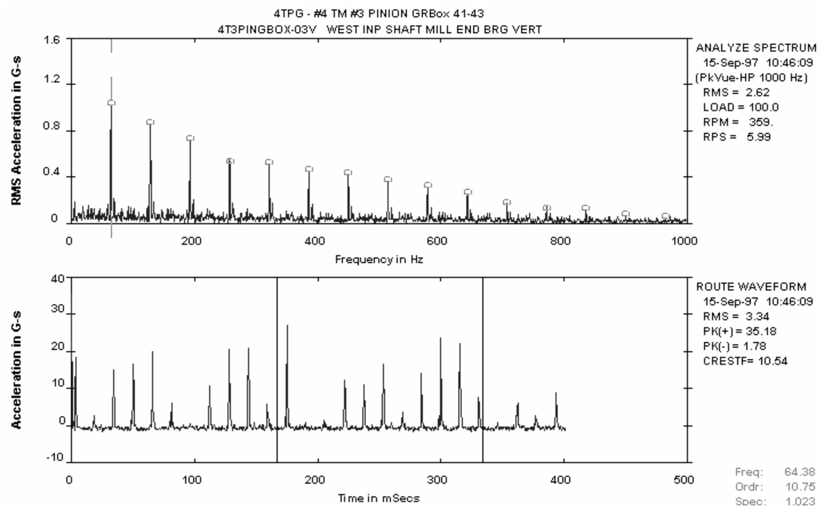


Figure 10-56

The waveform peak parameter represents the maximum PeakVue signal level and should be used for severity assessment and alarming. Peaks and patterns displayed in the PeakVue waveform and spectrum represent stress wave sources which correlate with calculated fault frequencies.

The Spike Energy method

The Spike Energy method was developed by the IRD Mechanalysis company. The principle of operation is very similar to all of the methods discussed that far (although it should be noted that the Spike Energy method was one of the first, if not the first, method to be developed). Like the Shock Pulse method, it relied on capturing the vibration from the resonance of the accelerometer. The difference is in the processing of the data. Instead of applying a "simple rectifier/low pass filter," a special technique was developed (in analog circuitry) to capture the peak amplitude and decay the signal in a repeatable way. The processed signal looks similar to the previous versions in this illustration; but it is actually different.

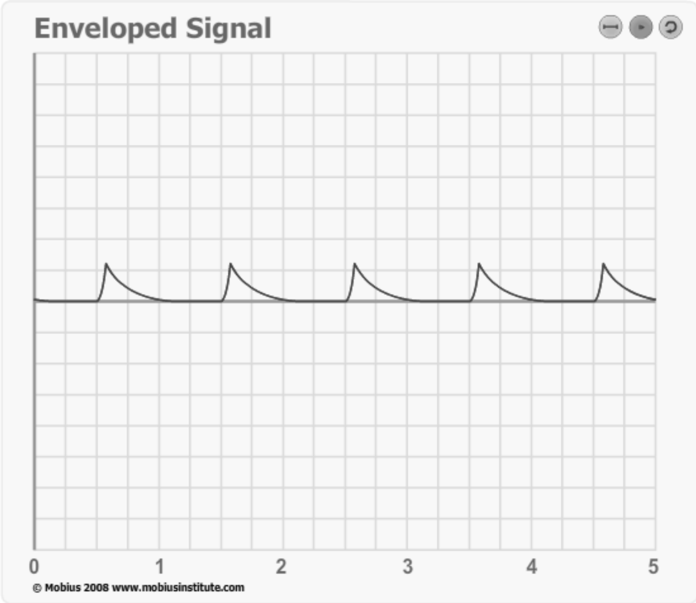


Figure 10-57

The Spike Energy technique was revolutionary in its time. You could trend a "gSE" value that gave an indication of bearing wear, and you could view the time waveform and spectrum.

It is still successfully used today, however there is an important difference today. IRD used to manufacture its own calibrated sensors (just like SPM does today). Therefore when you quoted a gSE level it meant something. Today general-purposes sensors are used. The sensitivity of sensors, and the amount of amplification at the resonant frequency, now varies greatly. If you use one sensor for an extended period of time then your trends will be good. But if you change sensors, your trend will have a step change.

Slow speed bearings

Slow speed bearings (below approximately 200 RPM) represent a monitoring challenge, but the good news is that these high frequency techniques may provide a workable solution. Although the rotation rate of the bearing may be slow, a high frequency shock pulse is still generated when a ball or roller contacts a defect. This may not generate enough energy to cause the bearing housing to ring and therefore one must be careful with normal enveloping techniques.

Ultrasound equipment is sensitive to low level shock pulses and friction and may provide diagnostic information on bearings rotating as slow as 15 RPM.

When monitoring low speed machines remember that vibration amplitudes are very low and may be difficult to monitor with standard accelerometers. High frequency vibration does not travel far, so the mounting and placement of the sensor is critical.

Also remember that we are probably not particularly interested in monitoring the machine for things like unbalance as at very low speeds even a large unbalance will generate very little force. So although the machine turns slowly, we may get the best information using high frequency techniques.

Stage Three Bearing Faults

The following points summarize the characteristics observed when a bearing is in Stage Three:

- More significant damage:
    - Minor damage through to more significant damage
    - Bearing can fail in many ways for many reasons
- Very high frequency vibration continues to increase in amplitude.
- Envelope (demodulation) spectrum will be effective
    - Filters must be set up correctly
- Classic patterns appear in spectrum:
    - Harmonics due to impacts
    - Modulation (sidebands) due to cyclical change in load

The spectrum is now very important. The acceleration spectrum and log velocity spectrum may have exhibited peaks at the bearing defect frequencies, but there is no doubt that the velocity spectrum will have vibration patterns that indicate the fault.

Figure 10-58

Outer race fault (inner race rotating)

If there is a spall on the outer race (probably right in the load zone), there will harmonics of BPFO present in the data. To begin with it is possible that the harmonics are weak, but as the fault develops the harmonics will get stronger, and will likely increase in amplitude as the frequency gets higher.

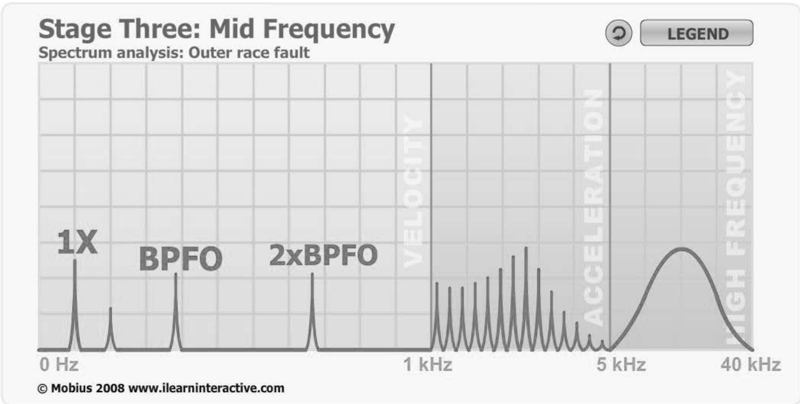


Figure 10-59

Outer race fault (outer race rotating)

If the outer race is rotating, then the pattern will change. There will still be harmonics; there always are harmonics when impacts occur. The forces in the impact will no longer be consistent. When the damaged area travels through the load zone, the impacts will be stronger compared with the impacts (and vibration level) when the damaged area is at the top of the bearing where it may be lightly loaded.

This rise and fall in the vibration amplitude generates sidebands in the spectrum. Because the outer race would be rotating at the turning speed, the sidebands would be spaced at 1X.

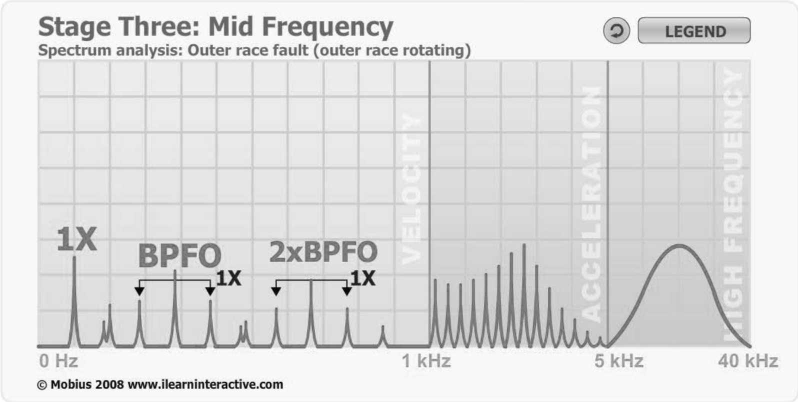


Figure 10-60

Inner race fault

If there is damage on the inner race on the bearing, we will observe three important characteristics:

1. Because there will be impacts, we will see harmonics of the BPFI frequency.

2. Because the damage is on the inner race and the vibration must travel through the balls to the outer race and on to the measurement location. Therefore the amplitude may be lower.

3. Because a spall travels around the bearing once per revolution, the impacts will not be equal in amplitude. So we will witness 1X sidebands. If the outer race is rotating then there will be no sidebands because the force behind each impact will be equal.

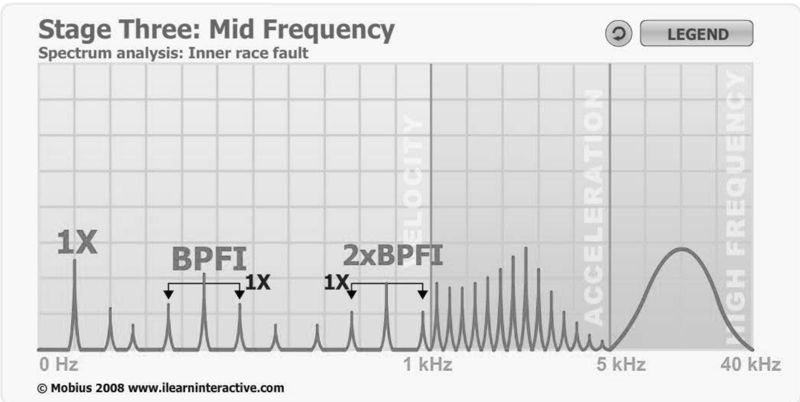


Figure 10-61

Ball or roller damage

If there is ball or roller damage, then we expect to see a peak at BSF with harmonics. It is likely that 2xBSF and harmonics will be stronger as the damaged area on the rolling element may make contact with the inner and outer race each rotation.

There will also be FT (cage rate) sidebands. As the balls move around the bearing they will go into the load zone of the bearing at the rate defined by the fundamental cage frequency, which will be less than half turning speed if the inner race is rotating, a little over 0.5X if the outer race is rotating.

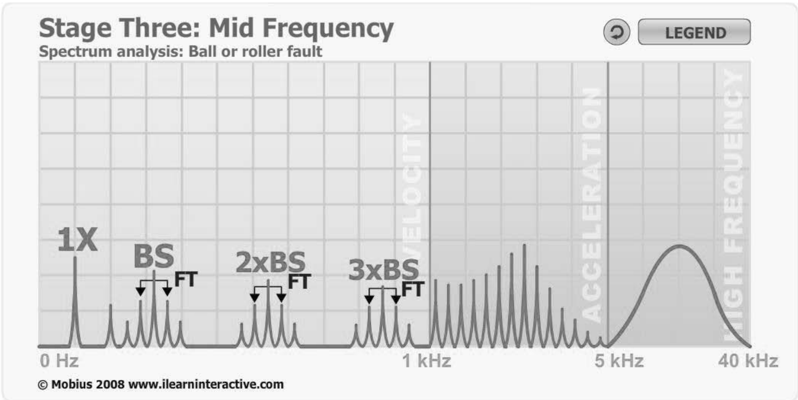


Figure 10-62

Knowing the forcing frequencies can be very helpful; especially if there is more than one bearing close to your measurement location. However knowing that the defect frequencies are non-synchronous, and that they generate harmonics, and that 1X sidebands indicate an inner race fault, and FT sidebands indicates a ball or roller fault, then we can determine with a lot of confidence that the bearing fault exists.

Overview of techniques

The effectiveness of each bearing detection technique is summarized below:

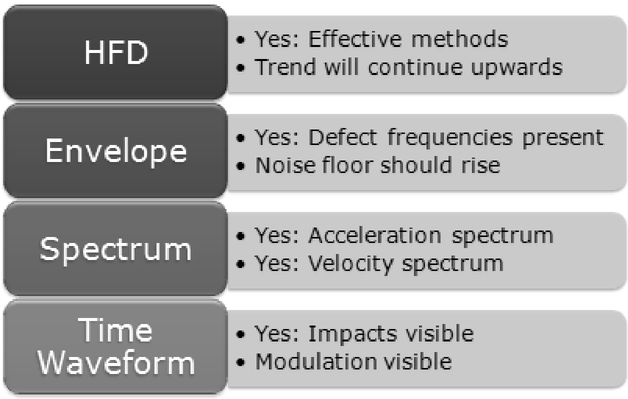


Figure 10-63

The following table summarizes the possible actions you could take at this stage:

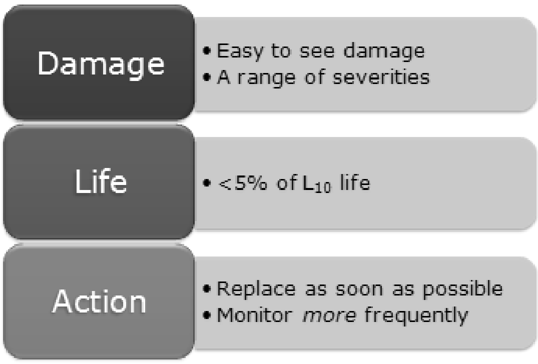


Figure 10-64

Spectrum analysis

Remember that bearing defect frequencies are non-synchronous. In order to quickly identify which peaks are synchronous and non-synchronous it is helpful to order normalize the data. It is also important to have enough resolution in the spectrum to separate bearing defect frequencies from shaft rate harmonics or other forcing frequencies, especially in multi shaft machines. If you routinely take one spectrum with a high Fmax (say 60x) then it is a good idea to zoom in and view the data from 0-10x or 0-20x so that you can clearly see the bearing tones.

Logarithmic graph scales

One of the tricks for quickly identifying bearing problems is to format the data in a way that accentuates them. Log and dB graph scales on the “Y” or vertical axis serve exactly this purpose. In brief, when a graph is presented in a Log or dB format, one is able to see very small peaks in the presence of very large peaks. In a machine, bearing tones may be much smaller in amplitude than 1X or a vane pass frequency. When looking at the graph in a linear scale, one might only see these dominant peaks and the bearing tones might be down in the noise.

With a Log or dB amplitude scale one can see two peaks in the same graph even if one peak is 1,000x higher in amplitude than the other! If you want to try to visualize this, consider trying to take a photograph of a person standing next to a very high tree. The problem is, you always have to have the top of the tree in the photo. If you stand far enough back to capture the top of the tree, the person will look like a little speck of dust at its base. This is your bearing tone compared to 1x. As mentioned, if you simply change the scale to Log or dB you will be able to clearly see the person and the entire tree in the image even if the tree is 1000x taller that the person.

Even if a bearing tone has a low absolute amplitude, if its amplitude doubles or triples then it indicates there is a serious problem – so it is very important to be able to see these tones even if they are quite small.

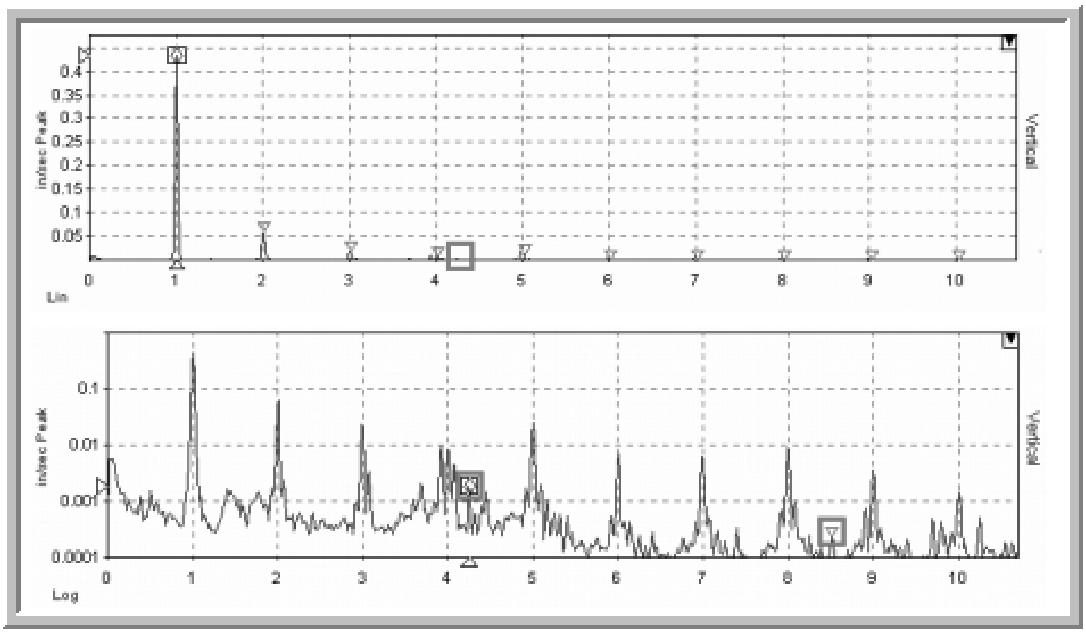


Figure 10-65

The two plots above are the same data, the bottom plot is in a Log format; the top one is linear. A bearing tone and harmonic are indicated by the red boxes – note that the tone is not visible at all in the top graph. The bottom graph also shows signs of looseness (1x harmonics) and a raised noise floor which is not visible in the top graph. Finally, also note that the data in both graphs are order normalized and there is plenty of resolution (and space) between the shaft rate harmonics so it is easy to find non synchronous tones.

Time waveform analysis

The time waveform will have clear signs of the defects. A waveform in units of acceleration will contain each of the impacts. Modulation will also be clearly visible in the data.

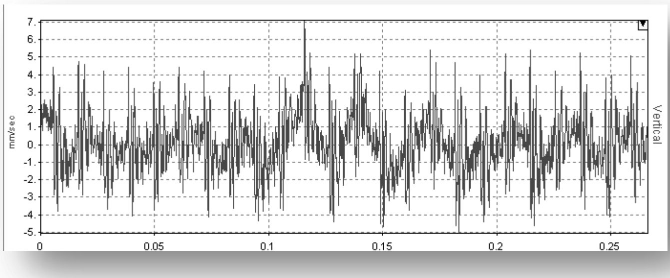


Figure 10-66

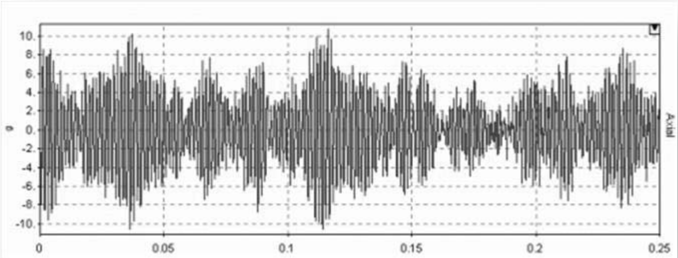


Figure 10-67

Time waveform analysis is particularly effective with low speed machines.

Stage Four Bearing Faults

The following points summarize the characteristics observed when a bearing is in Stage Four:

- Significant damage
    - Damage far more extensive
    - Damage in one component causes damage in other components
    - *Failure is imminent*
- Very high frequency vibration may trend *downwards*.
    - Smoothing of metal reduces sharp impacts
- Spectrum, time waveform and envelope spectrum analysis still effective – at first…

Figure 10-68

As we move into the stage four faults the vibration patterns will start to change dramatically; due in most part to the excessive wear. The vibration becomes less and less periodic – there are too many areas of damage. The vibration starts to become noisy which lifts up the noise floor – we call them "haystacks", but they are really just resonances.

As the fault develops further the defect frequency peaks will disappear altogether, and the spectrum will change to become more like the pattern expected with rotating looseness. This is because there has been so much metal loss that the clearances begin to grow.

Once you observe these changes you must take action; the bearing has less than 1% of its life remaining.

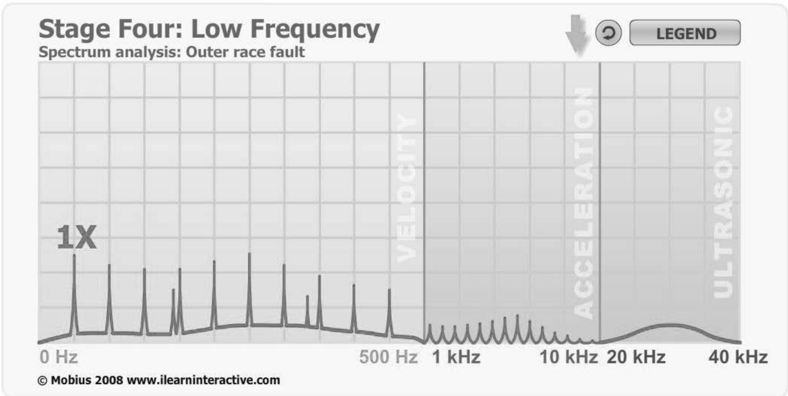


Figure 10-69

At this stage the high frequency techniques become far less effective. The shock pulse and gSE trends may dip. The spectrum will provide enough evidence of the fault condition. A person standing next to the machine could tell you that the bearing was damaged.

The effectiveness of each bearing detection technique is summarized below:

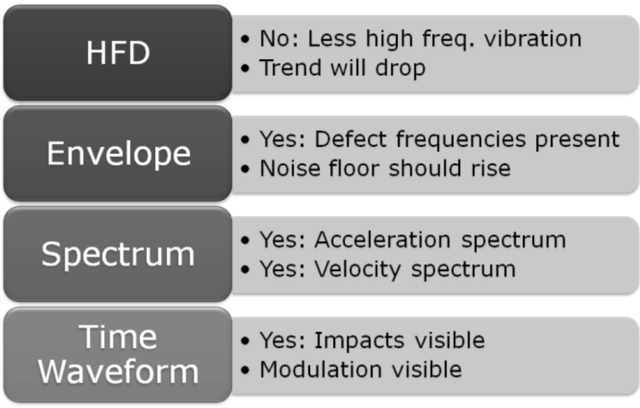


Figure 10-70

The following table summarizes the possible actions you could take at this stage:

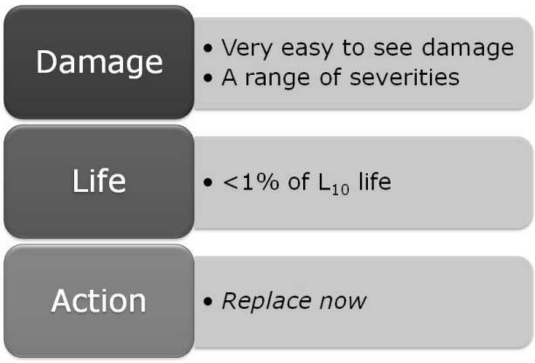


Figure 10-71

Optimizing your results

Now that we have nearly completed this section on detecting bearing defects, please remember that detecting defects is only part of the story. It is just as important to try to remove the root cause of bearing failure and extend the life of the bearings. This begins in the design phase, making sure one is using the correct machine and bearings for the job, and it continues into purchasing, shipping, handling, installation, acceptance testing, alignment, balancing, lubricating etc. In short, there is bearing wear and there is bearing care and we should be interested in both of them.

Case Study: Air Washer #1

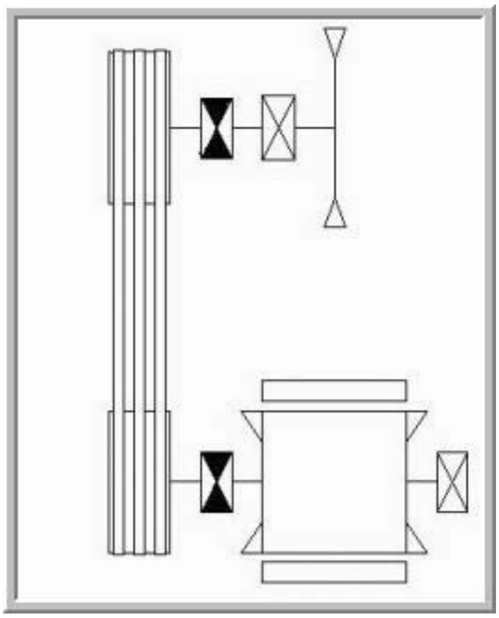


Figure 10-72 - Motor driving fan; 2:1 belt ratio

Speed = 1770 CPM; 1XM
40 HP motor
Overhung fan
16 blades on the fan; 16XF or 8XM
11 cooling fan blades on the motor
Motor sheave 6.5" (165 mm)
Fan shaft = 0.5XM; 885 CPM; 1XF
Fan sheave diameter 13" (330 mm)
Distance between centers 47" (1194 mm)
Belt rate 0.1636XM

When monitoring commenced on 4/5/99, the vertical and horizontal data did not show very much at all. The amplitude levels are quite low. The time waveform does show some high frequency content, suggesting the earliest signs of bearing wear. This is a belt driven machine, so there are peaks associated with the belt rate and the fan shaft. The data is from the motor.

The axial direction does show a small non-synchronous peak at 3.06 orders. The high 1x indicates the sheaves may be misaligned.

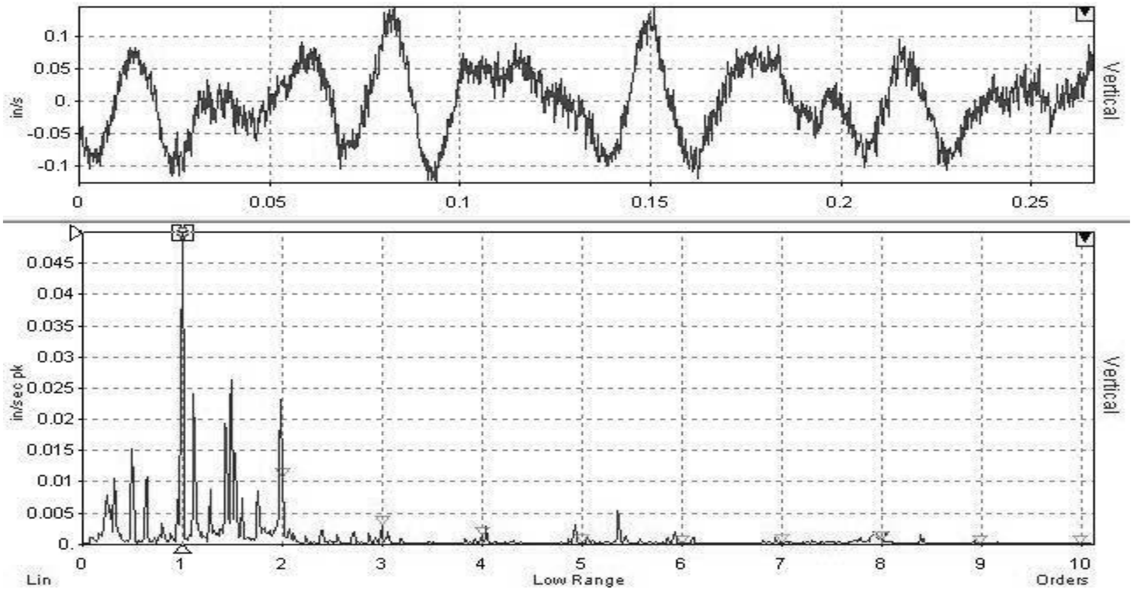


Figure 10-73 - Vertical and Horizontal show very little signs of bearing frequencies.

May 3 shows early signs of bearing wear, the levels were increasing, and the time waveform showed even more signs of bearing wear. It is difficult to see in a linear plot; however, in log view a peak at 3.1X with harmonics is visible. The bearing is at Stage Two now.

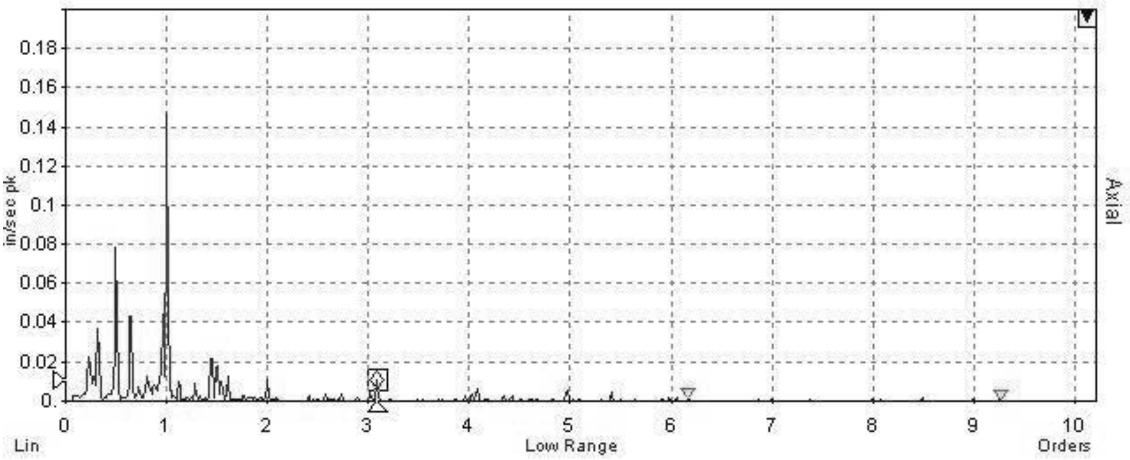


Figure 10-74 - Axial shows peak at 3.06 orders.

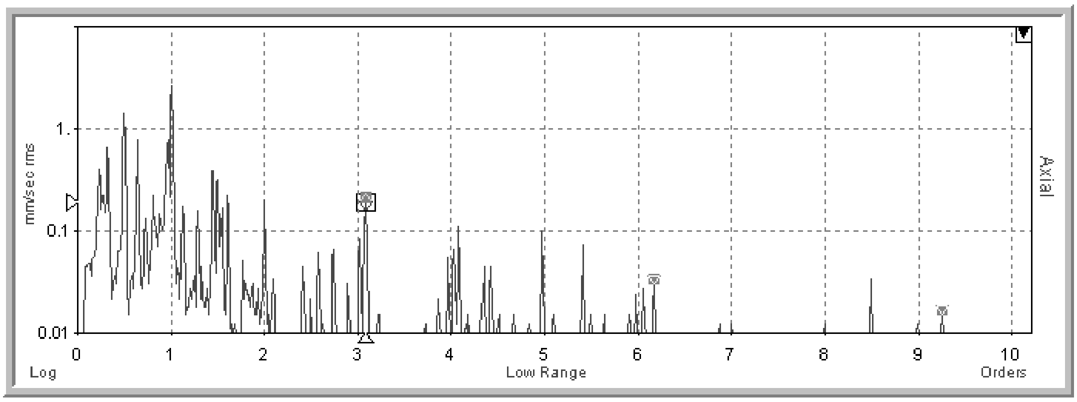


Figure 10-75 - The peaks at 3.06 and its harmonics are more visible in Log scale.

On 6/1/99, the signs of bearing wear are becoming quite strong. The time waveform is a classic. The impacts are occurring every 3.1X of running speed.

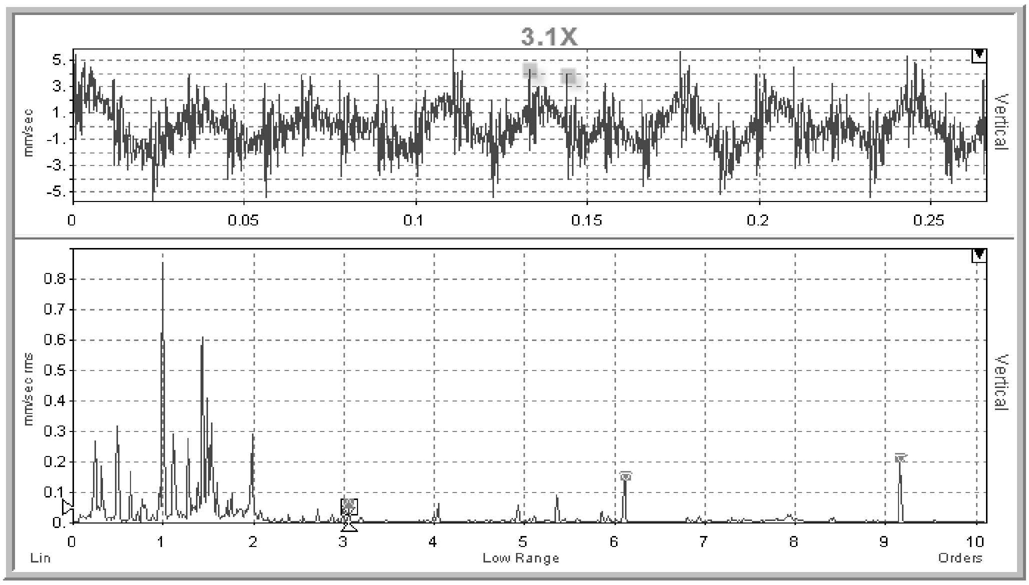


Figure 10-76 - The classic waveform shows the impacting at every 3.1 orders.

The high range of the spectrum shows there are harmonics of the bearing frequency to greater than 50 orders.

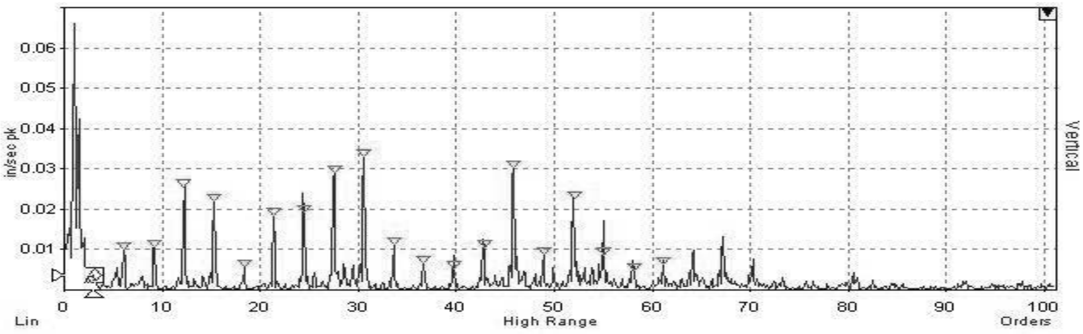


Figure 10-77 - High range shows bearing harmonics to greater than 50x.

Where are the sidebands?

Looking closely at the data, even in log, there are no strong sidebands. How can this be? Recall that sidebands occur if there is a ball fault (sidebands of FT) or an inner race fault (sidebands of 1X). Given that the bearing tone is low in frequency (BPI = 0.6 x # balls, BPO = 0.4 x # balls) it would be suggested that the fault is on the outer race. So there are harmonics, just no sidebands.

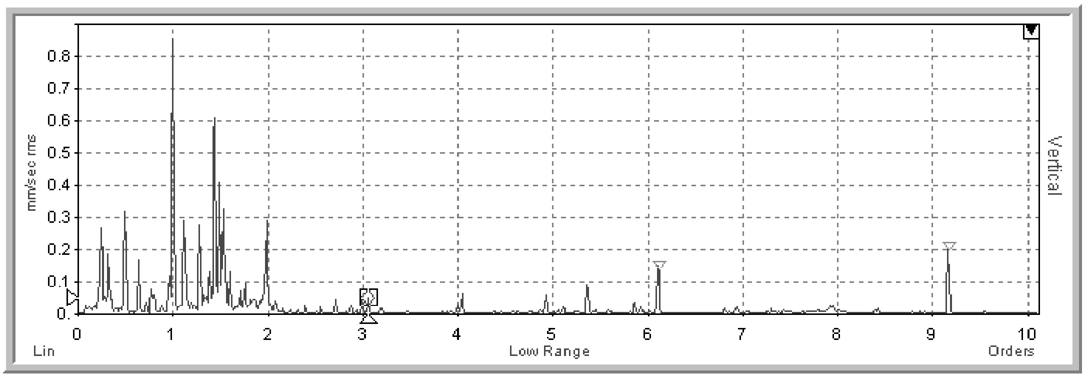


Figure 10-78 - There are no sidebands. This is an outer race fault which typically does not vary in amplitude like inner race frequencies

Notice that we have been able to determine, with some confidence, the nature of the fault without even knowing what bearing is installed in the machine.

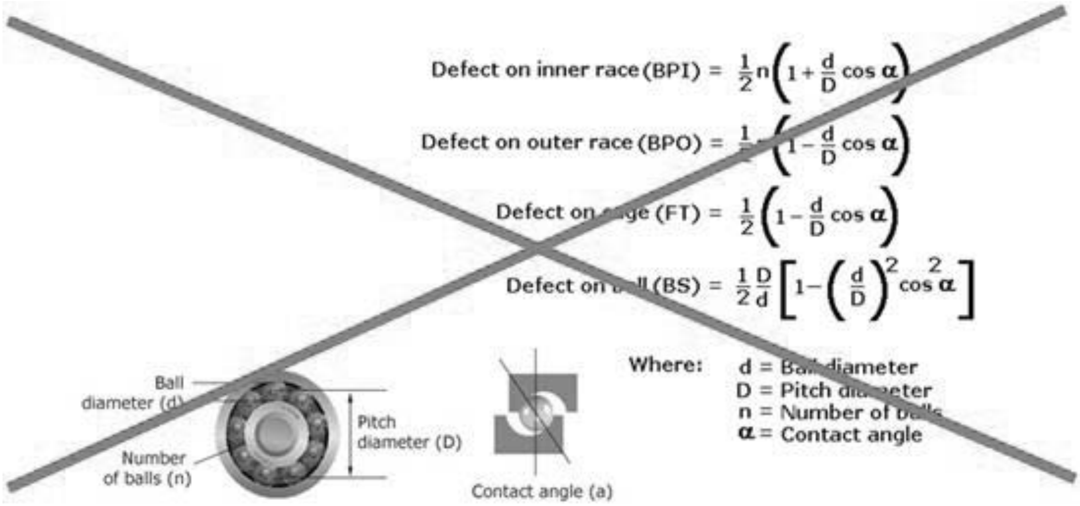


Figure 10-79 - Even though the bearing was not known, we were still able to determine the nature of the fault

On 5th October, 1999, the harmonic levels increase, haystacks appear, and the noise floor rises. This is the data from the high range, vertical direction. Fig. 15-32 shows the high range spectrum. Notice the bearing harmonics to more than 60 times turning speed.

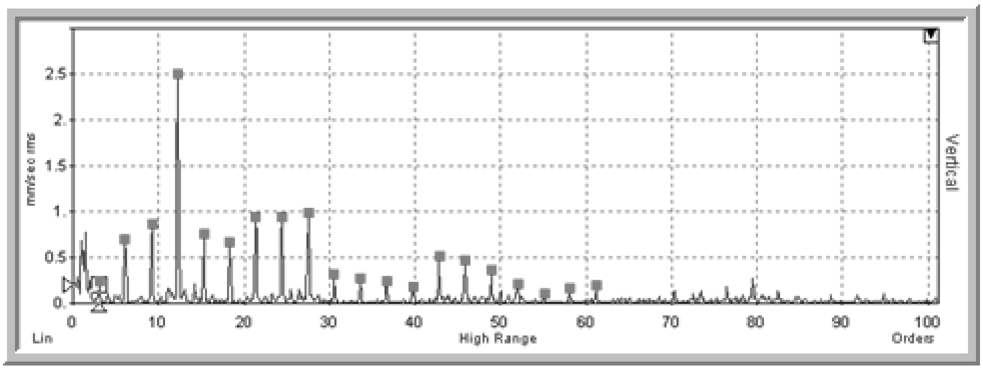


Figure 10-80 - High range showing bearing frequency harmonics to greater than 60x.

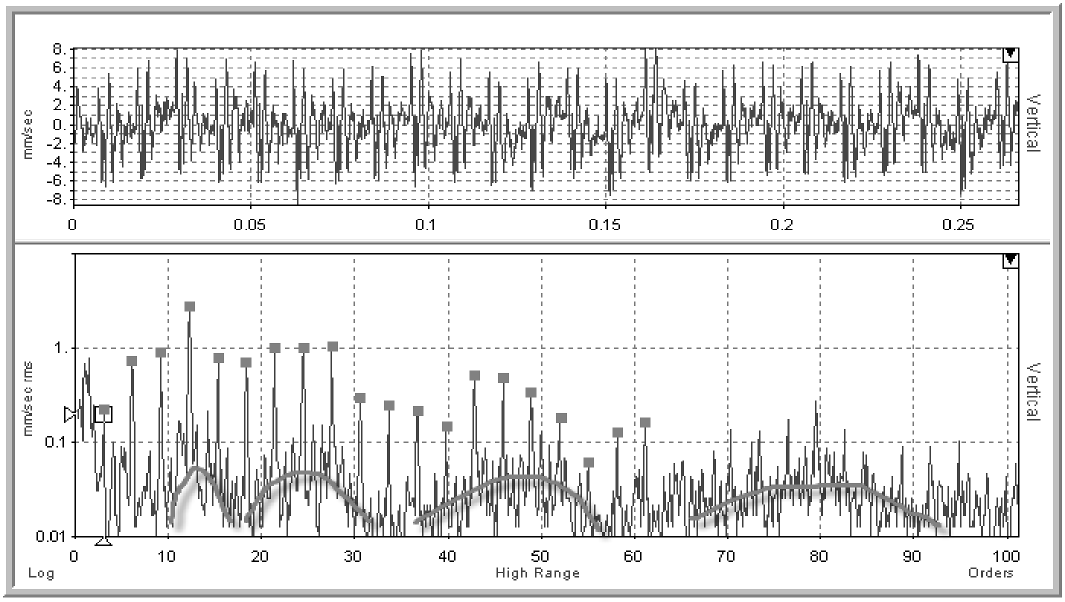


Figure 10-81 - by October, haystacks are evident

The 1/12/00 data looks very different! The time waveform does not have the same periodic spikes, the spectrum does not have the harmonics, but the noise floor is considerably higher.

The bearing is at Stage Four. The bearing is almost dead - in fact, the machine failed catastrophically just two days later!

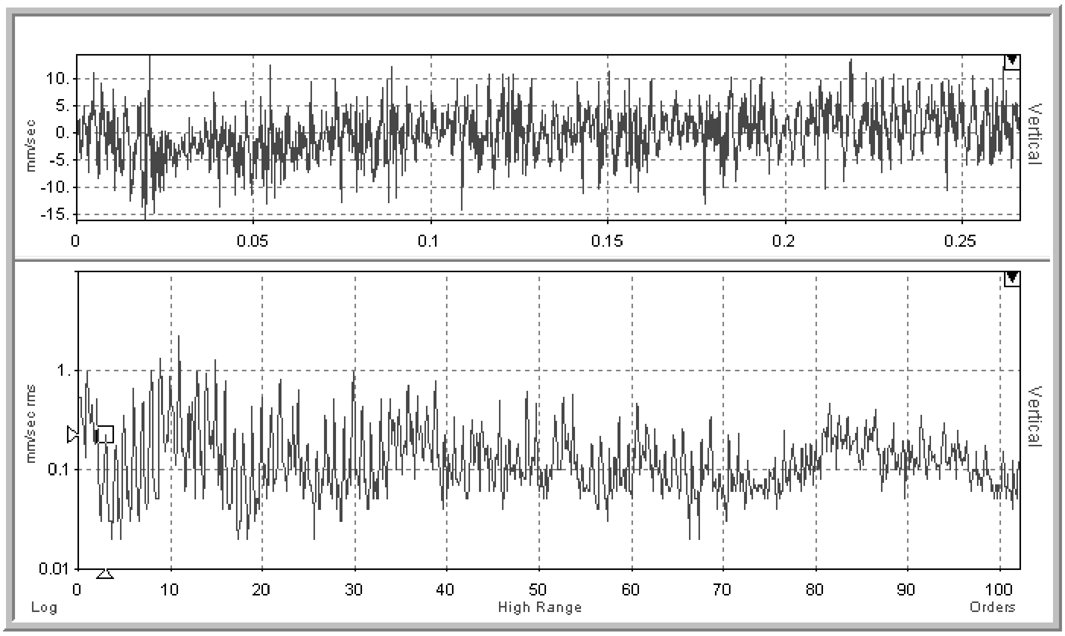


Figure 10-82 - Stage Four - the waveform is different, the spectrum floor is raised. The machine failed 2 days later.

Chapter 11

Journal Bearing Analysis

Objectives:

- What are journal bearings?
- Non-contact eddy-current probes
- Transducers, signals and conventions
- Keyphasors
- Slow roll subtraction
- Direct versus filtered orbits
- Orbit diagrams and centerline diagrams
- Fault diagnosis

Acknowledgement: many images have come from “Rotor Dynamic Measurement Techniques”, Product Note: Agilent 35670A-1 by John Jenson and “What Are Shaft Orbits Anyway?” by Mark A. Jordan, Sr. Rotating Equipment Vibration Engineer, Industrial Machinery Diagnostics, LLC

Diagnosing faults in machines with journal bearings

In this chapter we will learn about how to monitor journal bearings; also known as “fluid-film bearings”, “plain bearings”, “white metal bearings”, “sleeve bearings”, and possibly by other names. Journal bearings are used in a wide variety of rotating machinery such as turbines, generators and compressors. These machines are typically critical to the plant’s operation; therefore detecting conditions such as rubs, instabilities, unbalance and misalignment are very important.

While machines such as motors, pumps and smaller compressors may also have journal bearings, in this section we are focusing on machines that have non-contact eddy current probes installed (also known as Proximitor™ probes).

Figure 11-1

Accelerometers may be installed on journal bearings for one of two reasons:

In order to monitor the vibration on smaller machines with journal bearings when the cost of installing the displacement probes may not be justifiable. In this case many of the typical rules of vibration analysis can be used to detect unbalance, misalignment and other fault conditions. The fluid film between the shaft and the bearing does attenuate the vibration, so using an accelerometer is not as effective. Only displacement probes can see inside the bearing and report on the position of the shaft relative to the available clearance, and display the orbit and precession of the shaft.

Accelerometers may be installed on journal bearings *in addition to* displacement probes. The displacement probe measures the movement between the bearing casing and the shaft, however in certain situations, the bearing casing may be moving relative to ground. Adding an accelerometer will provide an indication of how the casing is moving, and therefore how the shaft is moving relative to ground.

Journal bearings have been used for many years; in fact they are the first type of bearing used in order to facilitate shaft rotation with reduced friction.

The basic design involves a cylinder of metal, coated with a soft metal Babbitt lining that encloses the shaft, with a film of lubricant to reduce friction. One of the greatest challenges has been to support rotation at a range of speeds without the fluid-film becoming unstable. If the "oil whirl" condition exists the vibration levels can become very high and damage can occur. If the "oil whip" condition exists then the entire system will become unstable very quickly, potentially resulting in catastrophic failure.

Over the years journal bearings have evolved in design in order to deal with fluid film stability. Examples of bearing designs include plain bearing, lemon/elliptical bore, pressure dam, offset, and tilting pad. It is beyond the scope of this course to discuss these designs any further.

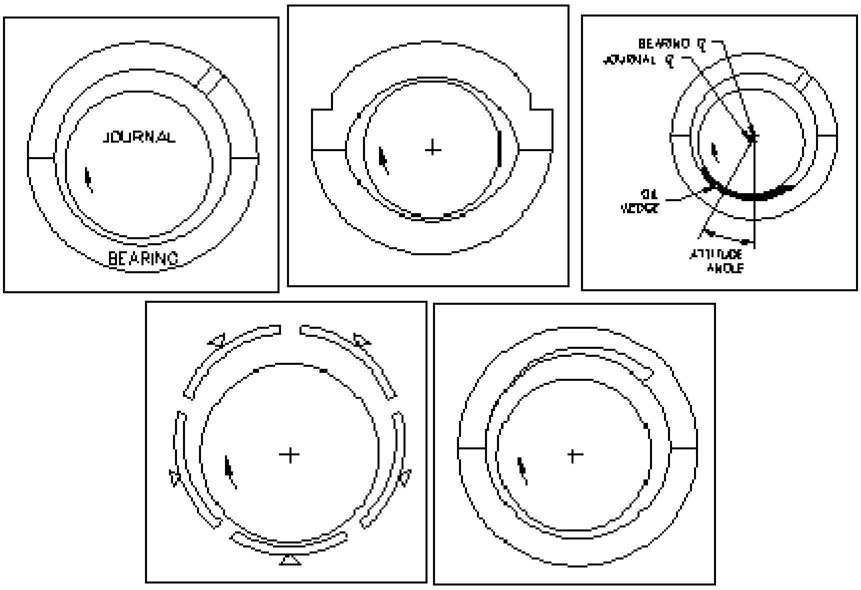


Figure 11-2

Non-contact eddy current probes

Non-contact eddy current probes are used to measure the displacement (distance) between the tip of the probe and the shaft. The probe is permanently installed in the bearing housing; therefore the displacement measured is an indication of the movement of the shaft within the bearing.

Non-contact eddy current probes utilize an external, high frequency oscillator to induce eddy currents in the surface of the shaft. As the name suggests, the probe does not contact the shaft. The electromagnet field induces currents in the surface of the shaft which in turn generate an electromagnet field that is detected by the probe. The signal is strongest when the shaft is closest to the tip of the probe. When operated in the transducer's linear range, the strength of the signal is directly proportional to the distance between the shaft and the tip of the probe.

Figure 11-3

Here is a sample product, showing the signal conditioning unit (oscillator and demodulator) and the probe.

Figure 11-4

Signals available from displacement probes

The signal from the probe is "demodulated"; the variations in the high frequencies amplitude from the probe (used to induce the eddy currents) are converted into lower frequency signals that tell us about the dynamic movement of the shaft (the A.C. component) and the average gap between the probe tip and the shaft (the D.C. component).

The A.C. signal provides valuable information about the dynamic movement of the shaft; i.e. how it is vibrating. We can use this signal like any other A.C. signal; we can display it as a time waveform, and we can perform an FFT to display a spectrum. As we will soon learn, we can use two probes, mounted at 90° to each other, to construct an orbit display that shows how the center shaft moves as it rotates.

The D.C. signal is a measure of the average shaft position within the bearing. For example, while the A.C. signal may tell us that the shaft is rotating with a circular motion, the D.C. signal tells us where the shaft is rotating; at the top of the bearing, at the bottom of the bearing, to the left or to the right. We can learn a great deal about the state of the machine (and any fault conditions) via the position of the shaft within the bearing and the motion path of the shaft. As we will learn, we can do this with the centerline plot and orbit plot.

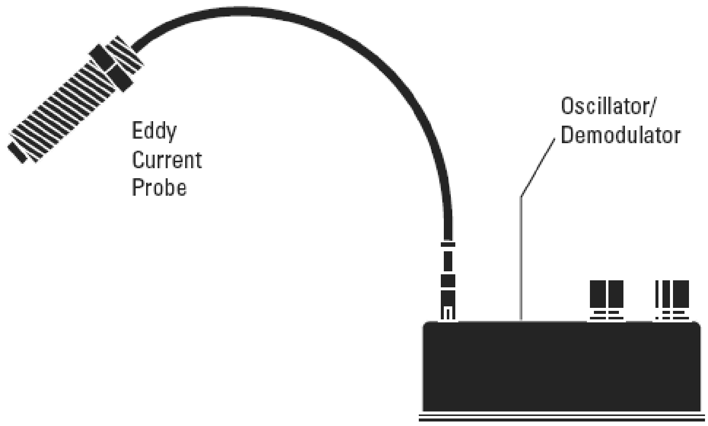


Figure 11-5

Displacement probe sensitivity

The displacement probe will have a documented "sensitivity" telling us the voltage output measured for a given displacement. For example, the sensitivity may be 200 mV/mil (or 7.87 V/mm), which means that if the shaft were to move 1 mil (one thousandth of an inch) we would see a change in voltage of 200 mV (or 0.2 volts). If the shaft were to move 5 mil pk-pk, we will measure 1 Volt pk-pk.

The transducer will have a documented "linear range". Within that range of displacement the voltage output will be directly proportional to the displacement – for example, if you were to double the displacement the voltage would double.

The following is a sample from GE Power (ex. Bently Nevada):

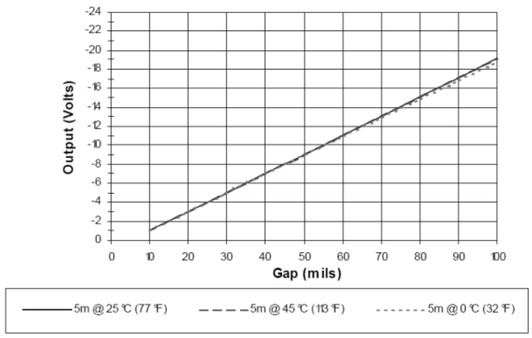


Figure 1: Typical 3300 XL 8 mm 5m or 1m System over API 670 Testing Range

Figure 11-6

The displacement transducer should be installed so that its "set point" is in the middle of the linear range of the transducer. If this were not done, the shaft may move outside the linear range of the transducer and the voltage measured would not provide an accurate measurement of displacement.

Two displacement probes are typically installed on the bearing. It is important that both transducers have the same sensitivity.

Displacement probe polarity

It is important to note that the polarity of the displacement transducer is such that as the distance between the probe tip and the shaft is *decreased*, the voltage will *increase*. This is because the strength of the electromagnetic field will be stronger, and thus the eddy currents will be stronger.

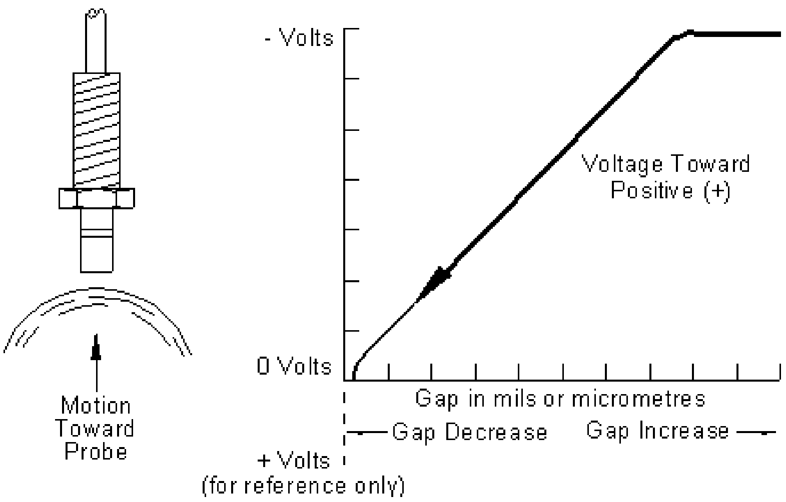


Figure 11-7

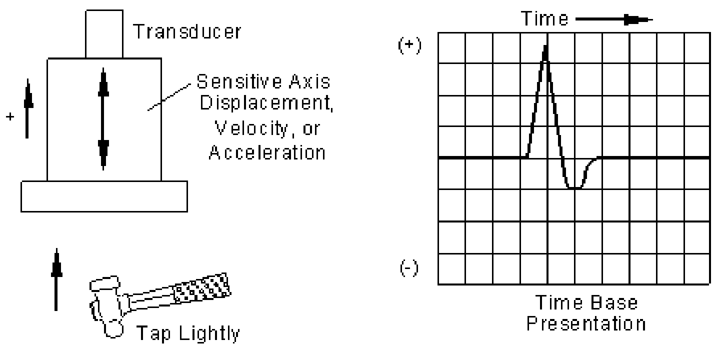


Figure 11-8

Produced by Sumico Technologies

Displacement probe conventions

In order to easily generate the orbit display, and determine the position of the shaft in two dimensions, it is necessary to mount the displacement probes orthogonally – i.e. 90° apart.

Ideally one would be mounted vertically and the other horizontally; however this is not always possible. The construction of the bearing may require the probes to be mounted 45° to the right and left of the top of the bearing, or any other position for that matter.

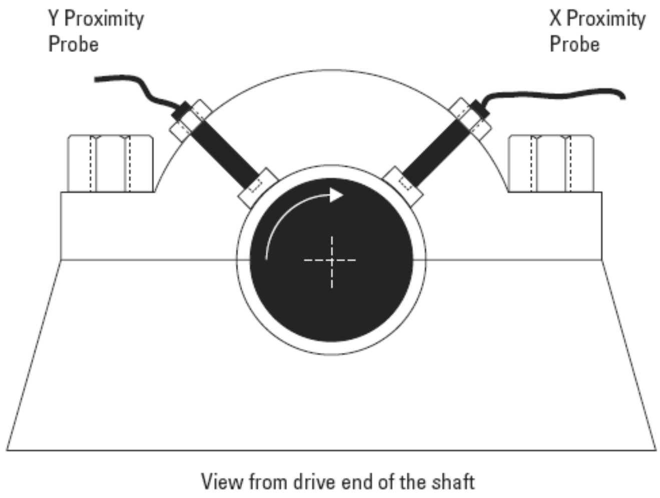


Figure 11-9

The vertical probe is called the "Y" probe, and the horizontal probe is called the "X" probe; in accordance with the orientation of most graphs the y-axis is vertical and the x-axis is horizontal. However when the probes are not mounted vertically and horizontally, it is necessary to utilize a convention to determine which is the "X" and "Y" probe.

The following illustration shows the "X" and "Y" probes for a range of possible installation positions. Note that the convention is based on observations from the driver (e.g. the turbine) component.

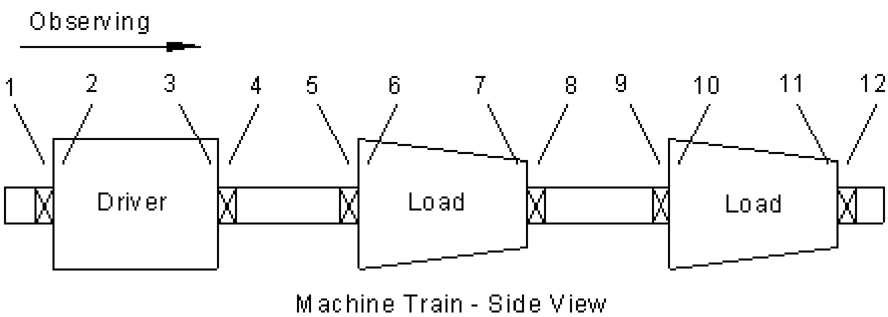


Figure 11-10

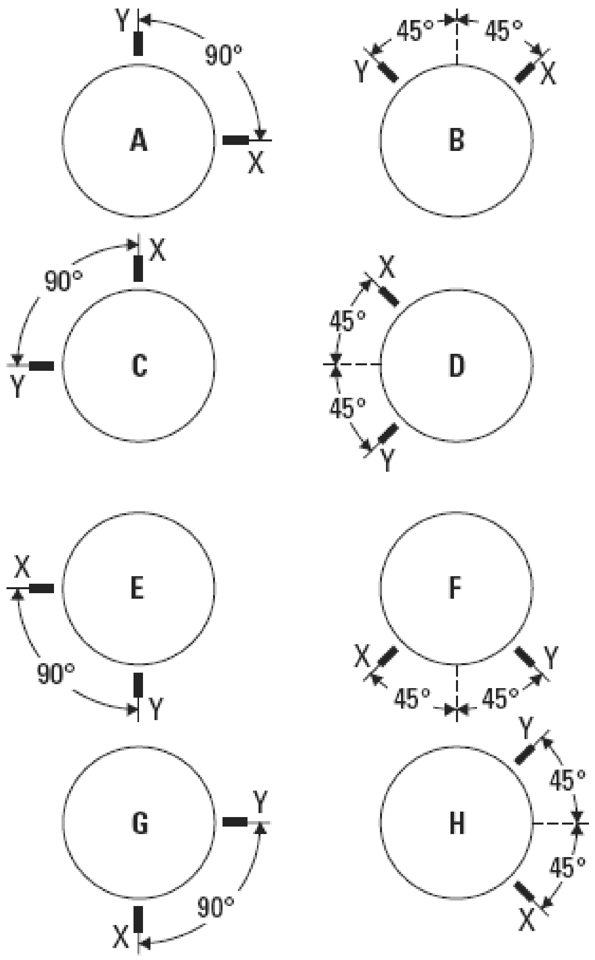


Figure 11-11

* Figures courtesy of Machinery Malfunction Diagnosis and Correction, Vibration Analysis and Troubleshooting for the Process Industries, Robert Eisenmann Sr. P.E. & Robert Eisenmann Jr. PTR Prentice Hall, Nov. 97. , chapter 6 and from "Standardized Rules for Measurement on Rotating Machinery", Bently Nevada

Keyphasor: once-per-revolution reference

In order to perform orbit analysis (and balancing), it is necessary to have a once-per-revolution timing reference. We need this reference so that the measurement (and display) electronics can determine when the shaft has completed one complete revolution, and so that we can determine how the shaft is moving in relation to the reference. For this reason orbit displays will typically indicate the position of the timing reference via a glowing "dot' on the orbit shape, as discussed in greater detail later in this chapter.

The once-per-revolution signal and be achieved in one of two ways. We can use reflective tape (or some other type of reflective or contrasting surface) and a laser or photo-electric optical tachometer. While this method is common when monitoring "balance of plant" machines, it is less common in this application.

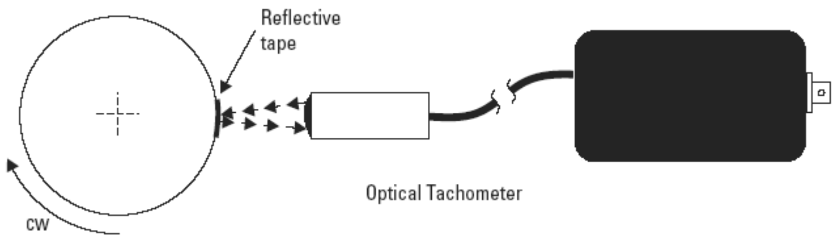


Figure 11-12

More commonly we can use a third non-contact displacement probe which is in line with the keyway. The Bently Nevada company calls this a "Keyphasor™".

If the key protrudes then we will measure a rising voltage once per revolution.

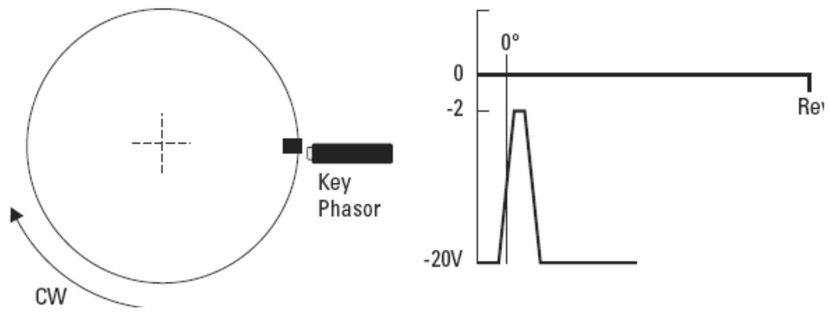


Figure 11-13

If the keyway is recessed, then we will measure a drop in voltage once per revolution.

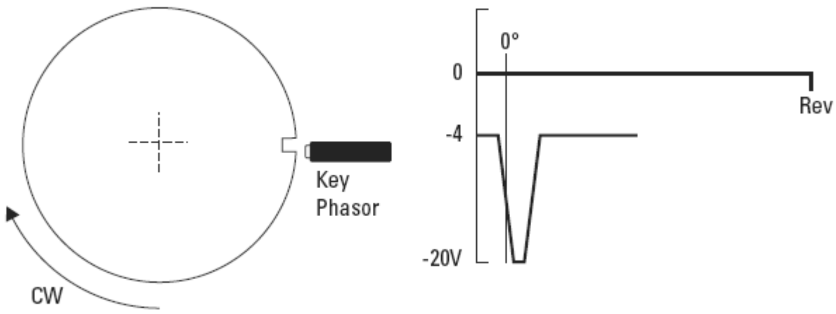


Figure 11-14

Vibration analysis of journal bearing machines

It is true that an accelerometer mounted on the case of the bearing will provide a useable signal that can be used to determine if the vibration has changed substantially. And it is true that the acceleration signal can be integrated twice to displacement; however it must be understood that the displacement readings would not be the same as the readings from the displacement probe. The accelerometer measures the vibration on the casing of the bearing. The casing of the bearing will vibrate as a result of the forces within the bearing, however due to the dampening effect of the fluid film, the vibration will be attenuated, and the low-frequency response will be limited.

The acceleration measurement can be used to aid in the assessment of conditions when displacement probes are not installed. And when displacement probes are installed the acceleration measurement can still provide useful data about how the casing is moving relative to the earth. Without a doubt, even though they are expensive to install, the non-contact eddy current displacement probes are by far the most useful transducer for this application. As already stated, displacement probes provide information about the absolute position of the shaft within the bearing; how the shaft is moving during each rotation and the physical (average) position of the center of the shaft.

Understanding the displacement readings

We must first be crystal clear about the information provided by the displacement probes. Visualizing the movement of the shaft within the bearing can be confusing. The displacement probes provide a measure of the displacement between the tip of each probe and the edge of the shaft. We are measuring a very small gap. There is an easier way to visualize what we are trying to measure.

Shaft centerline analysis using the D.C. "gap" voltage

First, imagine a point in the center of the shaft. At rest (i.e. the machine not operating), the shaft will make contact with the bottom of the bearing – this is the lowest possible position for the shaft, and thus it is the lowest point that the center of the shaft can go. If we were to measure the voltage from the two displacement probes we can utilize this reading as the reference point for the bottom of the bearing.

If we were now able to roll the shaft around the inner dimensions of the bearing so that it always made contact with the bearing, and continually read the D.C. gap voltage, we would be able to draw out a circle (assume a circular bearing profile) which represents the limits of movement for the shaft. Even though our readings are between the probe tip and the surface of the shaft, we can extrapolate to the center of the shaft. We now have a circular plot that shows the limits of movement of the center of the shaft.

Of course, we cannot perform that test in reality, however knowing the clearances of the bearing we can still create the same graph. This circle is the same circle that is typically shown on centerline diagrams. Therefore, from the D.C. gap readings we can determine where, in relation to the limits of the bearing, the shaft is located. We can tell how it has moved whilst running up to speed, and when we add the orbit information, we can determine if the shaft is rubbing the inside of the bearing (or threatening to do so).

Therefore centerline plots represent the movement of the center of the shaft (the average center, ignoring the dynamic movement), and the orbit plot shows the dynamic movement of the center of the shaft.

Performing the centerline test

When the machine is at rest or it is turning over slowly (less that 100 rpm), the reference reading is taken to represent the bottom of the bearing. As the speed of the machine increases, subsequent readings are taken from the two probes and from the tachometer (Keyphasor) in order to determine the speed. The changing position of the centerline of the shaft is plotted and the RPM is annotated on the plot.

In normal circumstances, the shaft should lift upwards and move to the left.

Eccentricity ratio 'ε'

The position of the shaft can be represented by a number called the "shaft eccentricity ratio". It is a ratio that indicates where the shaft is located relative to the center of the bearing. A value of '0' indicates that the center of the shaft is located at the center of the bearing. A value of '1' indicates that the center of the shaft is at one of the extreme limits of the bearing – i.e. it is touching the surface of the bearing.

A trend of decreasing eccentricity ratio ($\varepsilon \rightarrow 0$) indicates a potential stability problem. Conversely, an increasing eccentricity ratio ($\varepsilon \rightarrow 1$) suggests the rotor is approaching the constraints of the bearing wall.

Centerline diagrams are used to monitor the position of the center of the shaft as the machine runs up to speed, whilst it is running at speed and while it runs back down from full speed when it is shut down. We can learn a great deal about the machine by watching the path taken during run-up.

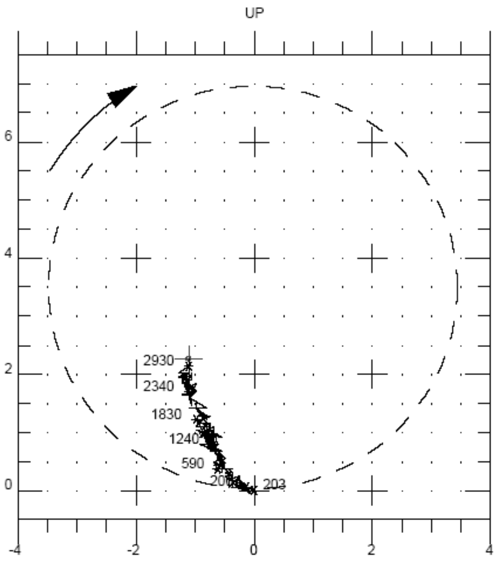


Figure 11-15

Analyzing the X and Y probe waveforms

We can simply view the time waveforms from each of the probes and perform a comparison. The waveform and spectrum do provide useful information.

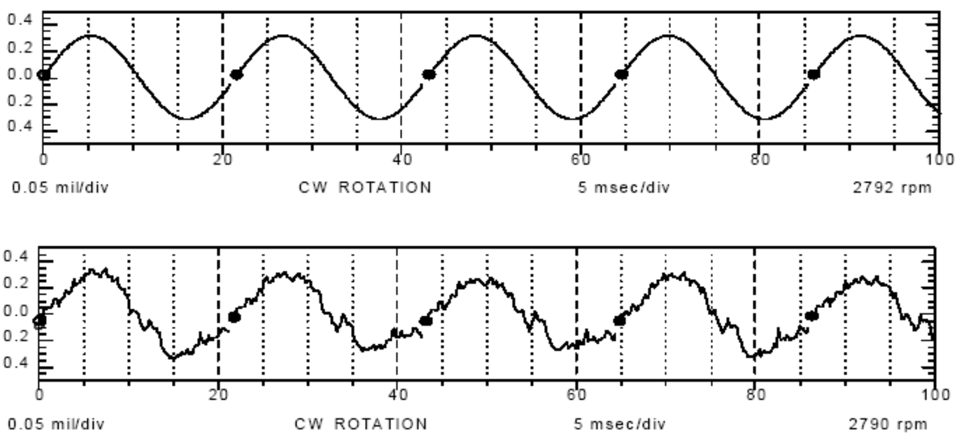


Figure 11-16

Orbit plots, or Lissajous figures

Orbit plots (also referred to as "Lissajous figures") provide an indication of the dynamic movement of the center of the shaft. We can learn a great deal about the operation of the machine (and condition of the shaft) from the shape of the orbit, from changes to the orbit from one moment to the next and from changes to the orbit over a longer period of time.

In summary, the orbit graph is generated by plotting the "Y" A.C. displacement signal against the "X" A.C. signal, and the position of the shaft reference (e.g. the keyway) is represented by a glowing dot.

For example, if the "X" and "Y" A.C. signals were equal in amplitude, the orbit would be circular in shape (because the shaft is moving an equal amount in the horizontal and vertical directions as it rotates.

That raises a very important point. If the "Y" probe is mounted at the top of the bearing (12:00 o'clock), and the "X" probe is mounted on the side of the bearing (3:00 o'clock), then the "Y" signal does represent the vertical movement and the "X" probe represents the horizontal movement. However, if the probes are mounted in different positions, the monitoring system (or software) must adjust the readings in order to provide a true position of vertical and horizontal movement; which is needed in order to relate the orbit shape to physical forces (e.g. preload).

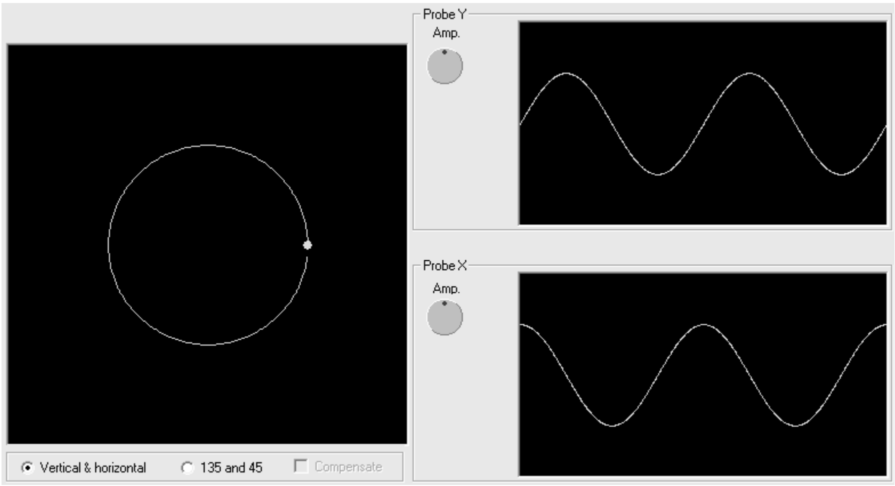


Figure 11-17

Using oscilloscopes

Orbit plots were originally generated using oscilloscopes. The "Y" probe was connected to the "Y" channel of the oscilloscope (or channel '1'), and the "X" probe was connected to the "X" channel of the oscilloscope (or channel '2'). The oscilloscope was then put into the orbit "Lissajous figure" mode where the "X" channel was graphed against the "Y" channel, and the result was the orbit display.

If available, the tachometer signal was connected to the "Z" channel, often called the "intensity input". The once-per-revolution pulse of voltage would cause the intensity (brightness) of the screen to increase, resulting in a bright dot on the orbit (Lissajous figure).

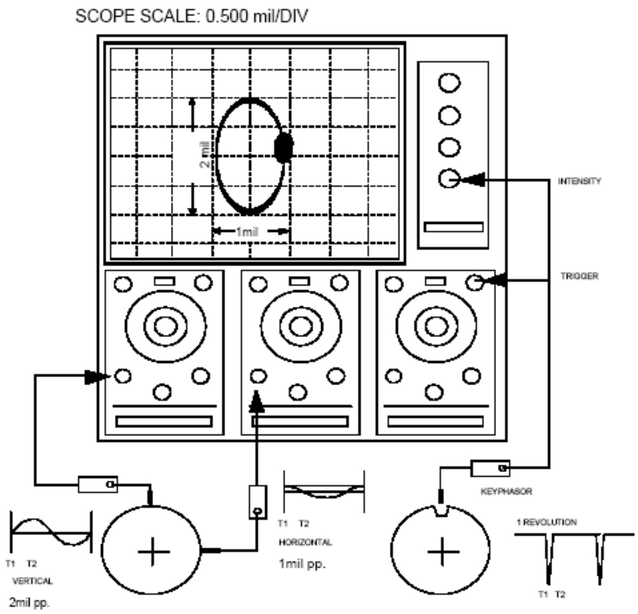


Figure 11-18

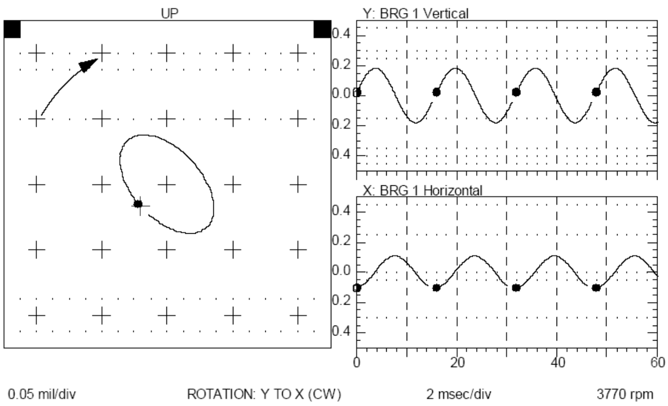


Figure 11-19

Monitoring systems

The widespread use of orbit diagrams (and centerline plots) was pioneered by the Bently Nevada company (now owned by General Electric). Bently Nevada conducted a great deal of research into the interpretation of orbit shapes and the shape of centerline plots. Bently Nevada also researched the dynamics of flexible turbine shafts. The Category IV course is designed to provide a much greater focus on rotor dynamics, orbits, centerline diagrams and more.

A number of manufactures now provide commercial products that capture data from displacement probes, and provide the software necessary to study the orbit plots and centerline diagrams. The primary focus of these systems is the protection of the machine; monitoring the displacement readings so that the machine can be shutdown (or "tripped") if predefined limits are exceeded.

Figure 11-20

Direct and filtered signals

The vibration signal contains all the vibration (dynamic changes in displacement) sensed by the displacement probes. As with vibration measured from accelerometers, it may contain vibration from a range of frequencies; vibration related to turning speed (1X), vibration from orders of turning speed (2X, 3X, and so on), non-synchronous vibration and noise.

When studying vibration from displacement probes we can learn a great deal about the machine by focusing on the 1X vibration, and we can learn from the orders/multiples of turning speed. For this reason, most systems will provide two sources of vibration; direct and filtered.

The direct vibration is the raw vibration from the displacement probes. It contains all of the vibration information. In some applications this data will have detail that can complicate the analysis process.

The filtered vibration is taken from the output of a tracking filter. The tracking filter will monitor the tachometer once-per-revolution signal and apply a band-pass filter so that only the vibration at 1X remains. The filtering system may also be capable of extracting 2X and 3X vibration – systems vary and you should check the capability of your system.

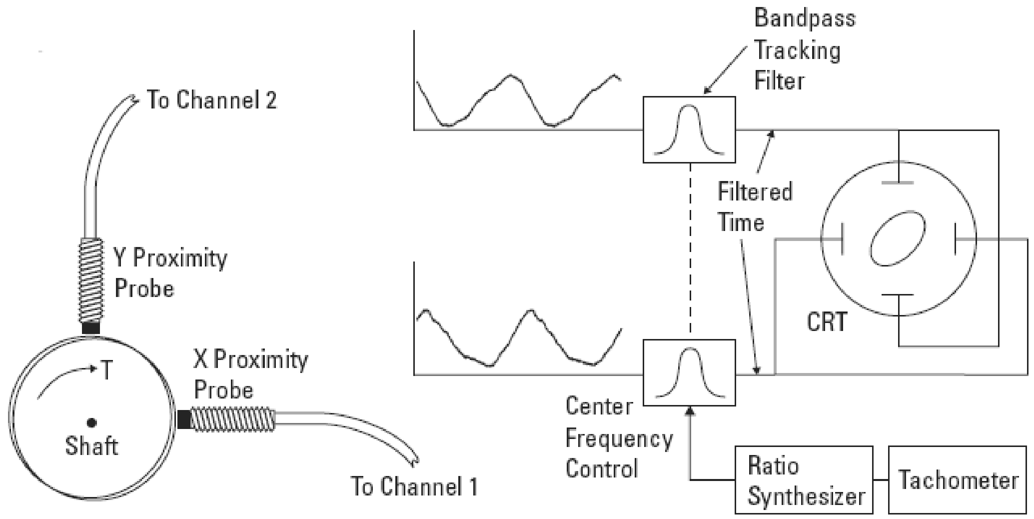


Figure 11-21

The following orbit displays were created from the same pair of displacement probes. The first set of data is the direct unfiltered data. The second set of data is filtered 1X vibration.

The shape of the orbit is similar in both cases, but the filtered orbit is much cleaner. The orbit is displayed as a single rotation of the shaft. (The orbit may actually be an average of the data from multiple rotations of the shaft, or it may actually have multiple orbits overlaid, or it may be a live display showing one orbit at a time.)

The live orbit display is also very important. As we will soon learn, the position of the tach reference (the Keyphasor dot) may move with time.

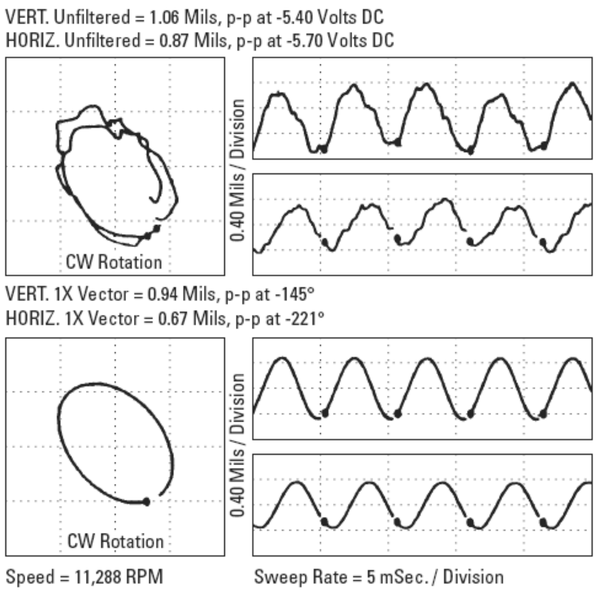


Figure 11-22

Slow roll or "glitch" compensation

Earlier we described how non-contact eddy current probes function; by inducing eddy currents into the surface of the shaft and then reading the electromagnetic radiation to gauge the distance between the probe tip and the shaft. The signal strength is greatly influenced by the displacement, however the signal is also affected by scratches in the surface of the shaft, and by flat-spots or any other imperfections.

The signal from the displacement probe will therefore contain voltage variations that are related to these "glitches". The orbit display is supposed to represent the movement of the center of the shaft, therefore we need to remove the "noise" generated by the glitches.

The noise is removed by first recording the vibration while the shaft is turning over very slowly; at a speed where there can be no dynamic forces and no movement due to unbalance, misalignment, etc. Using the tachometer signal as a reference, we record the position of the glitches for one rotation of the shaft. When the shaft is rotating at higher speeds the vibration we record is a combination of both the glitch data and the dynamic data (i.e. the data that represents the movement of the center of the shaft). If we subtract the recorded glitch data (or slow roll data) from the higher speed data, what remains is a record of the movement of the shaft (within the limitations of the process described).

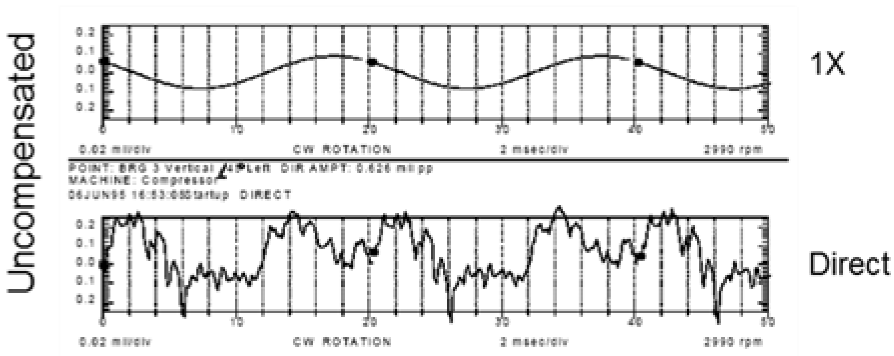


Figure 11-23

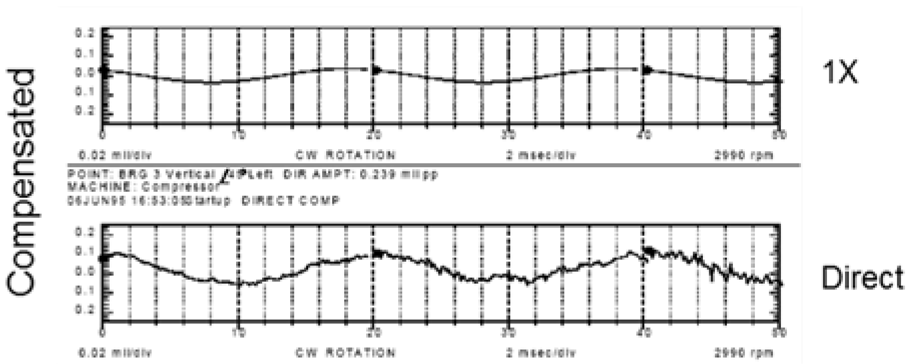


Figure 11-24

Using portable data collectors or analyzers to perform orbit analysis

Although there are protection systems and diagnostic systems available on the market that have been designed to specifically address the functions described thus far, it is also possible to perform orbit analysis with most two channel portable data collection systems (or analyzers). We will now investigate how these systems can be utilized to perform orbit analysis.

It should be said that we can create orbits from accelerometer data as well as from displacement probes. We are often interested in how a point on a structure is physically moving. For example, we know that if we compare vertical with horizontal data we expect a circular motion if the machine is unbalanced. We can check for a phase difference of 90 degrees, or we can use an orbit display and look for a circular plot.

Analyzers will typically give you the option to use direct data or filtered data. The analyzer would utilize a tachometer input in order to extract the 1X vibration. Alternatively we can use a low Fmax to filter out higher frequency data and low resolution data (e.g. 100 lines of resolution) in order to increase the speed of the measurement.

Determine the Fmax

In order to understand the relationship between Fmax, resolution and the number of shaft revolutions we need to review the formula provided below:

$$T = T_s \times N = \frac{N}{F_s} = \frac{N}{2.56 \times F_{max}} = \frac{LOR}{F_{max}}$$

The variable 'T' represents the time span; the amount of time required to collect the time record, which in turn is used to generate the spectrum. As you can see, if the Fmax is higher, the time span 'T' will be shorter, and vice versa. You can also see that if the lines of resolution (LOR) setting is higher, the time span 'T' will be longer and vice versa.

We are interested in the time span because that tells us how many times the shaft will rotate during the test. If we have just one rotation of the shaft the display will be very clear. Additional rotations will provide an indication of the consistency of the path (orbit) of rotation. The time for one rotation of the shaft is 60/RPM (or 1/[speed in Hz].) For example, if the machine speed was 1785 CPM (approximately 30 Hz), the time for one rotation is 1/30 or 0.033 seconds.

One way to compute the ideal setting is to start with a low resolution measurement; 100 lines of resolution. This ensures the fastest possible measurement – we do not need higher resolution.

For **one rotation of the shaft,** the Fmax is therefore equal to:

$$\frac{LOR}{T} = \frac{100}{\left[\frac{1}{Speed_{HZ}}\right]} = Speed\_HZ \times 100$$

For additional rotations of the shaft the equation becomes:

$$F_{max} = \frac{Speed_{HZ}}{[\#revolutions]}$$

Example: Speed = 25 Hz and LOR = 100

Fmax = 2500 Hz for one revolution

$$25 \times \frac{100}{1} = 2500Hz$$

Fmax = 250 Hz for 10 revolutions

$$25 \times \frac{100}{10} = 250Hz$$

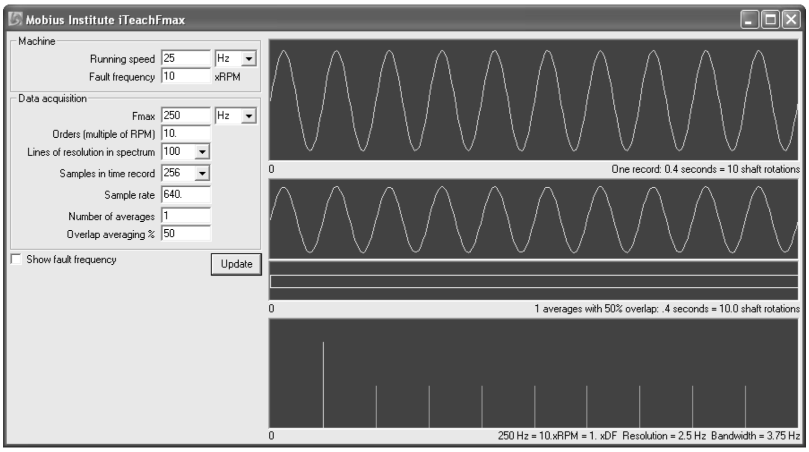


Figure 11-25

If you wish to filter out the vibration above 1X (assuming your collector/analyzer does not provide a filtered option), you may set the Fmax slightly above the speed of the machine. If the speed of the machine is 30 Hz, the Fmax could be 40 Hz (you will need to check your analyzer options). If you use 100 lines of resolution, you can compute the number of revolutions as follows:

$$\#revolutions = Speed_{HZ} \times \frac{100}{F_{max}}$$

As you can see, if Fmax is only slightly higher than the $Speed_{HZ}$ the number of revolutions will be close to the lines of resolution (LOR) – a high number. This may not provide a very clear orbit display.

Using time synchronous averaging (TSA)

If your data collector/analyzer offers the time synchronous averaging option, it should be utilized to provide the filtering that we desire. It will filter out all vibration that is not synchronous with the running speed of the machine. If the Fmax is set high, the orbit will include the 1X, 2X, 3X and other orders. If the Fmax is set above 1X and below 2X, the result will be a clean orbit display.

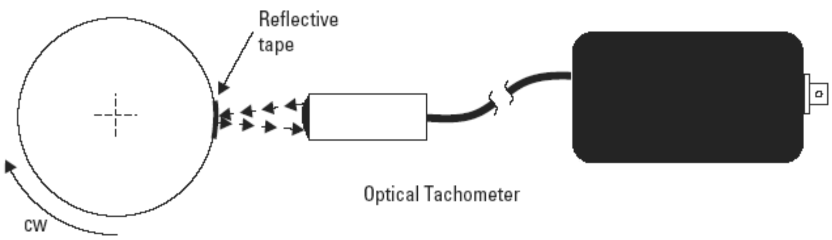


Figure 11-26

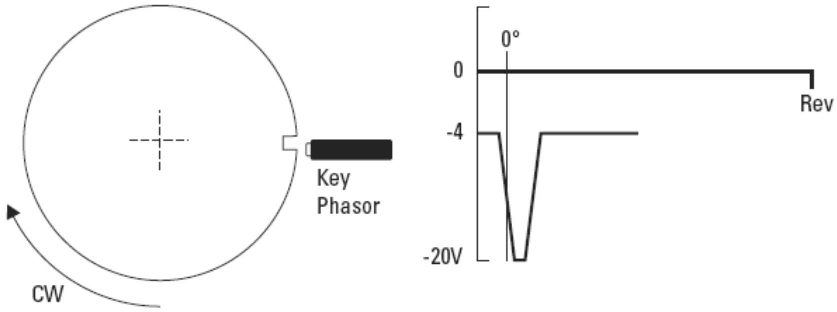


Figure 11-27

A general guide to orbit analysis

The following is a general guide to orbit analysis:

- Orbit plots typically show multiple revolutions of the shaft (e.g. 8 revolutions).
- Each Keyphasor dot indicates one revolution.
- A normal orbit should be slightly elliptical with one Keyphasor dot, and low vibration level.
- Distorted, or twisted orbits are usually due either to rubs or to runout.
- Multiple Keyphasor dots indicate a sub-synchronous vibration.

Diagnosing preloads with orbit analysis

Using a combination of centerline plots (to see how the shaft moves during start-up and when it rotates at full running speed), and the shape of the orbit, we can assess whether preload exists.

Example one: An orbit from a "good" machine

The following is an example of an orbit display and centerline plot of a machine operating correctly. Note the following characteristics:

1. The centerline has moved up and to the left (it is rotating clockwise).

2. The eccentricity ratio is approximately 0.5, when means the shaft is rotating around a point which is midway between the center of the bearing and the wall of the bearing.

3. The orbit is elliptical in shape.

4. The direct orbit is a similar shape to the filtered orbit, indicating that the vibration is dominated by 1X vibration.

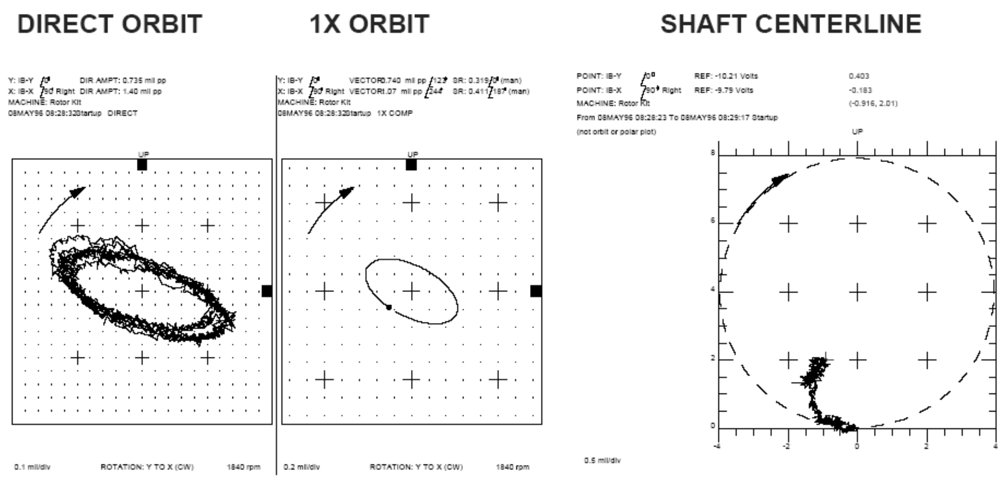


Figure 11-28

The orbit is wider than it is tall, which means that the forces are higher in the vertical plane (gravity) than horizontal – there is more movement in the horizontal plane.

Example two: An orbit from machine with preload

This set of data indicates that a problem exists. The average shaft center has moved higher in the bearing – it is now in the upper left quadrant. The eccentricity ratio is higher. This data indicates that a preload may exist. The orbit shape (at the inboard bearings) is common when misalignment exist between two components

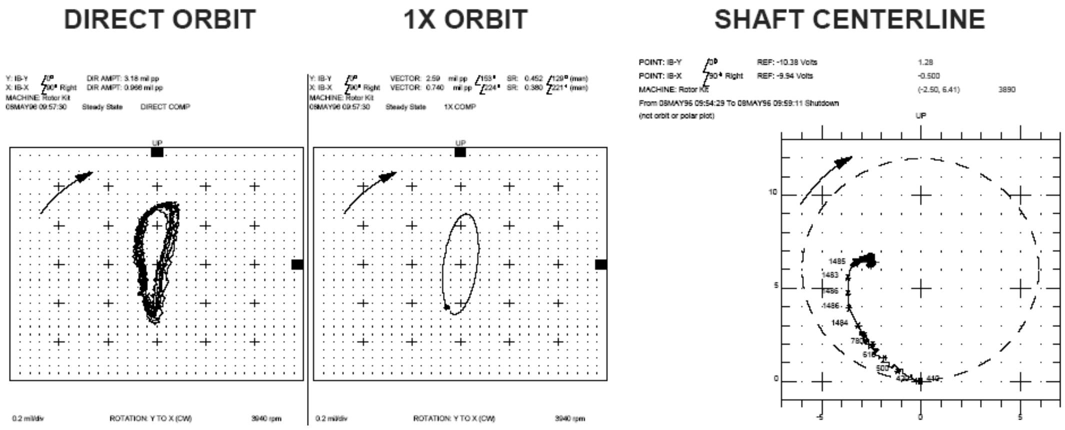


Figure 11-29

Example three: An orbit from machine with heavy preload

This data is representative of heavy preload. You can see that the shaft is close to the wall of the bearing. The eccentricity ratio is now approximately 0.95.

We can see from the 1X filtered orbit that there is significantly greater movement in the vertical direction compared to the horizontal direction.

The direct orbit shape is quite different to the filtered orbit, indicating that we now have other sources of vibration. In fact, due to the misalignment we now have higher 2X vibration.

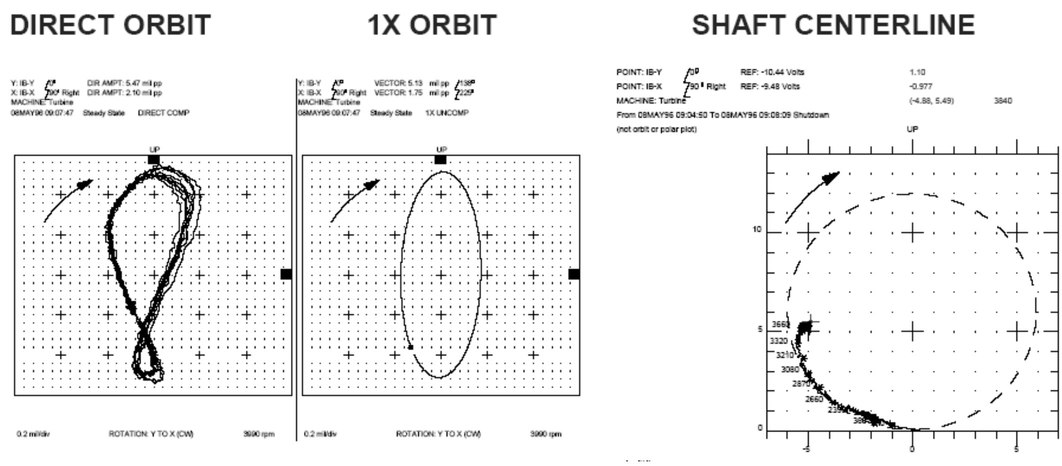


Figure 11-30

Based on the shape of the orbit we can visualize the shaft moving in the same direction as the direction of rotation for some of the time, and *against* the direction for part of the time. This is called *reverse precession* and will be explained in greater detail shortly.

Preloads summary

Radial preloads can be caused by gravity, fluidic forces, abnormal bearing loads (especially internally adjustable types), seals, misalignment, and the effects of pipe strain on the machine itself.

Centerline diagrams that show a high eccentricity are a good indicator of high preloads. Highly elliptical 1X orbits, and "figure 8" direct orbits are also very good indicators that high preload exists.

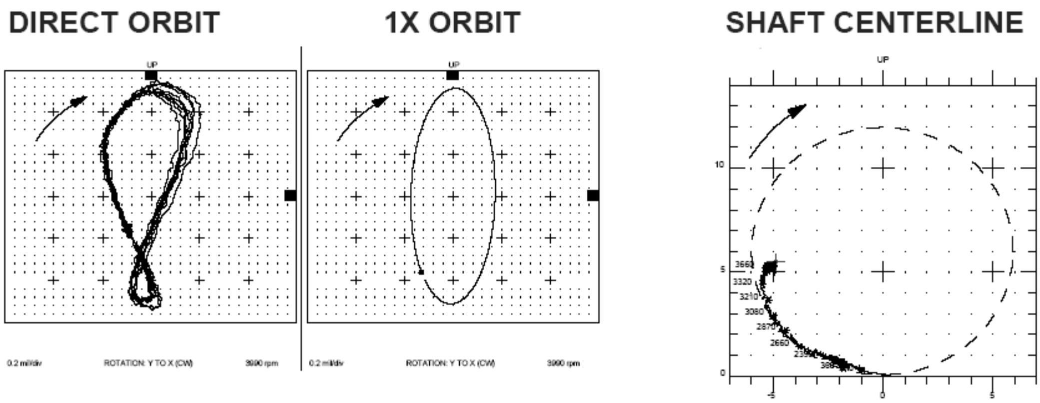


Figure 11-31

Orbit direction and vibration precession

The direction of the orbit should be the same as the direction of the shaft rotation. You can determine the direction of the orbit via the "blank-bright" order – the "bright" refers to the dot – the dot is the beginning of the cycle – the blank area (or the small gap in the orbit) precedes the dot. In these two examples, for counter-clockwise and clockwise rotation, you can see that the orbit is rotating in the same direction as the shaft.

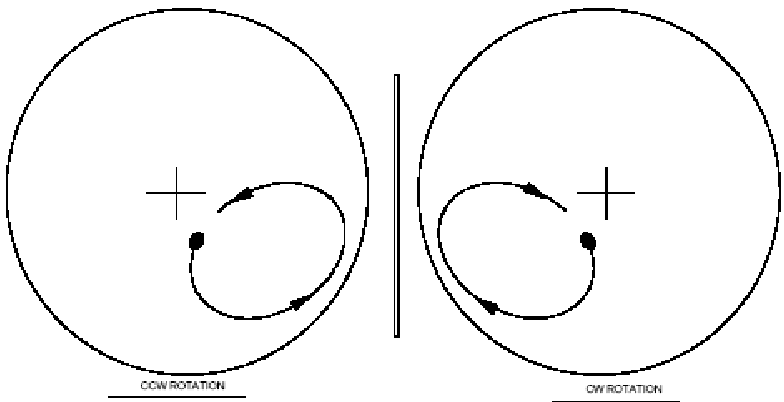


Figure 11-32

Diagnosing fault conditions with orbits

The orbit indicates the rotating path of the center of the shaft, and the shaft moves according to the forces applied to it, and according to the dynamics of the shaft. The dynamics of the shaft are a function of the mass, stiffness and damping. Clearly, the mass is greater in the vertical direction and the stiffness of the system is not equal in the vertical and horizontal directions. Therefore it is perfectly normal for the orbit to have an elliptical shape.

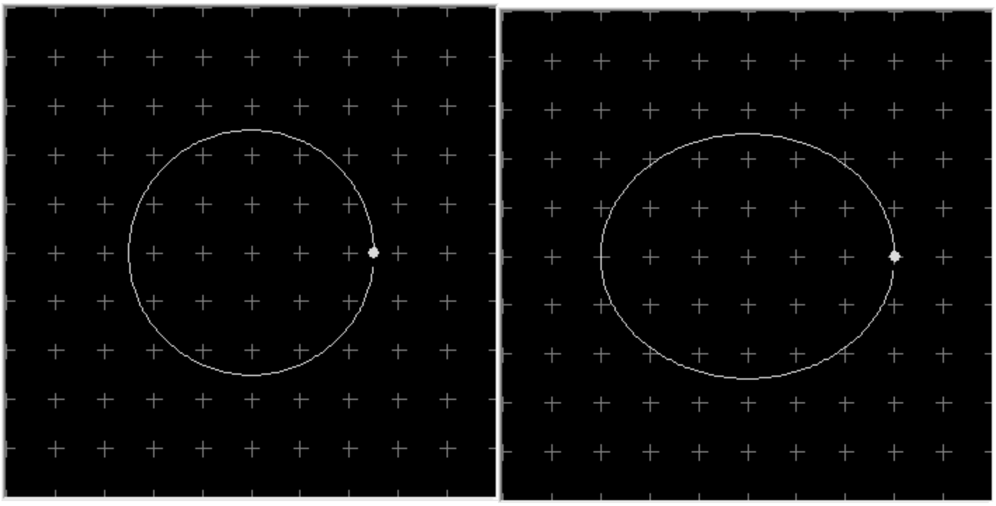


Figure 11-33

Diagnosing unbalance

If the machine is unbalanced the vibration level will increase (i.e. the size of the orbit will increase), however it will retain the elliptical shape. The ratio of the width to height will be as much as 4:1 (1:1 would be a circle).

The vibration is dominated by the 1X component so the direct and filtered orbits will be similar in shape.

The dot should not move in a live display.

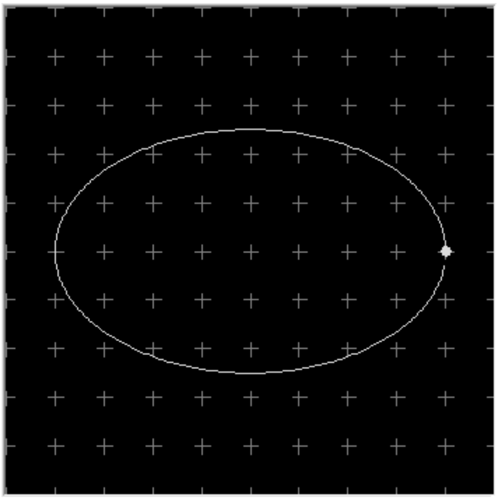


Figure 11-34

Diagnosing misalignment

If misalignment exists, the orbit shape will be flatter. The ratio will be greater than that described for unbalance; i.e. greater than 5:1.

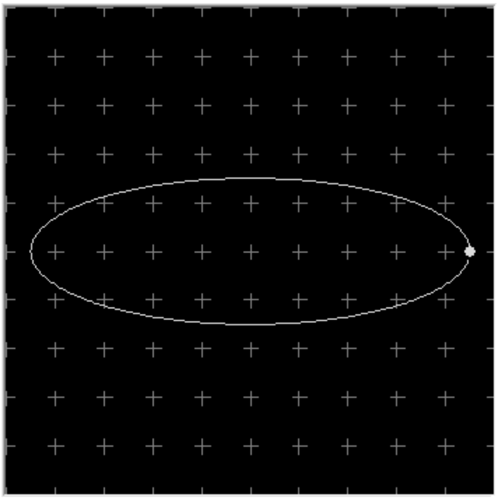


Figure 11-35

The angle of the ellipse will indicate the direction of the major force. In the following example the force on the shaft is coming from the top-left.

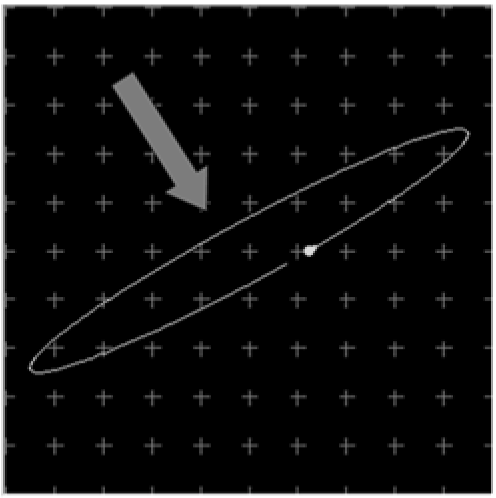


Figure 11-36

Detecting a loose rotating part

If there is a loose rotating part the vibration level will fluctuate. In one position the residual unbalance of the part may be in-phase with the residual unbalance of the rotor – the vibration will be highest. As the part rotates relative to the rotor the residual unbalance of the part may be 180 degrees out of phase with the residual unbalance of the rotor – the vibration will be lowest. For these reasons the phase angle of the vibration will gradually change. Of course, this will only be detected if a live display of the orbit is observed.

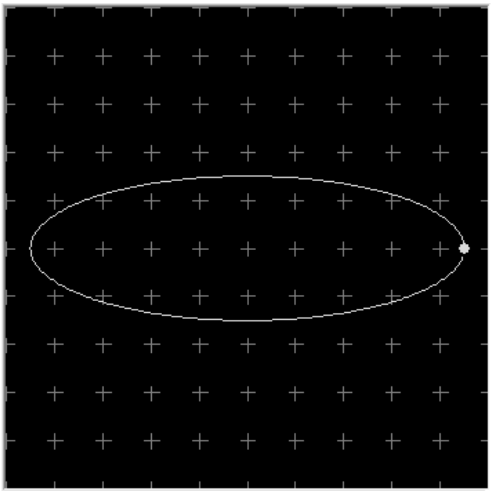


Figure 11-37

Studying loops and counting dots in the orbit

If there are two dots in the filtered orbit display there must be sub-synchronous vibration. In this case it is important to observe a live orbit display to observe whether the position of the dots on the orbit changes.

0.5 X – two dots stationary on live display

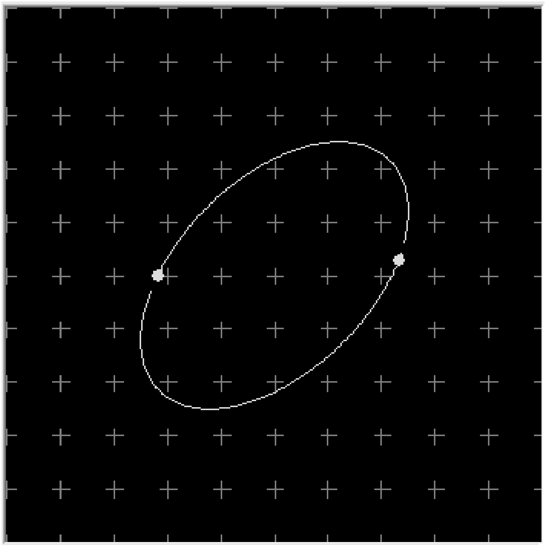
- Possibly a rub

<0.5X - dots move *against* direction of rotation

- Oil whirl (oil instability)

>0.5X – dots move in same direction

- Structural or rotor resonance



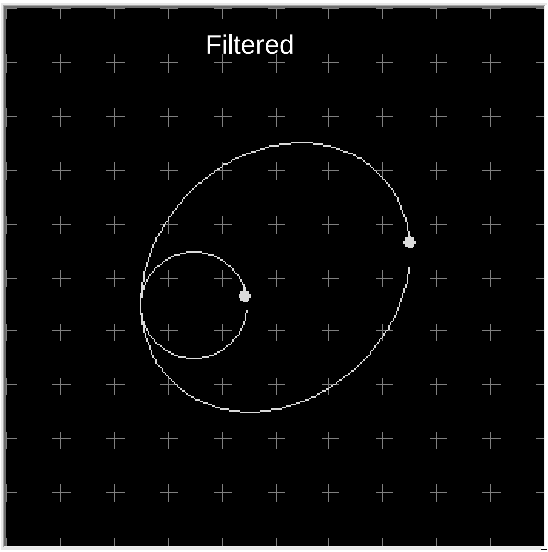


Figure 11-38

The direct orbit plot will either have internal loops, when precession is in the direction of the shaft rotation (forward precession), or outer loops, when precession is against shaft rotation (negative precession).

Fluid induced instabilities: oil whirl and oil whip

Oil whirl and oil whip are destructive conditions. We can use orbit analysis and spectrum analysis to detect the condition, and we can study the position of the shaft to determine whether these conditions may be induced.

Oil whirl

Oil whirl is a self-excited fluidic malfunction that typically occurs in plain sleeve bearings. It may occur for a number of reasons, including:

- Misalignment – bearing unloaded (eccentricity ratio ~0)
- Excessive clearance
- Lightly loaded, low damping

Oil whirl generates a characteristic source of vibration in the range of approximately 0.38X to 0.48X. Because it is lower than 0.5X we will witness two dots in the live orbit, and the direct orbit will have an inner loop. The dots will move in the same direction as the shaft rotation. The vibration precession will be positive (in the same direction as shaft rotation), therefore there will be an inner loop in the direct orbit. The motion of the shaft is circular, so the orbit shape will circular.

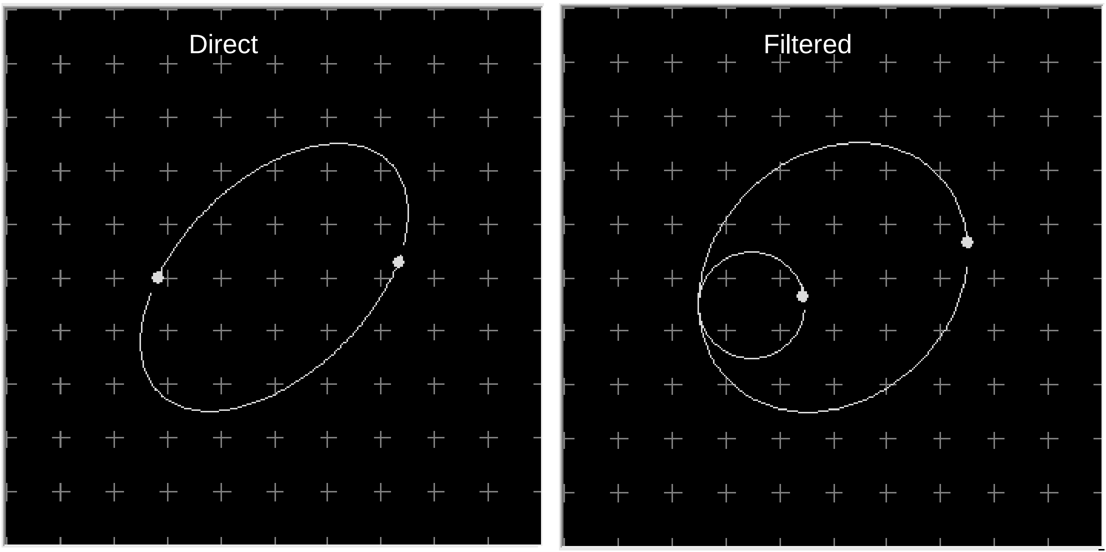


Figure 11-39

Spectrum analysis of oil whirl

If the vibration is displayed as a spectrum, there will be a peak in the range of approximately 0.38X to 0.48X. The height of that peak can become quite high.

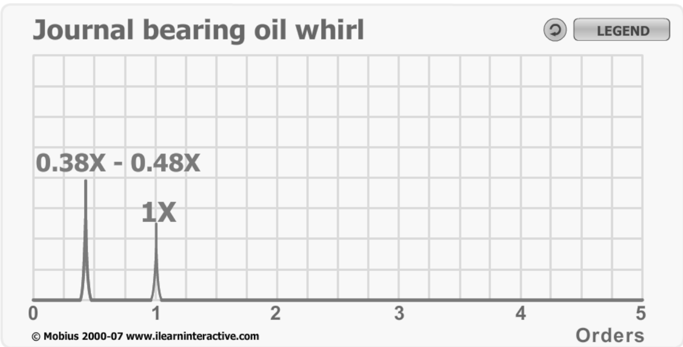


Figure 11-40

Oil (fluid) whip

Oil (fluid) whip is a very destructive condition. This condition can occur when the speed of the machine is over twice the critical speed (or the first balance rotor resonance); therefore the critical speed is close to the oil whirl frequency. The oil whirl excites the resonance and a violent condition arises. Vibration levels are very high and the machine must be stopped or else catastrophic failure may occur.

The filtered orbit will have multiple dots. A live display may indicate that the dots are not moving – this means that the vibration is half running speed

Shaft rubs

Using spectrum analysis, orbits and centerline diagrams we can determine if the shaft is rubbing against the bearing (or if any other rotating part is rubbing on a stationary part).

Detecting shaft rubs with spectrum analysis

When a shaft begins rubbing on the wall of the bearing or against a seal, we would expect to see a number of changes in the vibration:

1. The waveform will be "distorted".

2. The spectrum will have a large number of 1X harmonics.

3. There may be sub-order peaks and harmonics (e.g. ½X, ¼X, etc.).

4. The noise floor may lift.

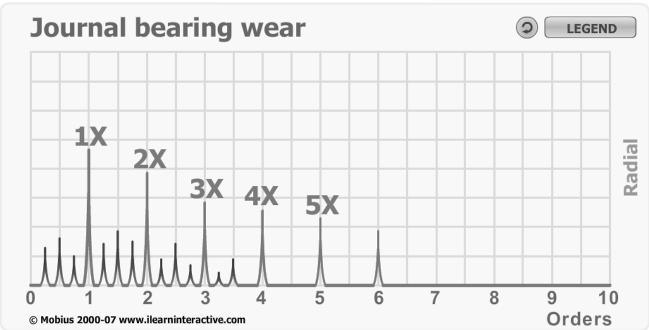


Figure 11-41

Detecting shaft rubs with orbit analysis

The orbit will also provide an indication if shaft rub is occurring. The centerline plot should also provide a clear indication if the shaft is rubbing the bearing wall – the average center will be close to the edge of the circle that is used to indicate the maximum clearance.

The filtered orbit can change shape slightly if it is just a light rub. The orbit path will change because the shaft is unable to rotate in its normal elliptical path. It may take on a "tear-drop" shape. It may just have a flat edge on the elliptical/circular shape. The orbit may also be flattened due to the preload or other force that is causing the shaft to rub.

If the rub worsens we will see the sub-order vibration, e.g. ½X and/or ¼X. The filtered orbit will have multiple stationary dots (if ½X) or rotating dots otherwise. The direct orbit will have multiple loops: the loops will appear inside the main orbit if there is forward precession, or outside the main orbit for reverse precession. The number of loops relates to the fractional order – 2 loops for 1/2X vibration, 3 loops for 1/3X vibration, etc.

Figure 11-42 - Forward precession (1/3X)

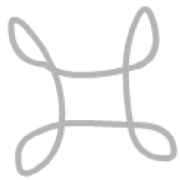
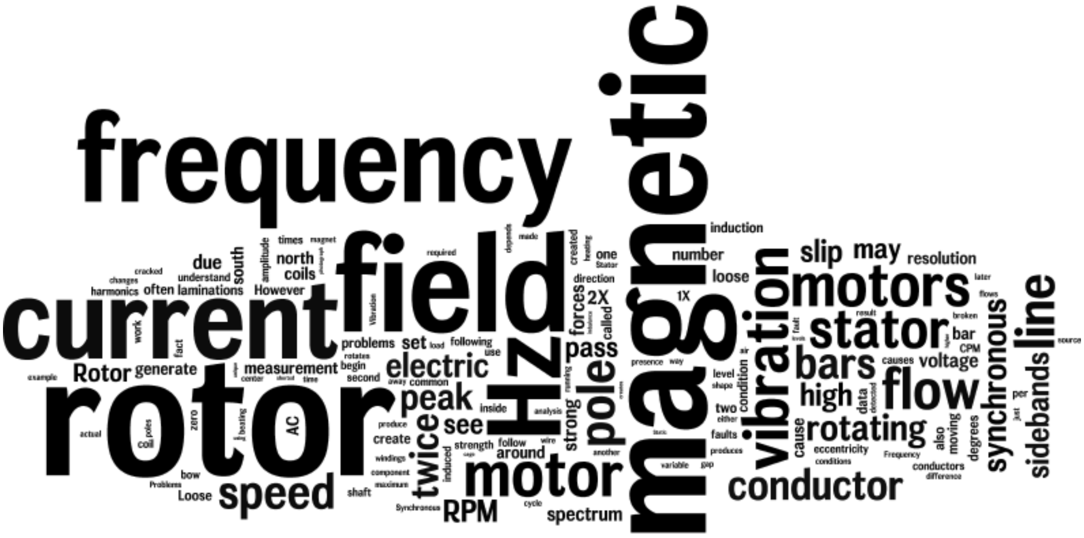


Figure 11-43 - Reverse precession (1/4X)

Summary

The combination of orbit analysis and centerline analysis provide a great deal of information about the operating condition of machines with fluid-film bearings. Live orbits, direct orbits, and filtered orbits provide very useful information.



Chapter 12

Electric Motor Analysis

Objectives:

- Understand the relationship between current flow and the creation of a magnetic field
- Understand how induction motors, synchronous motors, and DC motors work
- Understand how Variable Frequency Drives (VFD) work
- Describe how an electrical problem can look like a 2x or 4x peak in spectra
- Describe how misalignment can invoke an electrical problem in a motor
- Describe the causes of 3 types of sidebands from electric motors

Introduction

In order to fully understand electric motors and become skilled in diagnosing fault conditions it helps to understand how they work. Like any rotating machine they suffer from mechanical faults: unbalance, misalignment, bearing faults, bent shaft and more. However electric motors, whether they be synchronous, DC, induction, and whether they be driven from a fixed frequency or from a variable frequency, they all have vibration that result from the magnetic forces that make them turn. Understanding those forces can help you to understand the vibration patterns that we witness.

The basics of magnetism

The first basic concept to understand is that north attracts south, and vice versa. Electric motors are basically designed to create a rotating magnetic field that a magnetized rotor is compelled to follow. The difference between each type of motor is how it creates the rotating magnetic field, and how the rotor is magnetized.

Figure 12-1

Creating a magnetic field with current flow

When we pass an electric current through a conductor a magnetic field is created around the conductor – but only while the current is flowing. When the current flows in one direction a magnetic field will be set up with the north and south pole set up one way. When the current flows in the other direction, the north and south pole reverse direction. The strength of the magnetic field is proportional to the strength of the flow of current. And if the current flow stops, the magnetic field dies.

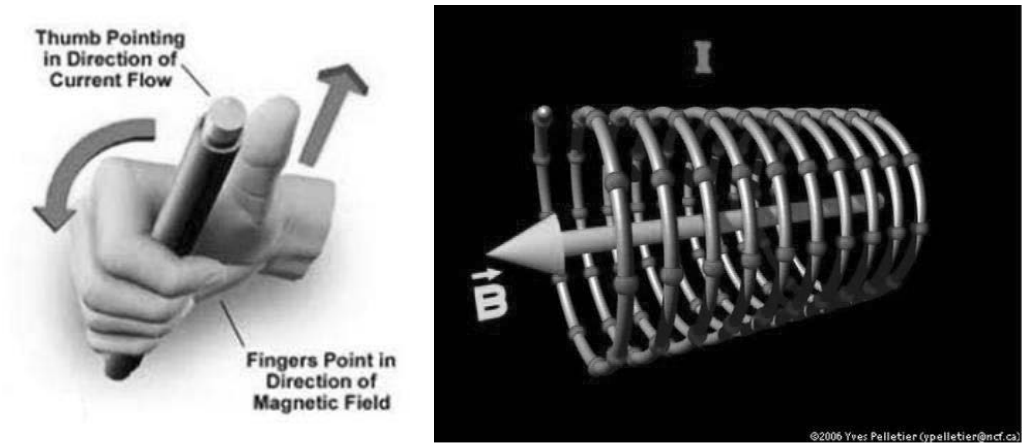


Figure 12-2

For example, if a DC voltage is applied to the ends of a conductor, current will flow and a magnetic field will be created. But moments later, if the current flow stops, the magnetic field will die.

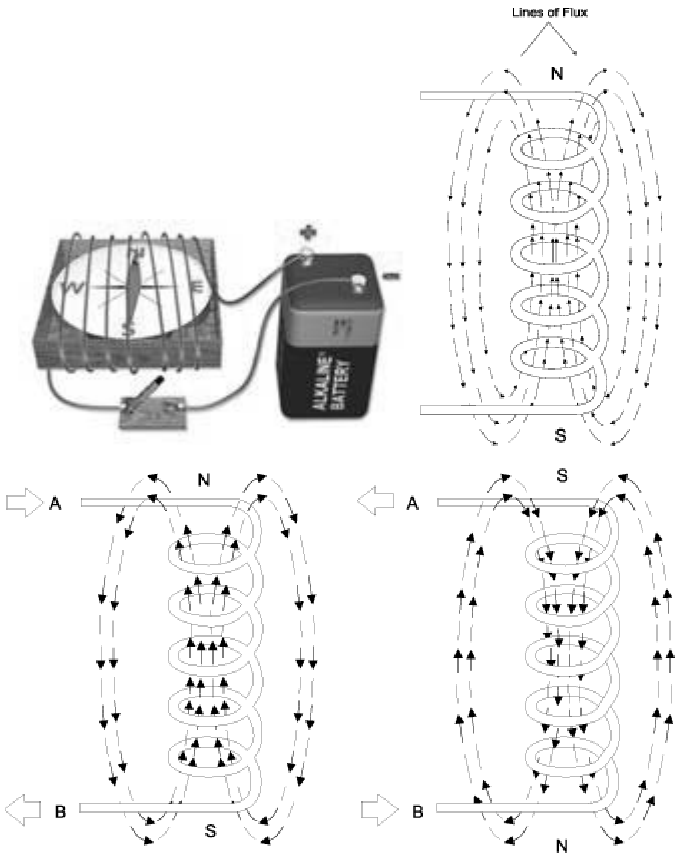


Figure 12-3

However an alternating current AC system is different. An AC voltage rises and falls in a sinusoidal pattern with a frequency of either 50 Hz or 60 Hz. As the voltage rises from zero to the full voltage 90 degrees later, the current flow will increase to a maximum value. As that happens a magnetic field will be established and will reach its maximum in phase with the current flow. But as the current falls back to zero, a further 90 degrees later, the magnetic field will fall again to zero.

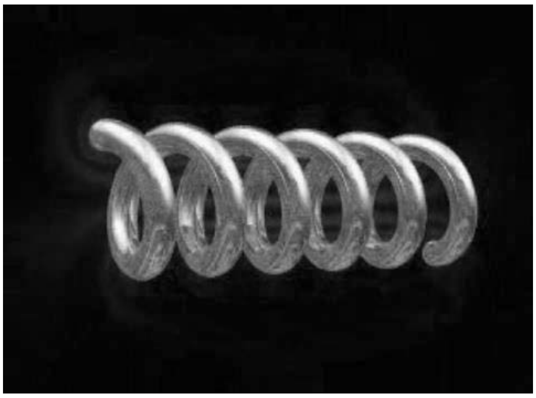


Figure 12-4

But then the current will begin to flow in the opposite direction. Therefore the magnetic field will again be established, but this time the north and south pole will be reversed. Again the magnetic field will reach its maximum when the current flow is greatest, at 270 degrees, and it dies again as the current comes back to zero at 360 degrees.

This happens with every cycle; either 50 times per second or 60 times per second.

If we look at that once more, we can see that the magnetic field becomes strong twice per cycle; once with the north and south oriented one way, and then with the north and south oriented the other way. That is why we often see a strong source of vibration that occurs at twice the line frequency; 100 Hz or 120 Hz.

Whenever a magnetic object, or ferrous object, is subjected to a magnetic field, it reacts to the forces. This phenomenon is called magnetostriction. Magnetostriction is the deformation of a magnetic material in the presence of a magnetic field, and it causes vibration at 100 or 120 Hz in all electric devices such as motors, generators, transformers, etc. It is therefore common to observe 100 Hz or 120 Hz vibration in electric motors

Coils and magnetic fields

The next issue to consider is the effect of taking the conductor and creating a coil. When current flows through a conductor a magnetic field is created, as just described. However when the conductor is shaped to form a loop, the magnetic field strength is doubled. If we make two loops, the magnetic field is further strengthened. Therefore electric motors use coils of insulated wire to create very strong magnetic fields. The coils are built into the stator and optionally the rotor (depends upon the type of motor).

Inducing current in a conductor

Another important fact is that when a conductor is subjected to a moving magnetic field, current will flow in that conductor. It works in reverse to what we have just described. Current flow creates a magnetic field, and that can induce current in another conductor. The amount of current flow depends upon the strength of the magnetic field and the length of conductor (the number of coils) that passes through the magnetic field.

Ignoring losses for a moment, if we use an AC current of 5A to create a magnetic field using 10 coils of conductor, and there was another coil of 10 conductors in the magnetic field, then 10 A will flow through the second coil. That is how transformers work, and that is how induction motors work. You see, when the 5A is induced to flow in the second coil, it generates a magnetic field as well. We'll discuss this in greater detail shortly.

The main point is that current flowing through a conductor will generate a magnetic field, and a conductor that is either moved through a magnetic field, or is placed in a moving magnetic field, will have current flow induced in that conductor.

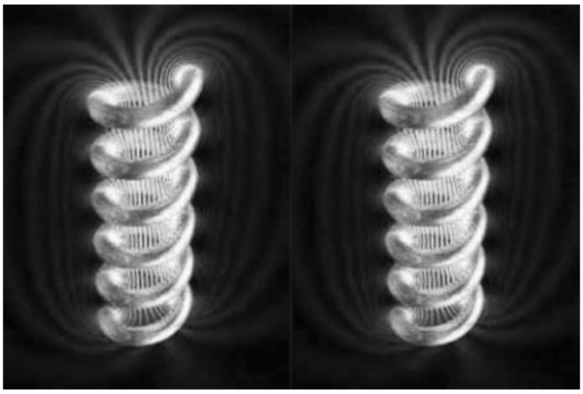


Figure 12-5

The application to electric motors

This characteristic of induced magnetism and current flow is put into practice when designing electric motors. The stator is constructed with coils of wire distributed around the inside of the motor case. When the AC voltage is applied to the coils they cause a rotating field to be created. The rate of rotation depends upon the number of poles in the motor. It is called the synchronous speed.

$$Synchronous\ speed, n_s = \frac{120f}{p}\ [rpm]$$

A two pole, 3-phase motor generates a rotating field that rotates at 60 Hz, or 50 Hz.

A four pole motor will create a rotating magnetic field at 30 Hz, or 25 Hz.

A six pole motor will create a rotating magnetic field at 20 Hz, or 16.6 Hz.

Note that if we use a variable frequency drive, instead of using 60 Hz (or 50 Hz), we can use any frequency that we choose.

Now that we have a rotating magnetic field we can put it to work to drive a rotor.

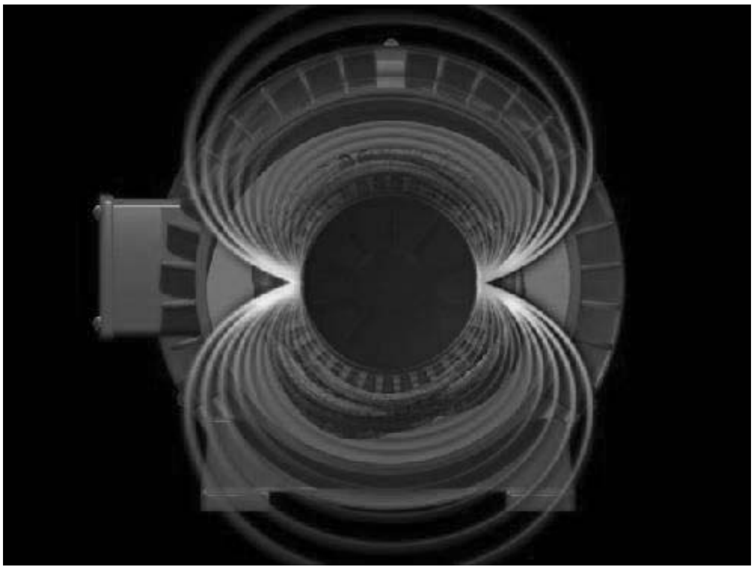


Figure 12-6 - Rotating magnetic field in motor stator

Synchronous motors

If the rotor is itself a magnet, then it will follow the rotating magnetic field at the synchronous speed. It is called a **synchronous motor**. It can be achieved one of two ways:

If we place a permanent magnet rotor inside the rotating field, then it will spin at the synchronous speed.

If we place a wound rotor inside the stator, and we supply the same AC voltage to the rotor, it will generate a magnetic field which will also follow the rotating magnetic field at the synchronous speed.

Induction motors

The third option is more complicated. Induction motors also include a rotor with conductors, however power is *not* provided to the windings. Instead, when the motor is started and the magnetic field begins to rotate, a current is induced into the conductors of the rotor. A strong current will begin to flow because it is stationary inside a strong moving field. As soon as the current begins to flow a magnetic field will be generated. This causes the rotor to behave as a magnet - so it rotates to follow the rotating magnetic field. However as the speed of the rotor approaches the synchronous speed, the conductors are no longer moving within such a fast moving magnetic field (relatively speaking). If the rotor was to turn at the synchronous speed, no current would flow, the magnetic field would drop away, and the rotor would begin to slow down.

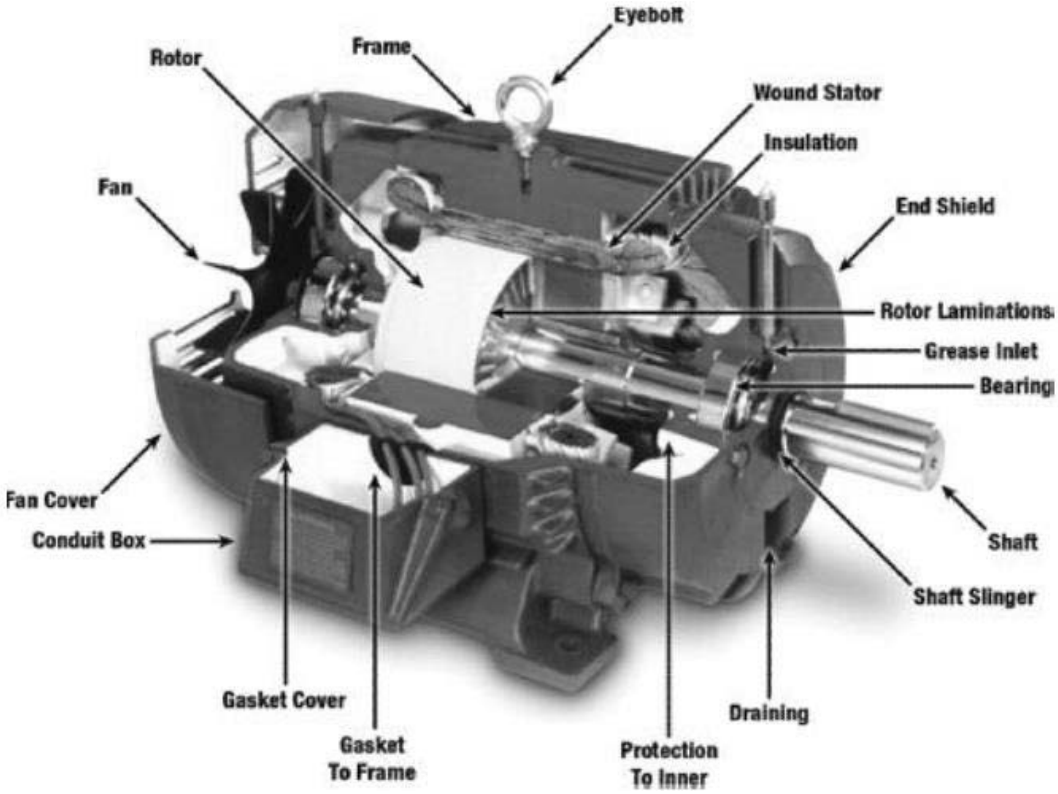


Figure 12-7 - Induction motor

The result is that the rotor travels more slowly than the magnetic field. The difference between the synchronous speed and the actual speed is called the *slip* frequency.

$$Rotor\ speed, n_r = n_s(1 - s)$$

$$s = \frac{n_s - n_r}{n_s}$$

Squirrel cage induction motors

Squirrel cage motors are very common in industry. In overall shape the rotor is a cylinder mounted on a shaft. Internally it contains longitudinal conductive bars of aluminum or copper set into grooves or “slots” and connected together at both ends by shorting rings forming a cage-like shape.

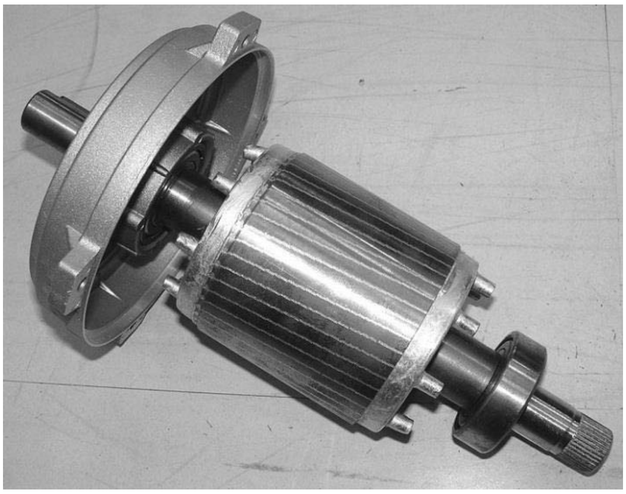


Figure 12-8

The core of the rotor is made up of a stack of iron laminations. The photograph below shows a rotor lamination and a stator lamination.

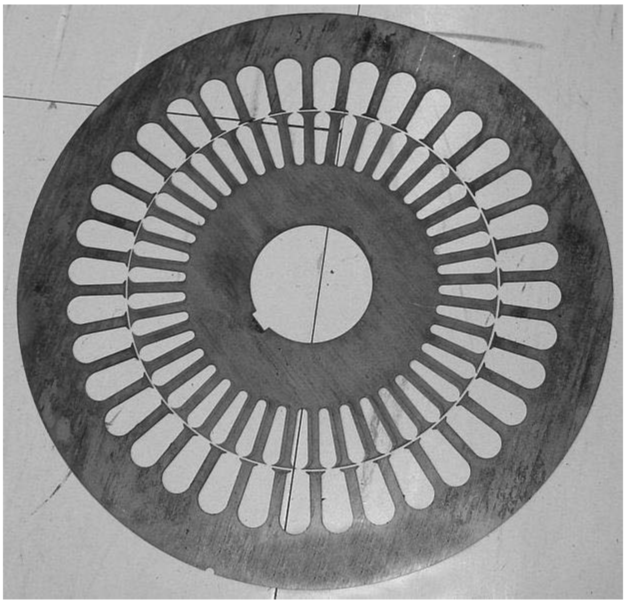


Figure 12-9

The following is a photograph of a stator.

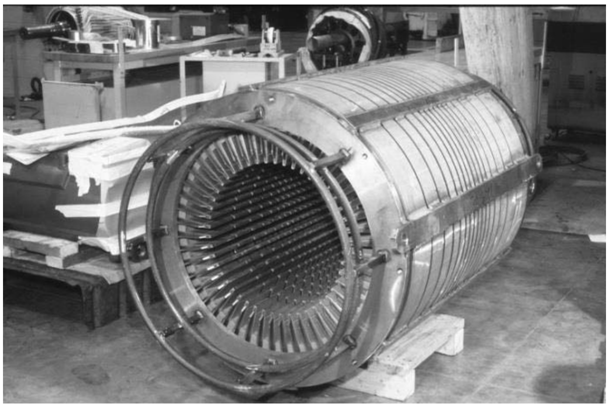


Figure 12-10

Fault diagnosis

As stated previously, all electric motors are susceptible to unbalance; the rotor may bend or warp due to overheating; they may be misaligned with the component they are driving; and their bearings may begin to wear. These, and other faults conditions, are mechanical in nature and they can be diagnosed in ways that have been covered elsewhere.

However, electric motors develop unique fault conditions due to the electromechanical forces that we have discussed. If the rotor is off center, or the rotor rotates eccentrically; or the rotor bars or stator slots are cracked or broken, then the motor will generate a unique set of vibrations (and potentially there will also be changes to the current flow) that can be detected.

Sources of Vibration in Electric Motors

Twice the line frequency (100 Hz or 120 Hz) is always a measurable vibration component in an electric motor. The magnetic attraction between the stator and rotor varies at this rate, and the iron itself changes dimension a little in the presence of the varying magnetic field due to "magnetostriction".

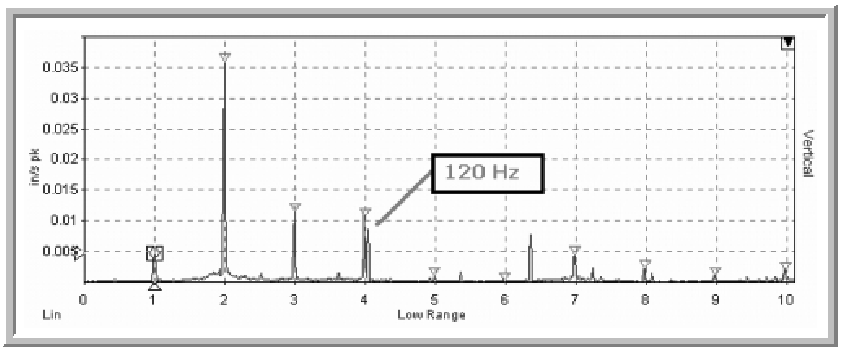


Figure 12-11 - A peak at 2x the Line Frequency is near the 4x peak. 2x line frequency is common in electric motors.

In two-pole motors (synchronous speed of 3000 CPM or 3600 CPM), it is sometimes difficult to distinguish the 100 or 120 Hz (twice line frequency) peak from the 2X peak in the vibration spectrum. This is another good reason for collecting spectra with high resolution.

One test to verify the presence of a 2X peak instead of 100 or 120 Hz, is to take a measurement while the motor is running, and then cut power to the motor. The 100 or 120 Hz peak will disappear, whereas 2X will remain (as the motor drops in speed).

The 2x line frequency peak can dominate the spectrum and can be very strong and dominate the spectrum.

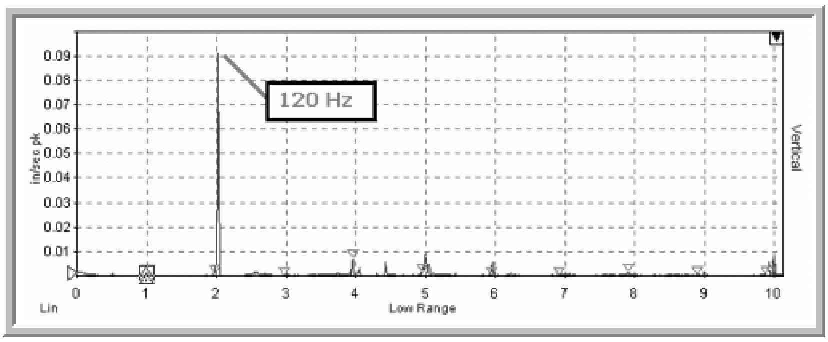


Figure 12-12 - The 2x line frequency can dominate the spectrum. It is near the 2x peak.

Variable frequency drives

A variable frequency drive works in a similar way, however, instead of 50 Hz or 60 Hz line frequency, the voltage is supplied at a different frequency, thus controlling the motor speed.

The 2xLF peak will still be present but it will appear at 2x whatever the actual line frequency is at the moment.

Stator Problems

Stator problems will generate high vibration at twice the line frequency (100 or 120 Hz).

Static eccentricity

Static eccentricity produces an uneven stationary air gap between the rotor and stator that produces a very directional source of vibration. Instead of the magnetic forces being balanced between the rotor and stator, there will be very strong forces between the rotor and stator at the point where they are closest. 180 degrees away the force will be weakest because the rotor and stator will be furthest apart.

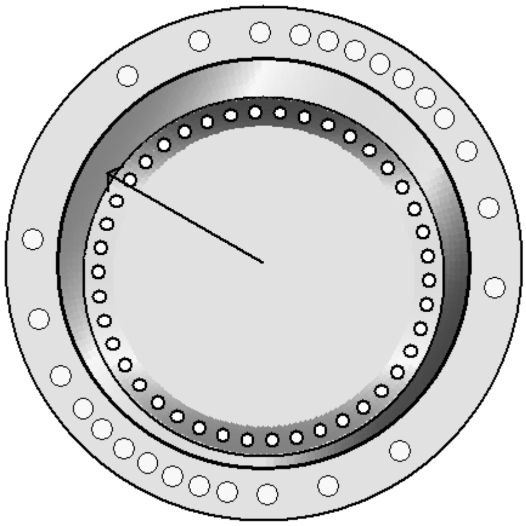


Figure 12-13

Because the magnetic forces rise and fall twice per electric cycle, we will see a high peak at 120 Hz or 100 Hz.

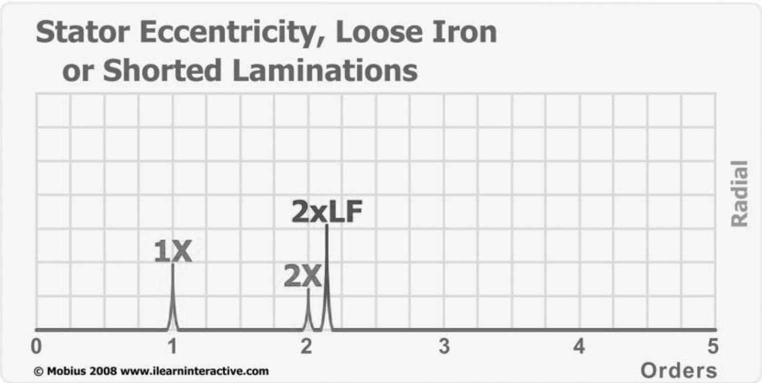


Figure 12-14 - Stator problems

Case study

The following set of data demonstrates how a spectrum may appear when the rotor is not spinning in the center of the stator.

In the first set of data we could easily think that we had high 2X vibration.

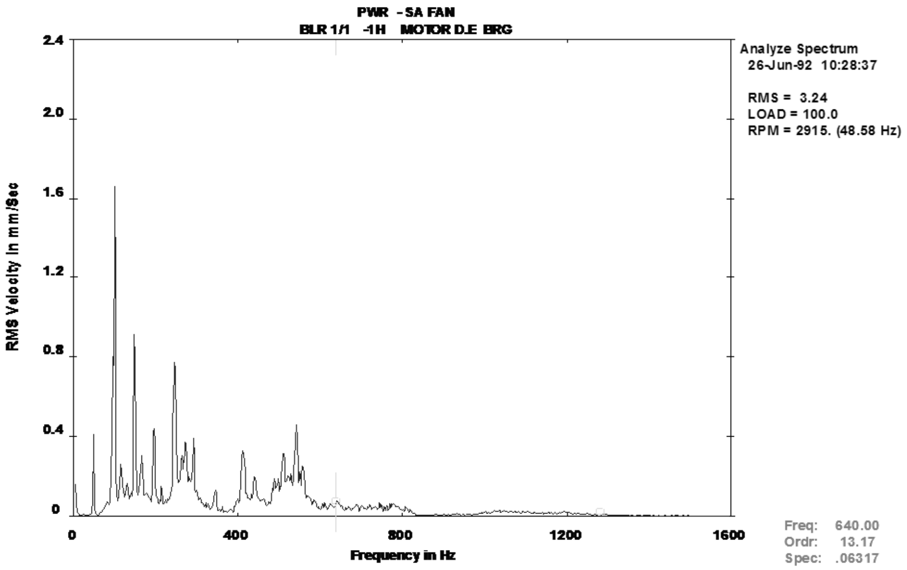


Figure 12-15

When we zoom in (or take a reading with higher resolution, it is apparent that there is in fact a 2X peak and a twice line frequency peak.

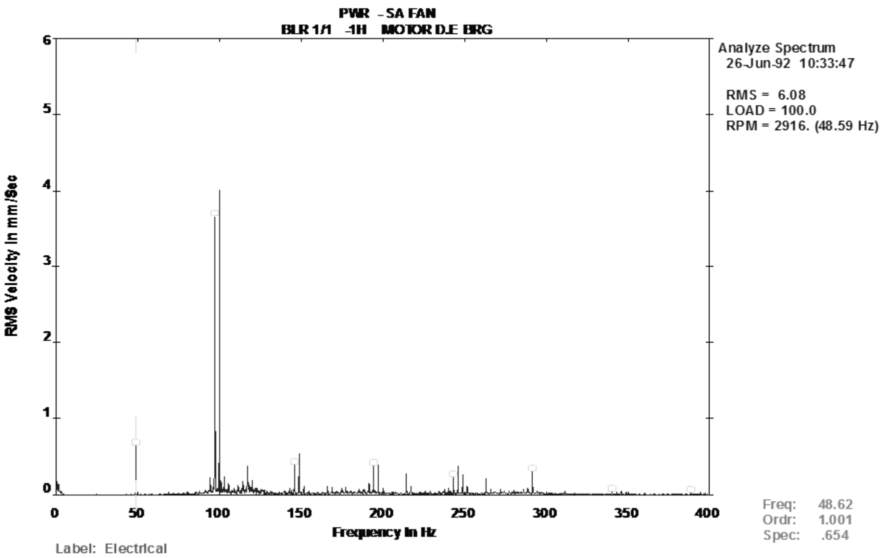


Figure 12-16

Soft Foot

Soft foot and warped bases can produce an eccentric stator. Altering the tension on the hold down bolts, correctly shimming the feet, reducing stress from conduit or piping, or repairing the foundation can resolve this problem.

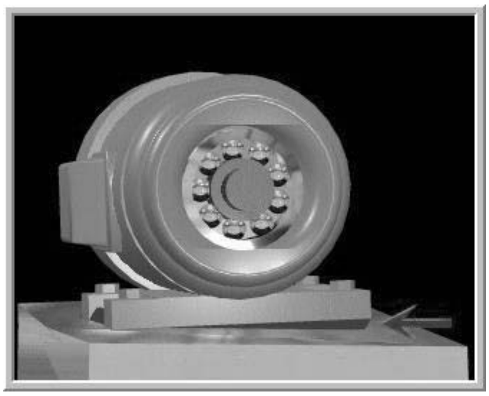


Figure 12-17 - Soft foot causes stator eccentricity

The vibration will again have a large peak at 120 Hz or 100 Hz. Depending upon the shape of the stator, it may be that there are two points where the vibration is highest – the two locations where the stator is closest to the rotor.

Any time you have static eccentricity or a soft foot problem, you may also hear a beating sound when testing the motor, especially on 2-pole motors. The beating comes from the interplay between the 2X vibration and the 2xLF vibration. The period of the beat will correspond to the slip frequency.

Rotor problems

There are a number of rotor-related faults that can be detected with vibration analysis and motor current analysis. They include rotor eccentricity, rotor bow, problems with rotor bars (cracking, breaking, etc.), loose rotor windings or rotor bars, and loose rotor on the shaft. Problems with the laminations or rotor bars may cause overheating which cause result in rotor bow. They are summarized below.

Eccentric Rotors

Eccentric rotors produce a rotating variable air gap between the rotor and the stator, which induces a pulsating source of vibration. Again you will see the twice line frequency component, however this time there will be pole pass sidebands around this frequency and the 1X peak.

The **pole pass frequency is the slip frequency times the number of poles**. The slip frequency is the difference between the actual RPM and the synchronous speed.

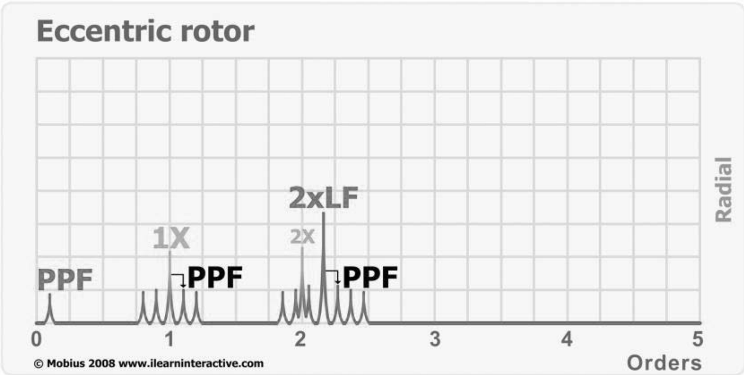


Figure 12-18 - Eccentric Rotors

Because the slip frequency is quite small, the pole pass frequency is therefore small, and thus a high resolution measurement may be needed to identify these sidebands.

Synchronous Speeds for 60 Hz & 50 Hz

2 pole: 3600 RPM or 3000 RPM
4 pole: 1800 RPM or 1500 RPM
6 pole: 1200 RPM or 1000 RPM
8 pole: 900 RPM or 750 RPM

Rotor bar problems

Broken, cracked, or corroded rotor bars can occur within induction motors, especially in motors that are frequently started and stopped under load. The starting current is much greater than the running current, and puts a strain on the rotor bars, causing them to heat.

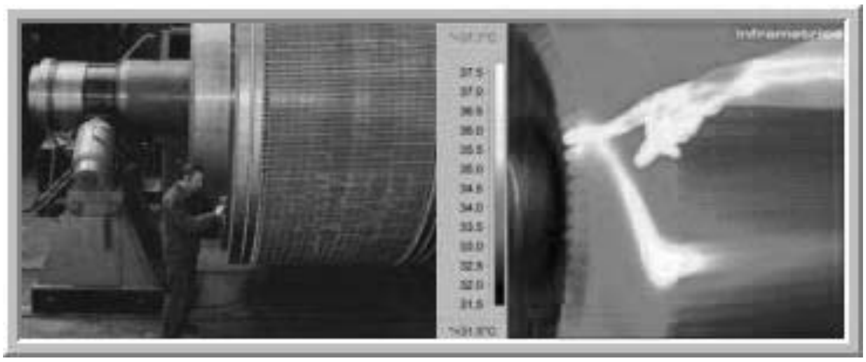


Figure 12-19 - Cracked or broken rotor bars short and produce heat

Rotor Bow

Uneven heating of the rotor due to unbalanced rotor bar current distribution causes the rotor to warp, or "bow", and rotor bow results in an imbalance condition with all its usual symptoms.

It can be detected by the fact that it goes away when the motor is cold. Local heating can be so severe in motors that the offending bar can actually melt and lodge in the air gap.

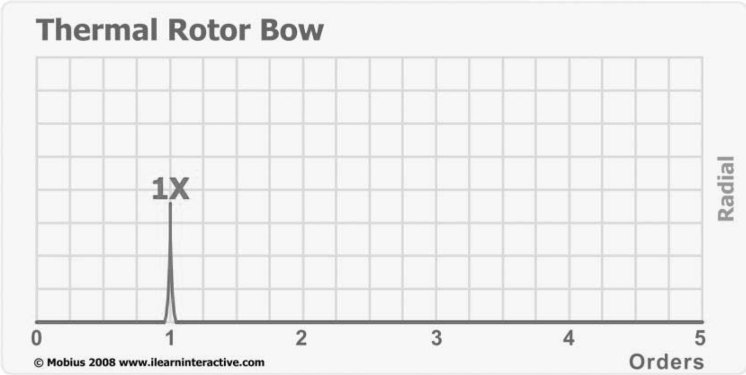


Figure 12-20 - Rotor bow looks like imbalance

Cracked Rotor Bars

Cracked rotor bars will generate pole pass frequency sidebands around 1X and its harmonics (2X, 3X, and so on). You will often see a very busy spectrum with harmonics of 1X, each with "skirts" of pole pass sidebands.

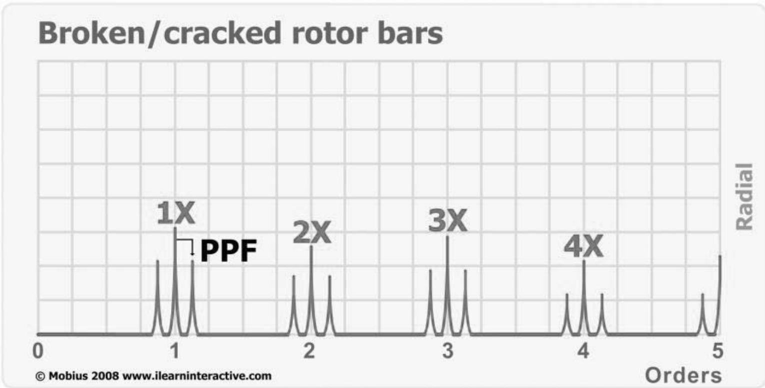


Figure 12-21 - Cracked rotor bars generate pole pass frequency sidebands around 1x and harmonics.

An induction motor with defective rotor bars will produce a vibration signature that slowly varies up and down in amplitude at twice the slip frequency of the motor. **This phenomenon is called beating,** and can often be heard as well as measured. The amplitude and frequency of the beats are dependent on the load on the motor (because that affects the slip frequency).

Loose Rotor Bars

If there are loose rotor bars there will be a peak at the rotor bar pass frequency, with sidebands of twice line frequency (100 or 120 Hz). Even if you do not know the number of rotor bars, if you see a high frequency with twice line frequency sidebands, you can be somewhat confident that this fault condition exists. There may be 35 to 96 rotor bars.

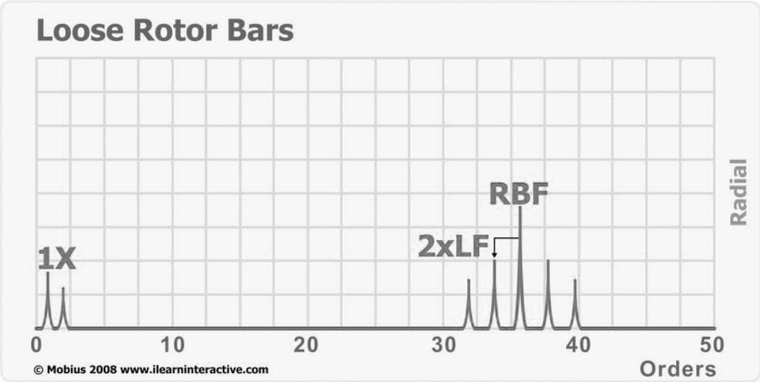


Figure 12-22 - It is normal to see RBF but a high RBF indicates cracked or broken rotor bars or Defective or loose rotor bar joints.

Rotor Bar Passing Frequency

It is common to see Rotor Bar Frequency (RBF - the number of rotor bars times the RPM) due to the fact that each rotor bar passes slight disruptions in the magnetic field due to the current's path through the rotor and stator. A high RBF indicates cracked or broken rotor bars or defective or loose rotor bar joints.

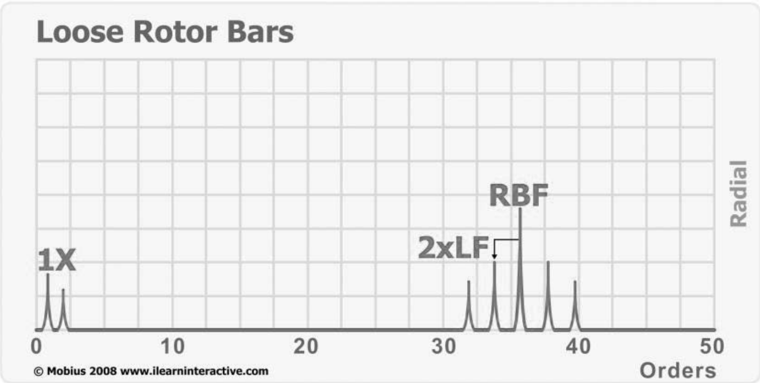


Figure 12-23

Loose Rotor

Sometimes the rotor can slip on the shaft, usually intermittently depending on temperature, and this causes severe vibration at 1X and harmonics. Abrupt changes in load or line voltage can instigate this condition.

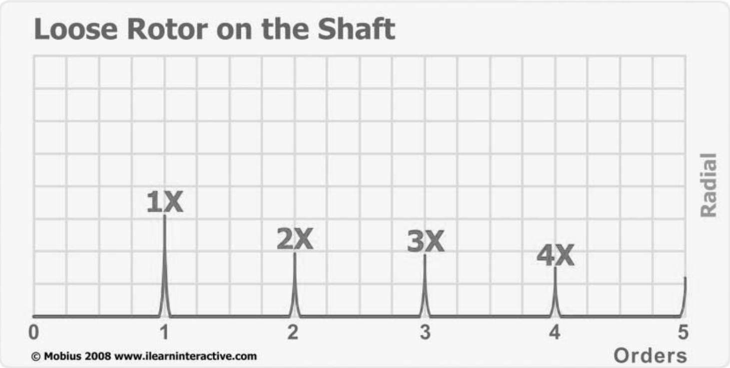


Figure 12-24 - Loose rotor produces harmonics of turning speed.

Loose Stator Windings

If the electrical windings of the motor stator are even a little loose, the vibration level at twice line frequency will be increased. This condition is very destructive because it abrades the insulation on the wire, leading to shorted turns and eventual short circuits to ground and stator failure.

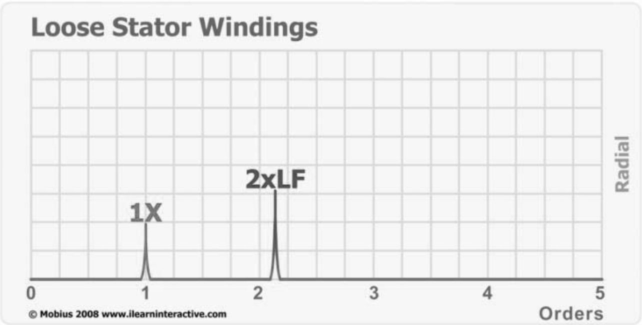


Figure 12-25 - Loose stator windings increase the 2x line frequency peak.

Lamination Problems

The rotor and stator of AC motors are made of thin laminations that are isolated from each other. If the laminations are shorted together, local heating and resultant thermal warping will occur.

Shorted laminations also cause higher twice line frequency vibration levels.

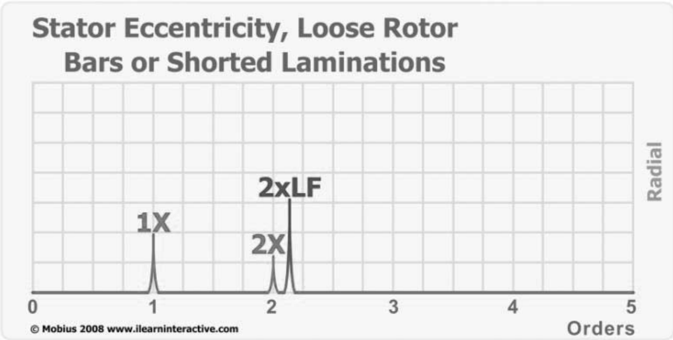


Figure 12-26

The warping can cause the 1X level to increase, and often pole pass sidebands are observed.

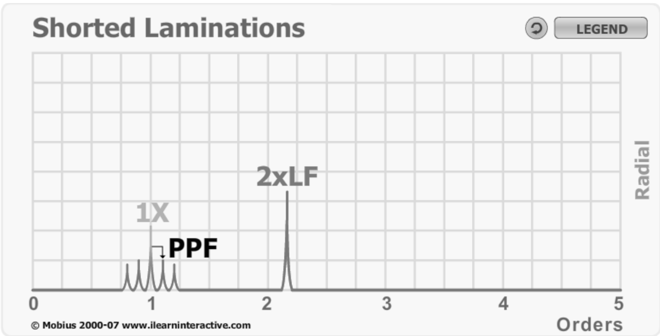


Figure 12-27 - Shorted Laminations

Loose Connections

Phasing problems due to loose connectors can cause excessive vibration at twice line frequency (100 or 120 Hz), with sidebands of one-third line frequency (16.67 or 20 Hz).

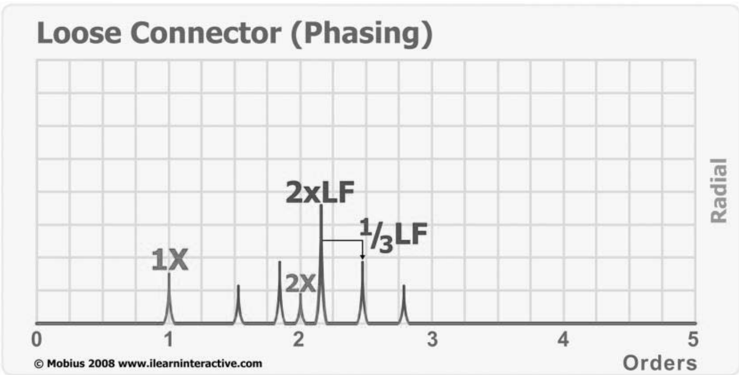


Figure 12-28 - Loose connections cause phasing problems

Motor Current Analysis

A test to check the condition of rotor bars that has become very common is to measure the motor current (just one phase is required) and generate a high-resolution spectrum. A portable current clamp or current transformer is used. The output is connected to the data collector (or spectrum analyzer) as usual (see your data collector manual for set up requirements).

Figure 12-29 - Current clamp

If levels are high a step-down current transformer (CT) may be required. A high resolution measurement is required centering around line frequency if possible. Measurements are often made at the motor control center.

It is better than vibration because the Line Frequency amplitude can accurately be compared to the sidebands for a more accurate indication of broken rotor bars. Vibration is influenced by imbalance and other sources but a current measurement is current only.

Some challenges are that a high resolution measurement is needed and it can be difficult to determine slip frequency if not at the machine.

The high resolution measurement is to see pole pass frequency sidebands around line frequency (50 or 60 Hz). Use 1600 lines and a frequency range of 0-70 Hz.

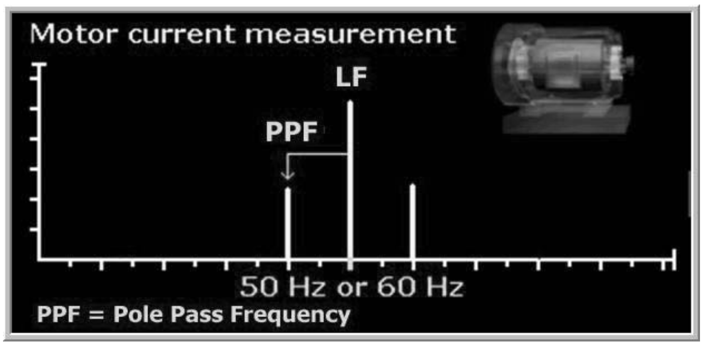


Figure 12-30

The following is a sample of real data.

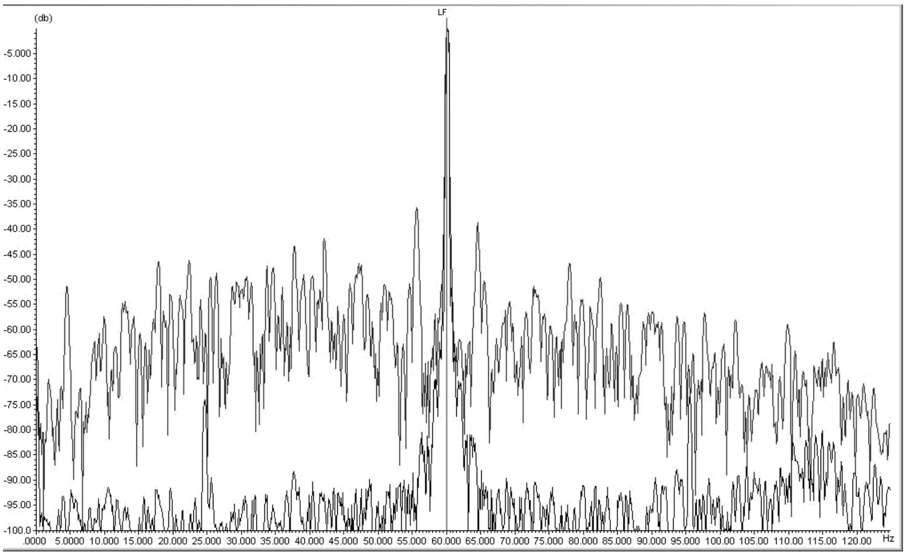


Figure 12-31

The goal is to compare the amplitude level of peaks at sidebands of the pole pass frequency (# poles x slip frequency), and compare those to the level of the line frequency (not twice line frequency).

For example, if the synchronous speed was 1800 RPM, and the actual running speed was 1760 CPM, the slip frequency will be 40 CPM, or 0.667 Hz. In this example (in the US), we would examine the peak at 60 Hz, and at 58.666 Hz and 61.334 Hz.

The following table provides a guide as to the relationship between the height of the line frequency peak and the height of the sideband.

-dB	Rotor Condition Assessment	Recommended Action
>60	Excellent	None
54 – 60	Good	None
48 – 54	Moderate	Trend Condition
42 – 48	High Resistant Connection or Cracked Bars	Increase Test Frequency and Trend
36 – 42	Broken Rotor Bars Will Show in Vibration	Confirm with Vibration, Plan Repair / Replace
30 – 36	Multiple Cracked/Broken Bars, Poss Slip Ring Problems	Repair/Replace ASAP
<30	Severe Rotor Faults	Repair/Replace Immediately

Table 12-1

Chapter 13

Pumps, Fans and Compressors

Objectives:

- Calculate Forcing Frequencies associated with pumps, fans, compressors
- Diagnose Vane Pass in spectral data
- Identify Cavitation in spectral data
- Describe the characteristics of Cavitation in Waveform Data
- Identify air flow problems in spectral data

Pumps, Fans, and Compressors

Pumps, fans and compressors can all suffer from the classic fault conditions such as imbalance, misalignment, bearing faults, and looseness. However, some faults are unique to these components.

There are many types of pumps in common use, and their vibration signatures vary over a wide range. When monitoring pump vibration, it is important that the operating conditions are consistent from one measurement to the next to assure consistent vibration patterns. Suction pressure, discharge pressure, and especially air induction and cavitation will affect the vibration pattern greatly.

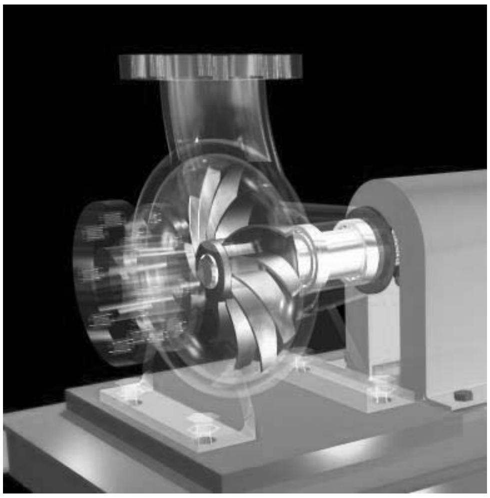


Figure 13-1 - Pumps, fans, and compressors have some unique vibration characteristics.

Centrifugal pumps always have a prominent vibration component at the vane pass frequency, which is the number of impeller vanes times the RPM. If the amplitude increases significantly, it usually means an internal problem such as erosion of impellers or a flow related problem, or possibly misalignment. Harmonics of vane pass are also common in such pumps.

Forces are generated by the pressure variations as a rotating blade, vane, or lobe passes a stationary housing or component.

The stationary component creates a non-uniform flow disturbance in the fluid or gas.

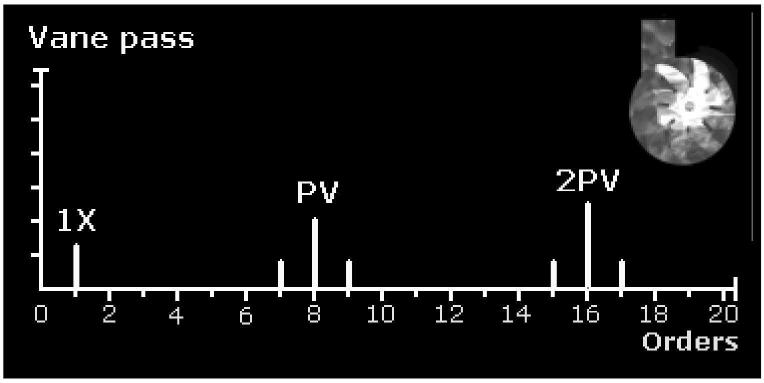


Figure 13-2 - Vane pass is common in pumps

Blade Passing Frequency

In its broadest sense, Blade Passing Frequency includes similar components such as vanes, lobes, and pistons. The forcing frequencies are:

- Number of blades x turning speed
- Number of vanes x turning speed
- Number of lobes x turning speed
- Number of pistons x turning speed

Causes of high blade passing frequency are:

- Rotor or housing eccentricity
- Non-uniform variable pitch blades
- Loose, bent, or misaligned housing diffuser vanes
- Blade or vane wear (abrasion or cavitation)
- Operation (improper performance parameters)
- Improper damping settings
- Dirty, damaged or missing filters
- Inlet or discharge line restrictions

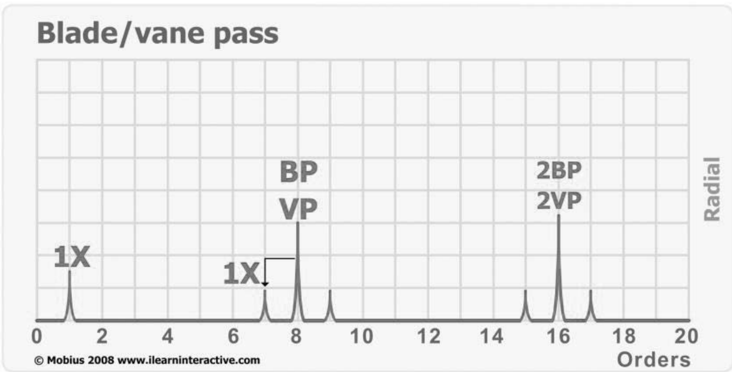


Figure 13-3 - Blade/vane pass and harmonics

Other faults include the impeller loose on the shaft and pump starvation.

When the impeller is loose on the shaft, there will be vane pass frequency with sidebands of turning speed.

Pump starvation may look like imbalance as the flow into the volute is uneven. The time waveform would show distortion due to the flow being uneven. The distorted time waveform produces harmonics of turning speed.

Cavitation

Cavitation normally creates random, higher frequency vibration or "noise". It is often observed as a "hump" in the vibration spectrum raising the floor from about 15X to 35X.

Cavitation normally indicates insufficient suction pressure or starvation - i.e. low inlet pressure. The liquid tends to vaporize while coming off the impeller creating vacuum bubbles that implode. The waveform can often sound like gravel in the pump.

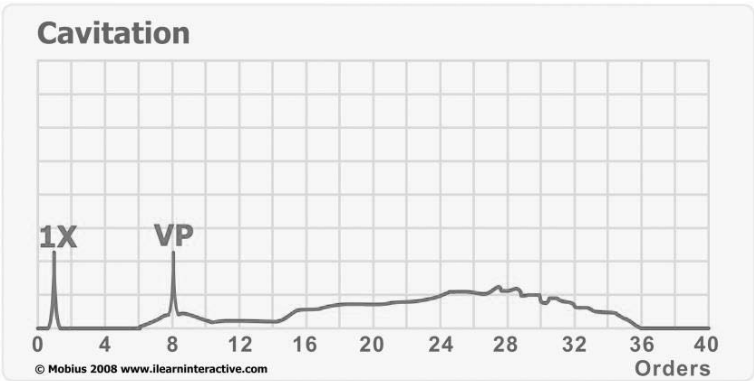


Figure 13-4 - Cavitation often has a hump of energy in the floor

The data in **Error! Reference source not found.** is from a 20 hp centrifugal pump that has a cavitation problem. It sounds like it is pumping gravel. The time waveform is a very useful analysis tool, as the high frequency bursts of energy are often clearly visible - however, more time data (i.e. a longer record) than normal may be needed.

The spectrum reflects the impacting in the raised noise floor and mounds of energy under the peaks. The highest peak is at vane pass.

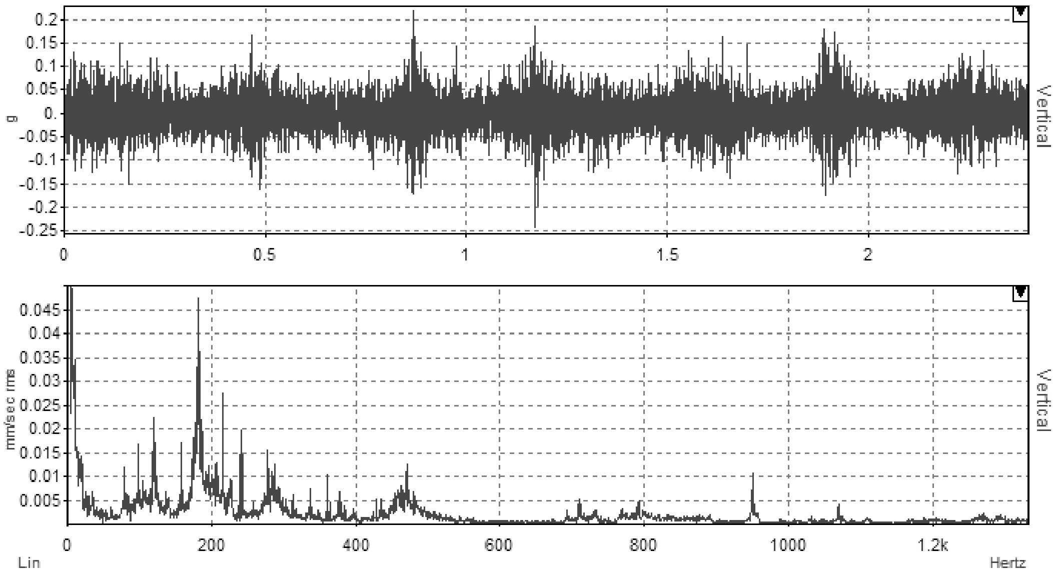


Figure 13-5 - Waveform has random bursts of energy. Spectrum has raised noise floor throughout

Turbulence

Turbulence in fans, pumps, and compressors is induced when the normal flow is restricted or impeded during operation.

It shows up in spectral data as low frequency, broad peaks between 0.5 and 0.8 orders.

Figure 13-6 - Flow turbulence in fans and pumps

Causes of turbulence are:

- Obstruction in air ducts or plumbing lines
- Sharp radius turns in piping
- Abrupt diameter changes in lines
- Over capacity operation.

Harmonics

Lobed blowers and screw compressors generate harmonics under normal conditions.

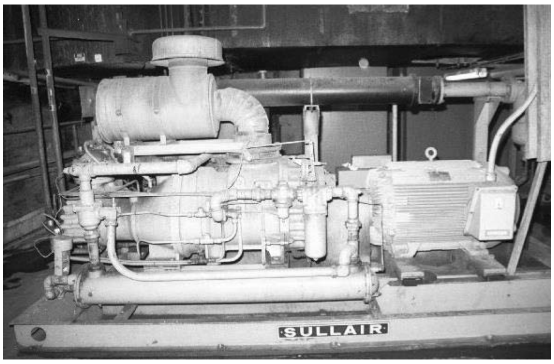


Figure 13-7

Sullair Rotary Screw Compressor, 1650 CFM

Speed is 1780 CPM

300 HP motor

Female rotor has 6 vanes and the male has 4 vanes.

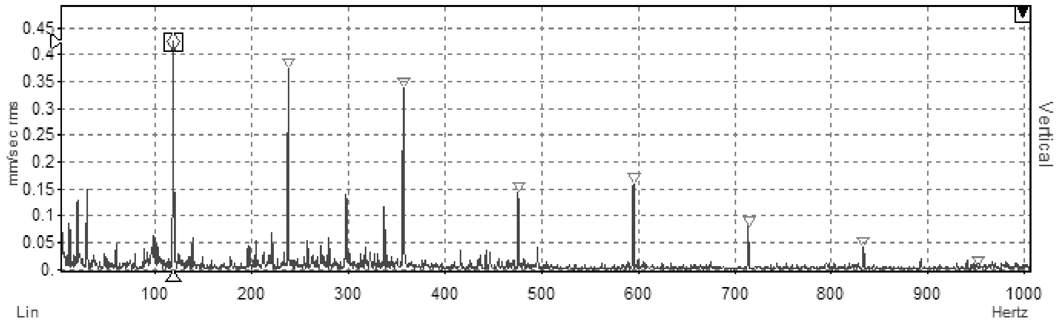
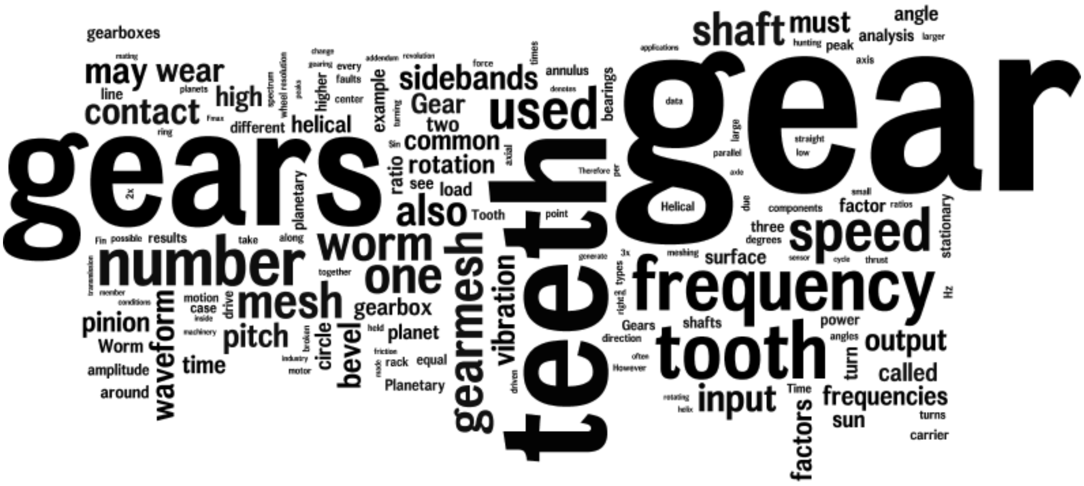


Figure 13-8



Chapter 14

Gearbox Analysis

Objectives:

- Review spur, bevel, helical, herringbone, worm and planetary gears
- Calculate gearmesh frequency, gear assembly phase frequency and hunting tooth frequency
- Understand the benefits of time waveform analysis and time synchronous averaging
- Learn how to set up an analyzer for the optimum waveform
- Diagnose a broken tooth, high tooth load, gear misalignment and other fault conditions

Gearboxes

Gearboxes are typically one of the most important pieces of rotating machinery in your plant. They are also one of the most challenging types of equipment to monitor. Many analysts take a simple approach; if the vibration levels increase in the spectrum then they call that the gearbox needs to be repaired.

There is a great deal you can learn with time waveform analysis, coupled with some knowledge about vibration patterns. In this section we will discuss how you can correctly test gearboxes and diagnose a wide range of faults.

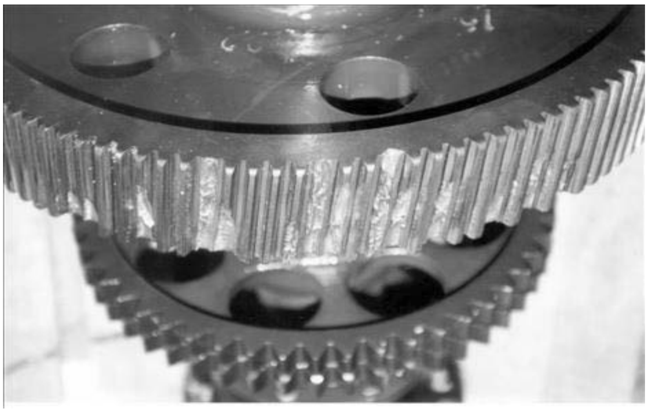


Figure 14-1 Gear with broken teeth

There are a number of reasons a gearbox may fail, and thus a number of fault conditions that we must be able to detect. Here is a partial list:

- Tooth wear
- Tooth load
- Gear eccentricity
- Backlash
- Gear misalignment
- Broken or cracked teeth
- …. And others

A recent study showed that 60% of failures could be attributed to lubrication skills and practical issues.

Understanding gearboxes

Before we get started on the diagnostic section, we need to be familiar with the terms associated with gearboxes.

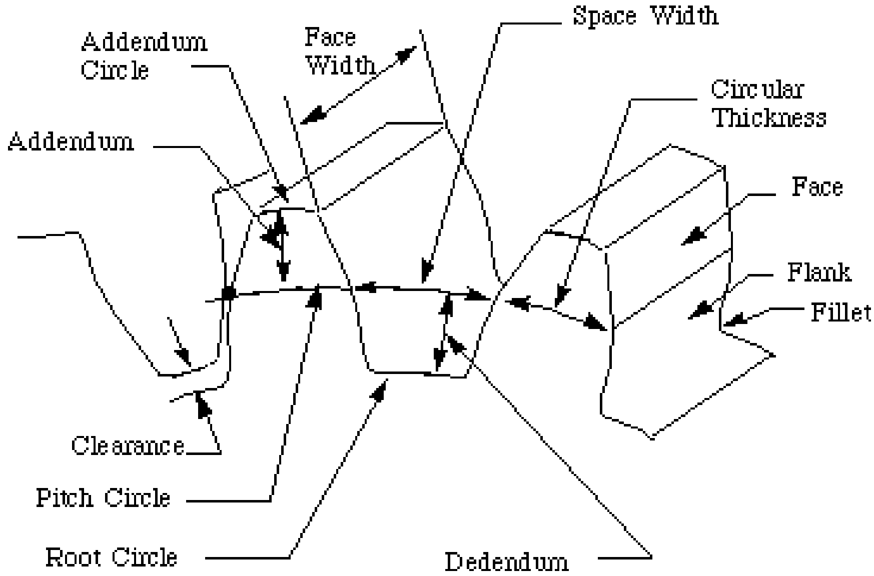


Figure 14-2 Annotated gear

Pitch surface : The surface of the imaginary rolling cylinder (cone, etc.) that the toothed gear may be considered to replace.

Pitch circle: A right section of the pitch surface.

Addendum circle: A circle bounding the ends of the teeth, in a right section of the gear.

Root (or dedendum) circle: The circle bounding the spaces between the teeth, in a right section of the gear.

Addendum: The radial distance between the pitch circle and the addendum circle.

Dedendum: The radial distance between the pitch circle and the root circle.

Clearance: The difference between the dedendum of one gear and the addendum of the mating gear.

Face of a tooth: That part of the tooth surface lying outside the pitch surface.

Flank of a tooth: The part of the tooth surface lying inside the pitch surface.

Circular thickness (also called the **tooth thickness**) : The thickness of the tooth measured on the pitch circle. It is the length of an arc and not the length of a straight line.

Tooth space: The distance between adjacent teeth measured on the pitch circle.

Backlash: The difference between the circle thickness of one gear and the tooth space of the mating gear.

Circular pitch p: The width of a tooth and a space, measured on the pitch circle.

Diametral pitch P: The number of teeth of a gear per inch of its pitch diameter. A toothed gear must have an integral number of teeth. The *circular pitch*, therefore, equals the pitch circumference divided by the number of teeth. The *diametral pitch* is, by definition, the number of teeth divided by the *pitch diameter*.

Correct gear mesh

A pinion tooth touches a wheel tooth at the red contact point (the knot) which moves up the line of action and along the teeth faces as rotation proceeds. Since contact cannot occur outside the teeth, it takes place along the line of action only between the points indicated by the black dots on the line of action and inside both addendum circles. The line between the black dots is named the ***path of contact.***

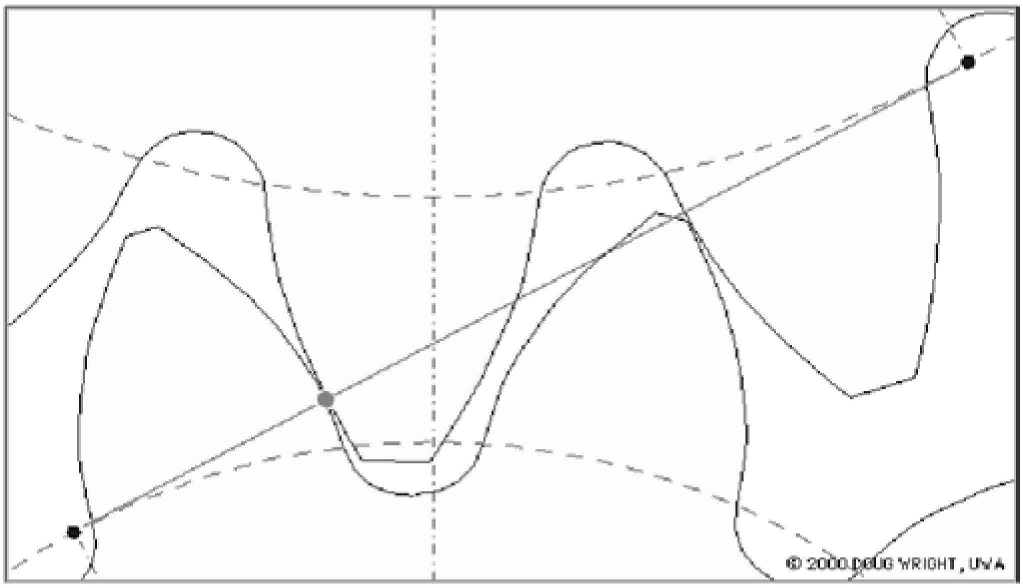


Figure 14-3 Perfect gear mesh

When teeth are meshing correctly, one tooth will be in mesh at all times, with a second tooth going into mesh before the first tooth goes out of mesh. It is written that in some designs, up to three teeth mash be in contact at any time; one going out of mesh, on in mesh, and the third coming into mesh.

If the gears are correctly aligned, and the gears are not eccentric, and the shafts are not bend, and the bearings are not loose, every tooth mesh will be equal. The vibration pattern will indicate a consistent pattern over every cycle of rotation.

But if these conditions did exist, the vibration patterns would have to change. If the gears moved closer together and then further away with each rotation, the vibration would rise and fall with each cycle.

It is expected that the surface of the teeth is smooth and well lubricated. The contact should be smooth and consistent.

But if there is wear, pitting, dents or cracks, the vibration will not be consistent. The contact may be rough and inconsistent. There may be just one point in the rotation when the cracked or broken tooth comes into the mesh – it may be only that time when the vibration changes. In fact, as we will discover, there may be a situation where the vibration changes only when one tooth on one gear comes into contact with a unique tooth on the other gear.

A number of unusual conditions can exist which we will explore in this module.

Gear types

There are a number of different types of gears. Each has a specific purpose. For the vibration analyst it is helpful to know what type of gear is in the gearbox because the forces generated are different. For example, spur gears generate a radial force, whereas helical gears generate an axial force.

Spur gears

Spur gears connect parallel shafts, have involute teeth that are parallel to the shaft and can have internal or external teeth. They cause no external thrust between gears. They are inexpensive to manufacture. They give lower but satisfactory performance. They are used when shafts rotate in the same plane.

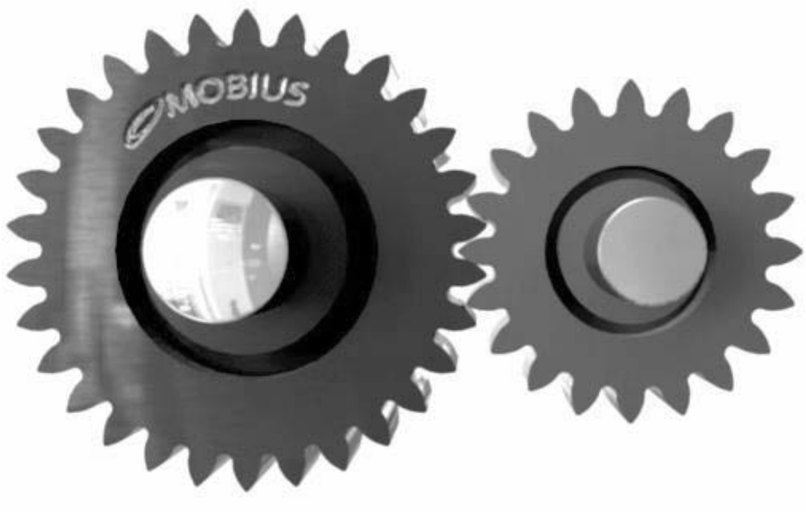


Figure 14-4 Spur gears

The main features of spur gears are dedendum, addendum, flank, and fillet. Dedendum cylinder is a root from where teeth extend, it extends to the tip called the addendum circle. Flank or the face contacts the meshing gear, the most useful feature if the spur gears. The fillet in the root region is kinetically irrelevant.

Characteristics

The speed and change of the force depends on the gear ratio, the ratio of number of teeth on the gears that are to be meshed. One gear among the two is on the input axle, the axle of the motor and the other gear of the pair is on the output axle, the axle of the wheel.

They have higher contact ratio that makes them smooth and quiet in operation. They are available for corrosion resistant operation. They are among the most cost-effective type of gearing. They are also used to create large gear reductions.

Helical gears

Helical gears connect parallel shifts but the involute teeth are cut at an angle to the axis of rotation. Two mating helical gears must have equal helix angle but opposite hand. They run smoother and more quietly. They have higher load capacity, are more expensive to manufacture and create axial thrust.

Figure 14-5 Helical gears

Helical gears can be used to mesh two shafts that are not parallel and can also be used in a crossed gear mesh connecting two perpendicular shafts. They have longer and strong teeth. They can carry heavy load because of the greater surface contact with the teeth. The efficiency is also reduced because of longer surface contact. The gearing is quieter with less vibration.

Gear Configuration

They can be manufactured in both right-handed and left-handed configurations with a helix angle to transmit motion and power between non-intersecting shafts that are parallel or at 90 degrees to each other. For shaft at 90 degrees, the same helix angles are used and the tooth contact area of the gear is very small. If the angle of gear teeth is correct, they can be mounted on perpendicular shaft by adjusting the rotating angle by 90 degrees. The inclination of the teeth generates an axial force. As the angle of inclination increases the axial force also increases. Thrust bearings can counter these forces.

Applications

These are highly used in transmission because they are quieter even at higher speed and are durable. The other possible applications of helical gears are in textile industry, blowers, feeders, rubber and plastic industry, sand mullers, screen, sugar industry, rolling mills, food industry, elevators, conveyors, cutters, clay working machinery, compressors, cane knives and in oil industry.

Disadvantages

A disadvantage of helical gears is the resultant thrust along the axis of the gear, which needs to be accommodated by appropriate thrust bearings. This can be overcome by the use of double helical gears by having teeth with a 'v' shape.

Helical Bevel Gears

Helical bevel gear is a toothed gear in angular design. The input side is provided with a motor flange or a free input shaft and the output side is provided with a free shaft end or a hollow shaft. Helical bevel gears are fitted with flanges of various sizes. Reciprocating tools cuts them.

The advantages of helical bevel gears are high efficiency and low reduction rate. The use of helical bevel gear saves energy and cost. Helical bevel gears are manufactured by an alloyed case hardening steel. The gear material is given an extremely strong, homogeneous structure.

They can replace worm gears in a variety of applications, particularly in modular machinery. They are also used as storage and retrieval unit. They are commonly used in modern differentials.

Herringbone or double helical gears

Herringbone or double helical gears conduct power and motion between non-intersecting, parallel axis that may or may not have center groove with each group making two opposite helices. The two helix angles come together in the center of the gear face to form a 'V'. In these gears the end thrust forces cancel themselves out. It is difficult to cut this type of gear but it is made easier by machining a groove in the face at the point of the apex of the 'V' creating a break in the middle of the herringbone gear teeth. They do not have any separating groove between the mirrored halves.

Figure 14-6 Herringbone gears

Action is equal in force and friction on both gears and all bearings. Herringbone gear also allow for the use of larger diameter shaft for the same volumetric displacement and higher differential pressure capability.

The most common application is in power transmission. They utilize curved teeth for efficient, high capacity power transmission. This offers reduced pulsation due to which they are highly used for extrusion and polymerization. Herringbone gears are mostly used on heavy machinery.

Bevel gears

They connect intersecting axes and come in several types. The pitch surface of bevel gears is a cone. They are useful when the direction of a shaft's rotation needs to be changed. Using gears of differing numbers of teeth can change the speed of rotation. They are usually mounted on shafts that are 90 degrees apart, but can be designed to work at other angles as well.

These gears permit minor adjustment during assembly and allow for some displacement due to deflection under operating loads without concentrating the load on the end of the tooth. For reliable performance, Gears must be pinned to shaft with a dowel or taper pin. Bevel gear sets consist of two gears of different pitch diameter that yield ratios greater than 1:1.

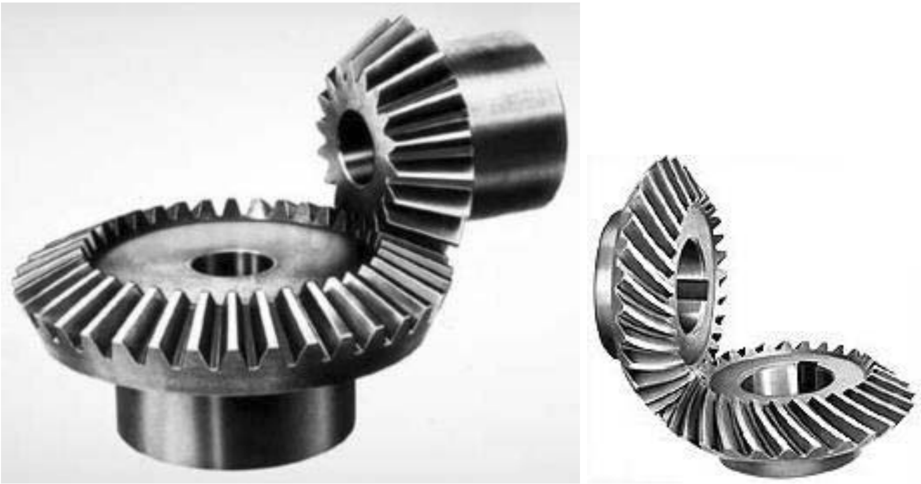


Figure 14-7 Bevel gears

Types

The teeth on bevel gears can be straight, spiral or bevel. In straight bevel gears teeth have no helix angles. They either have equal size gears with 90 degrees shaft angle or a shaft angle other than 90 degrees. Straight bevel angle can also be with one gear flat with a pitch angle of 90 degrees. In straight when each tooth engages it impacts the corresponding tooth and simply curving the gear teeth can solve the problem.

Spiral bevel gears have spiral angles, which gives performance improvements. The contact between the teeth starts at one end of the gear and then spreads across the whole tooth. In both the bevel types of gears the shaft must be perpendicular to each other and must be in the same plane.

The hypoid bevel gears can engage with the axes in different planes. This is used in many car differentials. The ring gear of the differential and the input pinion gear are both hypoid. This allows input pinion to be mounted lower than the axis of the ring gear. Hypoid gears are stronger, operate more quietly and can be used for higher reduction ratios. They also have sliding action along the teeth, potentially reducing efficiency.

Applications

A good example of bevel gears is seen as the main mechanism for a hand drill. As the handle of the drill is turned in a vertical direction, the bevel gears change the rotation of the chuck to a horizontal rotation. The bevel gears in a hand drill have the added advantage of increasing the speed of rotation of the chuck and this makes it possible to drill a range of materials.

The bevel gears find its application in locomotives, marine applications, automobiles, printing presses, cooling towers, power plants, steel plants, defense and also in railway track inspection machines. They are important components on all current rotorcraft drive system.

Spiral bevel gears are important components on all current rotorcraft drive systems. These components are required to operate at high speeds, high loads, and for an extremely large number of load cycles. In this application, spiral bevel gears are used to redirect the shaft from the horizontal gas turbine engine to the vertical rotor.

Worm gears

A worm gear is an inclined plane wrapped around a central axle. It is a gear with one or more teeth in the form of screwed threads.

Figure 14-8 Worm gears

Worm gears are made of two parts: the pinion and the worm gear. The pinion has small number of teeth and they wrap around the pitch cylinder. The worm gear has concave faces to fit the curvature of the worm in order to provide line of contact instead of point of contact. They are cut helically for better mating Worm gears can provide a high angular velocity between non-intersecting shafts at right angles. They are capable of transmitting high tooth loads, the only disadvantage is the high sliding velocities across the teeth. They provide ultimate power ratio.

Features

The efficiency of worm gear depends on the lead angle, sliding speed, and lubricant, surface quality and installation conditions. They offer smoothest, quietest form of gearing. They provide high-ratio speed reduction in minimal spaces.

Worm gears are used when large gear reductions are required. Worm gear has a unique property of easily turning the gear. The gear cannot turn the worm because the angle on the worm is shallow and when the gear tries to spin the worm, the friction between the two holds the worm in place.

Worm gears work under difficult conditions, presenting unique lubrication demands. The types of oils most commonly used to lubricate worm gears are compounded mineral oils, EP mineral gear oils and synthetics.

Operation of the worm gears

Worm gears are always used as the input gear. For the operation of worm gear, torque is applied to the input end of the worm shaft by a driven sprocket or electric motor. The worm and the worm shaft are supported by anti-friction roller bearings. Because of high friction worm gears are very inefficient. There is lot of friction between a worm gear and the gear being driven by the worm gear. When used in high torque applications, the friction causes the wear on the gear teeth and erosion of restraining surface.

Types

There are three types of worm gears:

1. **Non throated**- a helical gear with a straight worm. Tooth contact is a single moving point on the worm drive.

2. **Single throated**- has concave helical teeth wrap around the worm. This leads to line contact.

3. **Double throated**- called a cone or hourglass. It has concave teeth both on the worm and helical gear.

Applications

Worm gears are widely used in packaging machinery, material handling, machine tools, indexing and food processing. They are used widely in conveyor systems. They are also used in torsen differential, used on some high-performance cars and trucks. They serve as speed reducers in many different industries.

Rack and pinion

Rack Gears

A rack gear is a toothed bar into which a pinion meshes. Racks are gears of infinite pitch radius. They are used to translate rotary motion to linear motion or vice versa. They will mesh with pinions of the same pitch.

Racks are made of various materials. The commonly used materials for racks are stainless steel, brass, and plastic.

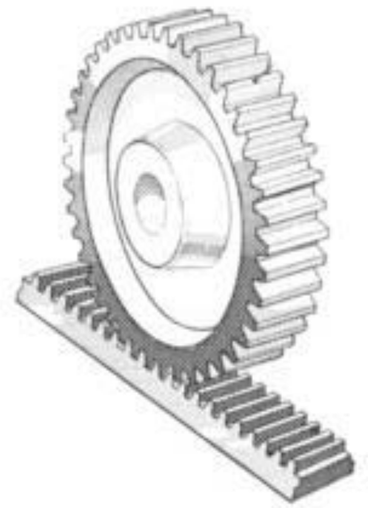


Figure 14-9 Rack and pinion gears

They are widely used in automobiles. The steering wheel of a car rotates the gear that engages the rack. The rack slides right or left, when the gear turns, depending on the way we turn the wheel. Windshield wipers in cars are powered by a rack and pinion mechanism. They are also used in some scales to turn the dial that displays weight.

Pinion Gears

It is a small cogwheel. The teeth fit into a larger gear wheel. Rotational motion is converted into linear motion when the pinion turns and moves the rack. Pinion gears are engineered to be the best gears.

Pinion gear system involves the use of a small round gear called pinion and a large flat gear called rack, more the number of teeth in the pinion gear, more is the speed of rotation. Pinion with smaller number of teeth produces more torque. Pinion is attached to the motor shaft with glue. Rotation of pinion is done by rotation of pinion about a fixed center that helps the rack to move in the straight line. If the rack is moved and the pinion rotates then the center of the pinion moves taking along the pinion with it.

Planetary gears

These are a set of gears usually two or more on or inside a larger gear. They make drastic gear ratio possible. They are used when one wishes to turn the input in the same direction as the output. The gear in the center of the larger gear, called the sun engages two or three smaller gears placed in the same large gear. These small gears are called planet and they are engaged to the inside of the large gear called the ring. Planet gears turn on movable center and the sun gears turn on a fixed center.

Figure 14-10 Planetary gears

Gearing Mechanism

In a planetary gearing, the member which receives motion from outside the mechanism is called the driver; the member from which motion is taken outside the mechanism is called the follower; the member which carries one or more bearing pins about which the planet gears rotate is called the train arm. There is only one member that is maintained in a fixed position. Planetary gears are also used to produce different gear ratios depending on the which gear is used as input, which one as output and which one is held stationary.

Applications

An automatic transmission uses planetary gearsets to create different gear ratios using clutches and brake band to hold different parts of the gearset stationary and change the input and the output.

Planetary gears are most commonly used gear train. Planetary gear trains have several advantages. They have higher gear ratios. They are popular for automatic transmissions in automobiles. They are also used in bicycles for controlling power of pedaling automatically or manually. They are also used for power train between internal combustion engine and an electric motor, and they are used in wind turbines.

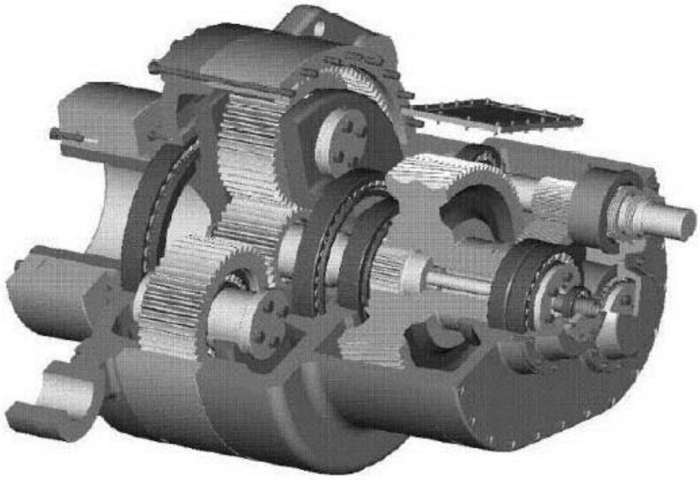


Figure 14-11 Wind turbine planetary gearbox

Forcing frequency calculations

There are three key frequencies involved, the input speed, the frequency of the gearmesh, which is the number of teeth multiplied by the speed of the shaft, and the output speed.

$$Gear\ mesh = Number\ of\ teeth\ \times Shaft\ speed$$

$$Output\ speed = Input\ speed \times \frac{Input\ teeth}{Output\ teeth}$$

Worm gear forcing frequencies

Worm gears are often confusing because there is sometimes a question as to how many teeth are on the input worm drive. In the case of a worm gear, it is not the number of teeth that is of concern (often a worm drive only has one tooth) but the number of *flights* on the worm gear.

Figure 14-12 Worm gear

The flights refer to the number of teeth that mesh with the driven gear during one revolution of the worm drive. This can be readily identified if the output gear speed, the number of teeth on the output gear and the input shaft speed are known. In this example an output drive gear with 24 teeth turning at 10 Hz is driven by a worm gear turning at 29.5 Hz. The number of flights (#F) on the input gear can be determined as follows:

GM = To x So = Fin x Sin

Fin = (To x So) / Sin

If you know the input speed [Sin], output speed [So] and the number of teeth on the output shaft [To] (i.e. the number of teeth on the worm wheel) then you can compute the number of flights [Fin].

Example:

Input speed [Sin]: 20 Hz

Teeth on worm wheel [To]: 30

Output speed [So]: 5 Hz

GM = To x So = Fin x Sin

Fin = (To x So) / Sin

Therefore: Fin = 30 x 5 / 20 = 7.5

It has 7.5 flights (7.5 *teeth that mesh with the driven gear during one revolution of the worm drive).*

Gear wear and common factors

Gear assembly phase problems are a gear wear pattern induced fault which results in the generation of fractional sub-harmonics of GMF. Gear assembly phase problems can originate in manufacturing problems in the gears, damage to the gears from contamination traveling through the mesh, or from re-orienting the gear teeth during overhaul.

One additional piece of information that we can gain from the tooth count is the type of wear that the gears will experience. We can "factor" the number of teeth on each gear and determine the number of common factors. If two meshing gears had prime numbers of teeth, for example 31 teeth and 19 teeth, then each tooth on the each gear must come into contact with every other tooth before it contacts the same tooth again. This ensures relatively even wear on all of the teeth.

Prime numbers: 1, 3, 5, 7, 11, 13, 17, 19…

If the number of teeth is not a prime number, then the wear pattern can be less than ideal. First we must determine the factors that make up each gear tooth count, and then find the common factors between each gear.

For example, if a gear has 33 teeth, the factors (sometimes called "prime factors") are 1, 3 and 11 (1 x 3 x 11 = 33). The largest factor is 11. If the other gear has 21 teeth its factors are 1, 3, and 7 (1 x 3 x 7 = 21). The largest factor is 7.

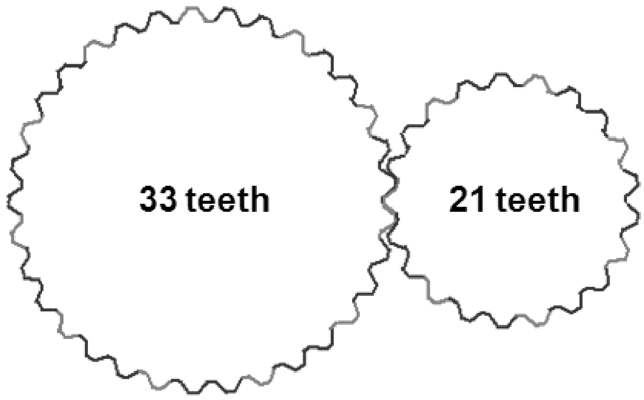


Figure 14-13 Example of common factors

The factors of 1 and 3 are common to both gears – so they are our "common factors". When the common factor is not one, the teeth will wear unevenly. A common factor of three will cause a tooth on one gear to mesh with every third tooth on the other gear. We will see this in the spectrum at a frequency of the gearmesh divided by the common factor.

However, if there is more than one common factor, we must multiply them together. For example, if a gear has 30 teeth, the factors (sometimes called "prime factors") are 1, 2, 3 and 5 (1 x 2 x 3 x 5 = 30). The largest factor is 5. If the other gear has 18 teeth its factors are 1, 2, and 3

(1 x 2 x 3 x 3 = 18). The largest factor is 3. In our example the product of these common factors is 6 (1 x 2 x 3 = 6). Notice that every 6th tooth comes into contact.

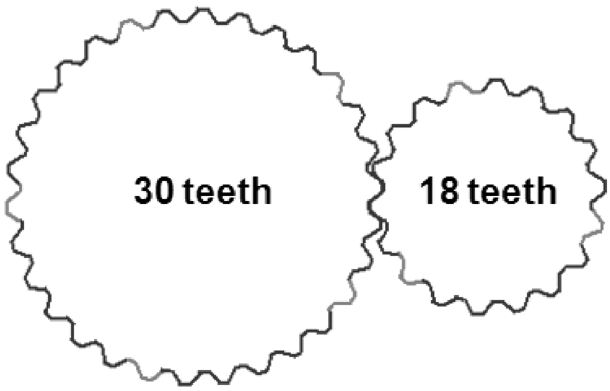


Figure 14-14 Example of common factors

The frequency generated in this example is the gearmesh frequency divided by the product of the common factors, 6 in this example. It is called the "gear assemble phase frequency". Another way to look at it is that the frequency is five times the speed of the driving gear (5 is the highest factor), and 3 times the speed of the driven gear (3 is the highest factor on that gear).

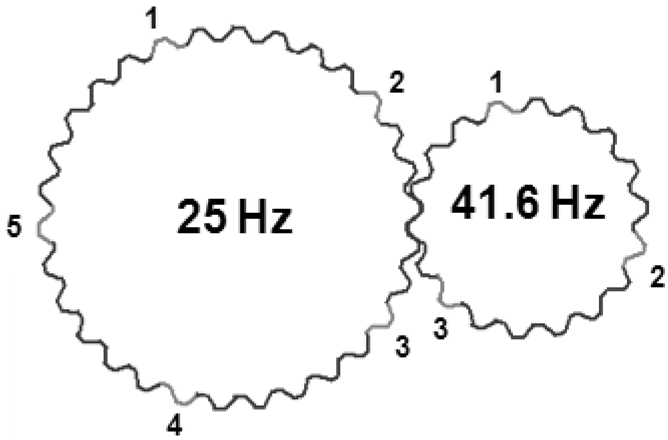


Figure 14-15 Example of common factors

Hunting tooth frequency

The uncommon factor identifies the number of similarly defective teeth on the gear and the number of revolutions the other gear must make before the same teeth mesh again. The frequency of this event is called the hunting tooth frequency. We can calculate the hunting tooth frequency a number of ways.

HTF = GMF x Na / (Tin x Tout)   Na = product of common factors

HTF = *Input speed / CF inner*

HTF = Output speed / CF outer

The so-called "hunting tooth frequency" is the rate at which a tooth in one gear mates with a particular tooth in the other gear. If the gear ratio is an integer such a 1, 2 or 3, the hunting tooth frequency will be the RPM of the larger gear, and the same teeth will be in contact once per revolution. This will cause uneven wear on the gears - a small defect in one tooth will repeatedly contact the same teeth in the other gear causing localized wear on those teeth.

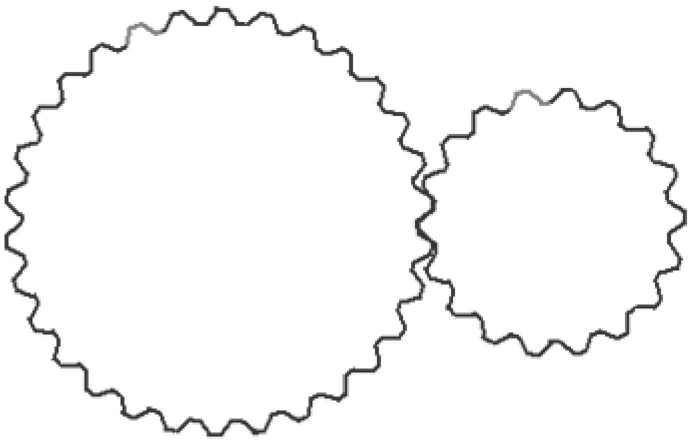


Figure 14-16 Gear with single tooth highlighted

For this reason, gearboxes are not made with these simple ratios unless absolutely necessary. Ideally, the hunting tooth frequency should be as low as possible to evenly distribute the wear around both gears. This means the number of teeth on each gear should be a prime number.

In practice, the hunting tooth frequency is used to detect faults on both the gear and pinion that may have occurred during manufacturing or as a result of mishandling. It is typically a low frequency, and you can hear a "growling" sound from the gearbox.

Ghost frequencies

Sometimes the vibration spectrum of a gearbox will contain components which cannot be related to any known geometry of the gearbox. These are called "ghost frequencies", and are caused by irregularities machined into the gears in the manufacturing process. Ghost components are independent of loading, and tend to disappear as the gears wear.

Vibration analysis

There will normally be peaks at the shaft speeds and gearmesh frequency; however they will be low level. There may be a 2X gearmesh peak, and there may be sidebands of shaft speed around the gear mesh frequency.

Some faults also show up at 3x gearmesh frequency. It is therefore necessary to set the Fmax high enough to see 3x gearmesh and sidebands.

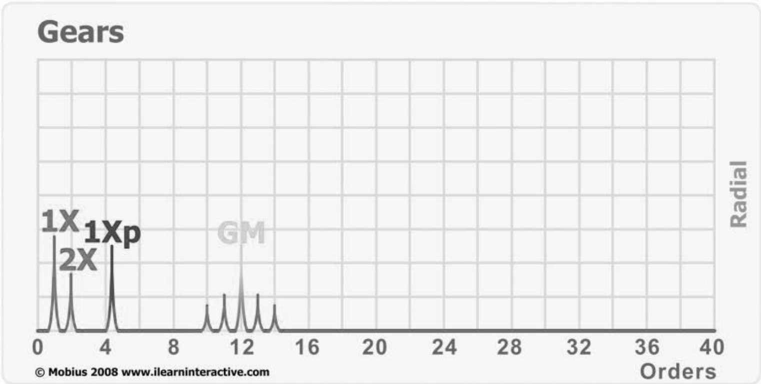


Figure 14-17 - Gears generate peaks of shaft speeds and gearmesh frequency.

These frequencies will be most prominent in the radial direction for spur gears, and in the axial direction for helical gears.

Most faults are detected by studying the gearmesh frequency and 2x and 3x gearmesh frequencies along with their sidebands. The following is a partial list of faults that can be detected through this study.

- Tooth wear
- Tooth load
- Gear eccentricity
- Backlash
- Gear misalignment
- Broken or cracked teeth

It is beyond the scope of the course to cover these in greater detail.

Waveforms and Gear Analysis

The waveform is a VERY powerful tool when attempting to diagnose gear faults.

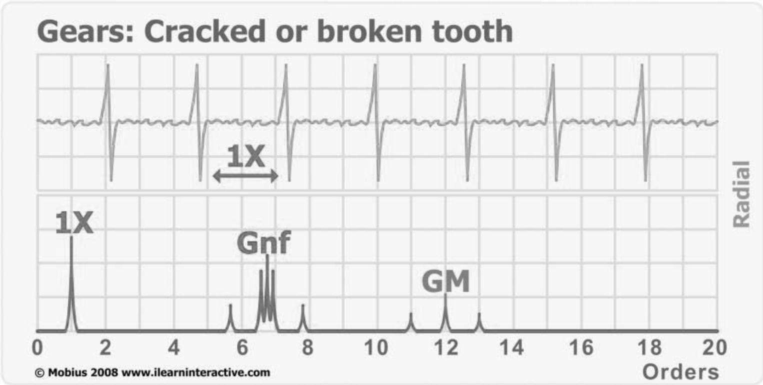


Figure 14-18 - A cracked or broken tooth will show up in the waveform as a pulse at gear rpm.

As each tooth meshes there is a pulse in the waveform. The teeth can typically be counted in the time waveform. Depending upon the nature of the fault, one of the impacts/pulses per cycle may be at a higher amplitude; or lower amplitude (if it were missing). See **Error! Reference source not found.**

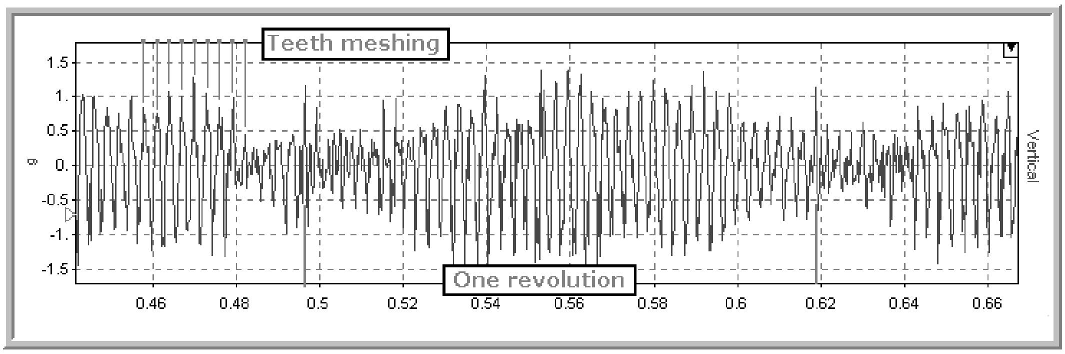


Figure 14-19 - Each impact relates to two teeth meshing.

The waveform data in **Error! Reference source not found.** shows impacts as each gear meshes. The variance in amplitude is due to the gears being out of alignment, causing varying loads and therefore varying amplitudes. The pulses relate to the contact between two teeth. This waveform also shows the cycle that relates to a complete revolution of the shaft.

The time waveform will be largely sinusoidal with limited modulation. In healthy gearboxes there should be no distinct transients.

A rule of thumb is to collect 6-10 rotations of each shaft with adequate resolution to see sidebands and detail.

Time synchronous averaging

Time synchronous averaging (TSA) is commonly used when attempting to diagnose gear faults. TSA will average away all of the vibration sources that are not synchronous with the tach pulse, which is taken from the shaft of the gear of interest.

This means that other sources of vibration, from bearings, the motor, resonances, and so on, are removed, leaving a clean time waveform. The TSA is time consuming (a large number of averages is required), and it is time consuming to set up in the first place. However the results are worth all the effort.

Tooth Wear

When teeth begin to wear, two things will happen. The first is that the sidebands of gear mesh will increase in level. The sidebands will correspond to the speed of the gear with the wear. The sidebands will develop as a result of amplitude or frequency modulation.

Figure 14-20 - Tooth wear

The second thing that occurs is that the natural frequency of the gear is excited due to the impacting of the gears meshing. This peak will also exhibit sidebands, and as a natural frequency, it is likely to have a broader base.

The 3x gearmesh frequency will increase in amplitude and multiple sidebands will appear.

Figure 14-21 Very badly worn gear

Tooth Load

The amplitude of the tooth mesh frequency is dependent on the alignment of the shafts carrying the gears, and the load on the gear. The higher the load, the higher the amplitude. A high peak at the gear mesh frequency does not necessarily indicate a problem. Excessive load may result in wear and damage in the future.

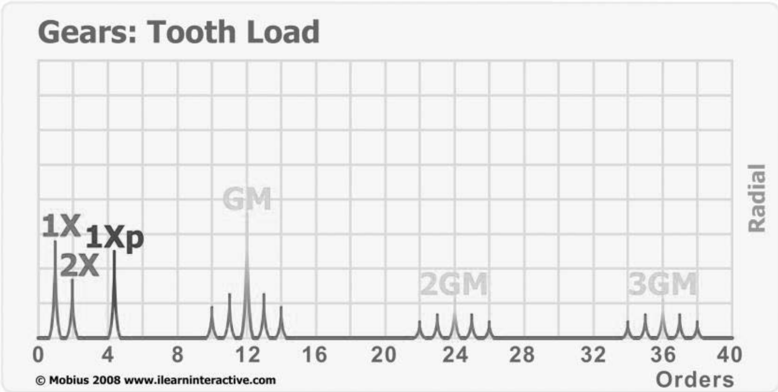


Figure 14-22 - Tooth load increases the gearmesh levels.

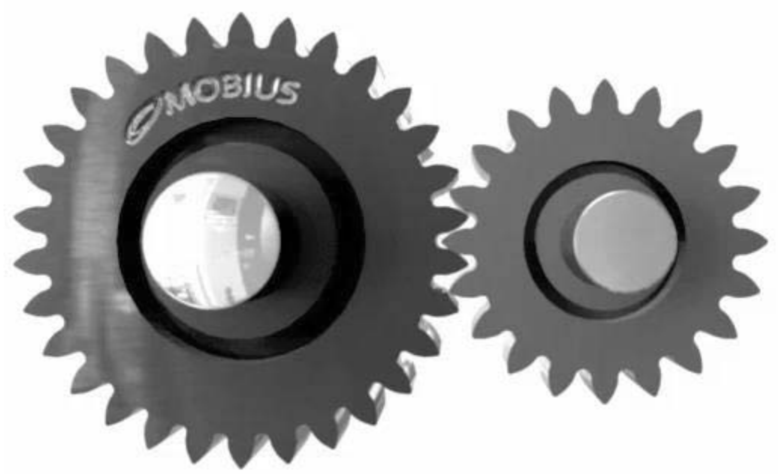


Figure 14-23 Gears meshing under load

Eccentric Gears

Eccentric gears and gears with bent shafts cause the load to vary at once per rev (amplitude modulation) producing sidebands spaced at shaft speed (of the offending gear) around gearmesh frequency. However, often you will only see a single sideband, rather than an entire family.

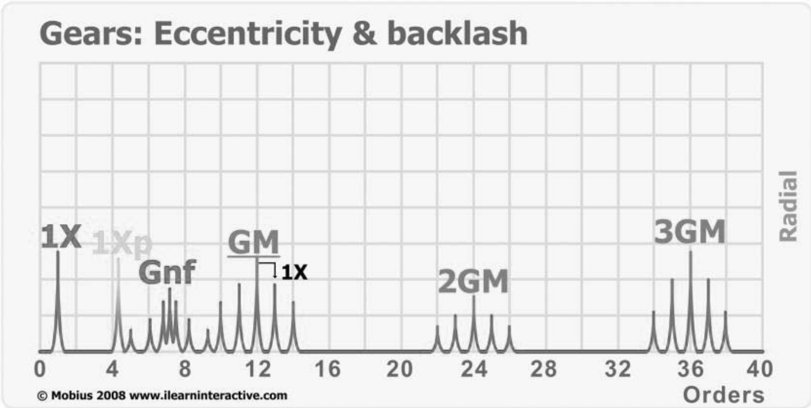


Figure 14-24 - Gear Eccentricity and Backlash

The 1x Gearmesh and the 3x Gearmesh frequencies will be dominant due to "non-linear" vibration. This is caused when the teeth slide in and out of contact.

Figure 14-25 Eccentric gears

Gear Backlash

Gear backlash also generates shaft speed sidebands around the gearmesh frequency. The gear mesh peak and the gear natural frequency peak will often decrease with increased load when this problem exists. See **Error! Reference source not found.**

Misaligned Gears

Misaligned gears also generate high gearmesh frequencies with sidebands, however it is common to have harmonics of gearmesh frequency, with higher levels at twice and three times gear mesh frequency.

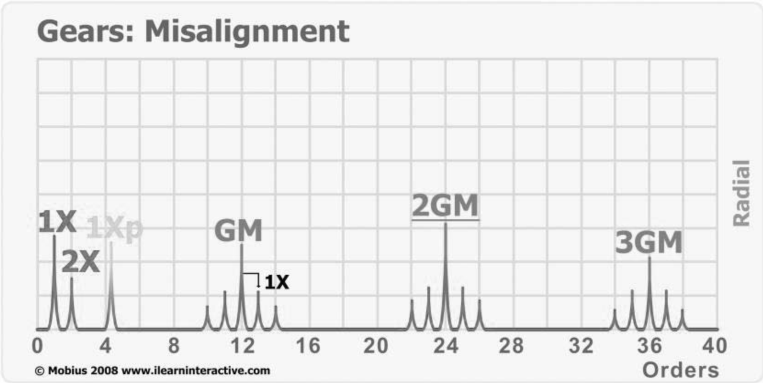


Figure 14-26 - Gear misalignment

Figure 14-27 Misaligned gears

Wear on the gear is uneven or skewed due to the angular contact between meshing teeth. This produces sidebands around 1x, 2x, and 3x gearmesh frequency may be at 2x RPM instead of 1x RPM. The sidebands may be lower below 2x gearmesh compared to the sidebands above 2x gearmesh.

The spectral data shows the sidebands when zoomed in around gearmesh frequency.

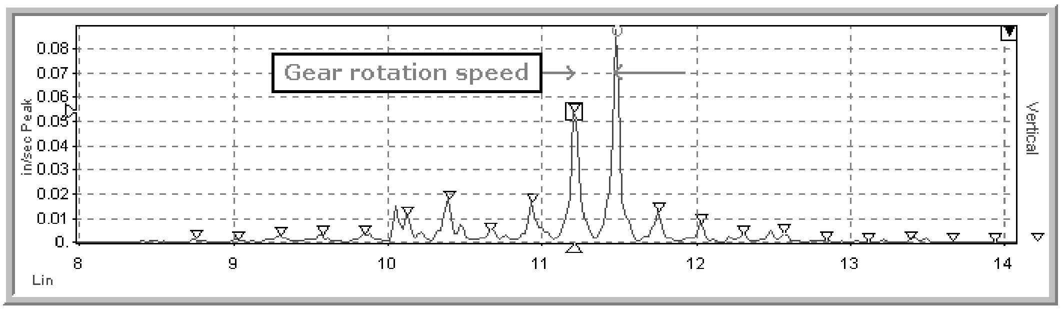


Figure 14-28 - Sidebands of shaft speed are evident when zoomed in around gearmesh frequency.

Harmonics of the gearmesh frequency indicate the presence of sidebands especially around 2x gearmesh frequency. 2x gearmesh frequency is also the highest in amplitude.

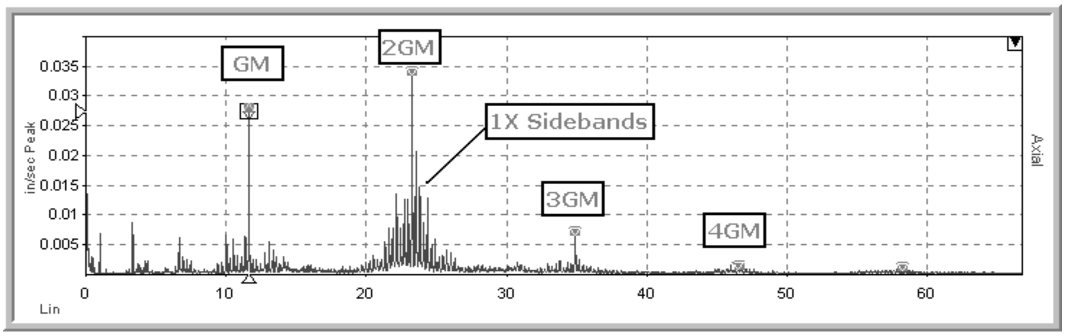


Figure 14-29 - 2x Gearmesh Frequency is highest and has many sidebands.

The sidebands would be clearer if the data were not integrated and would be even better in a logarithmic scale. This data actually came from the axial direction.

Time waveform analysis provides a very clear indication of the modulation that occurs.

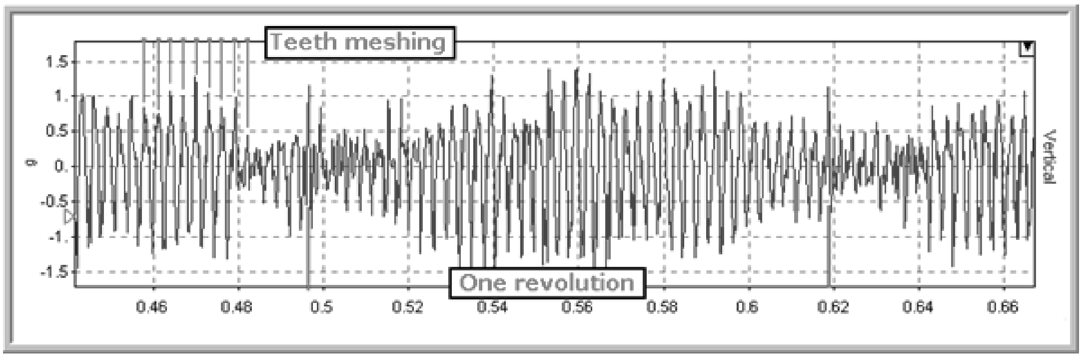


Figure 14-30 Time waveform from misaligned gears

Gears with spokes

Gears with spokes may not be perfectly round; there may be high spots and the point of each spoke. As that point goes into mesh, the vibration will rise momentarily. If there are six spokes, for example, there will be six pulsations per rotation.

Figure 14-31 Gear with spokes

Cracked or Broken Tooth

A cracked or broken tooth generates a high amplitude peak at the turning speed of that gear. And it will cause the gear natural frequency to be excited. There will be sidebands of turning speed of that gear around gearmesh frequency.

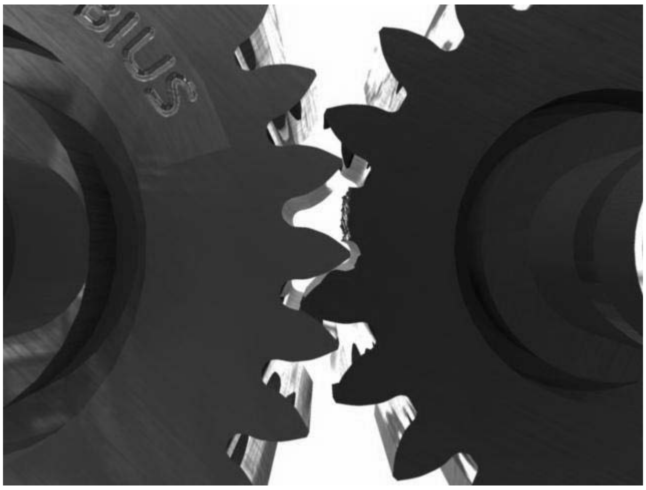


Figure 14-32 Gear with broken tooth

However, the best way to see a cracked or broken tooth is in the waveform. If there were 12 teeth, one of 12 pulses in the waveform will be very different from the others. The time difference between these pulses will be equal to the period of the turning speed of the gear because the tooth comes into contact once per revolution.

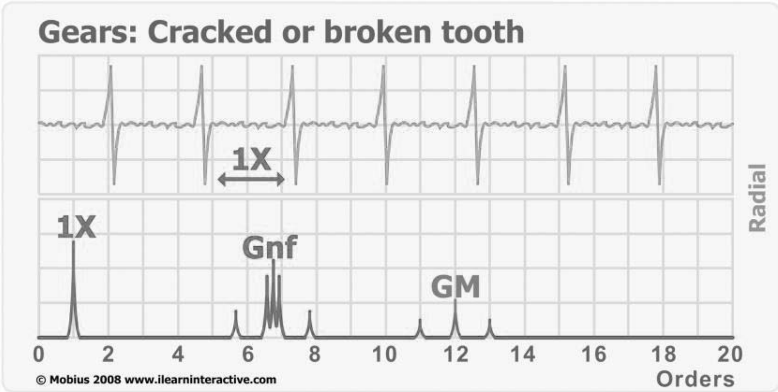


Figure 14-33 - Broken tooth shows up best in the time waveform as a pulse spaced at shaft speed.

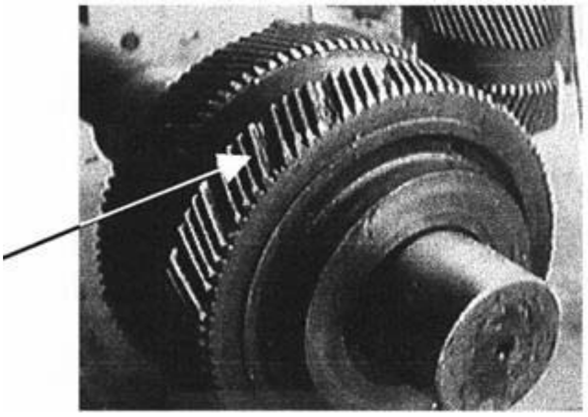


Figure 14-34 Gear with broken teeth

Time Waveforms and Gearbox analysis

Time waveform analysis is the key to success with gearboxes. Time waveforms show you what is happening with each tooth mesh. If the teeth do not mesh evenly, you will see it. If the teeth are worn; you will see it. If the teeth are damaged; you will certainly see it.

Here is an example of gears with spalling on the teeth. The spectrum does not make it very clear what the problem could be, but the waveform clearly reveals the problems. There is not a peak over 0.006 in/sec.

Figure 14-35 Damaged gear teeth

The spacing of the impacts in the waveform are at the RPM of the gear with spalling. In this case notice the low g levels that would not trigger an alarm.

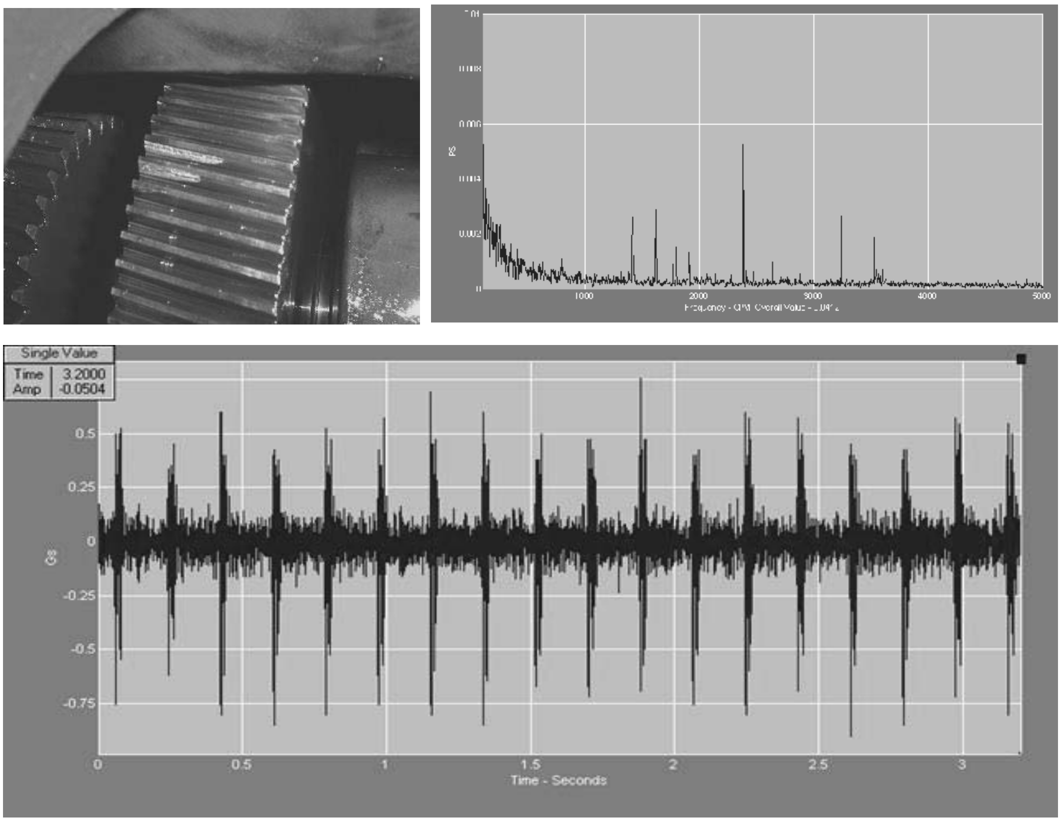


Figure 14-36 The impacts are spaced at the RPM of gear with spalling. No alarms were triggered.

Wear Particle Analysis

Wear Particle Analysis often provides a warning of gear damage well before vibration analysis does.

Note that wear particle analysis is typically separate from the routine oil analysis which misses abnormal wear particles. Cutting wear, abrasive wear, and sliding wear produce particles that are larger than the capabilities of spectroscopy.

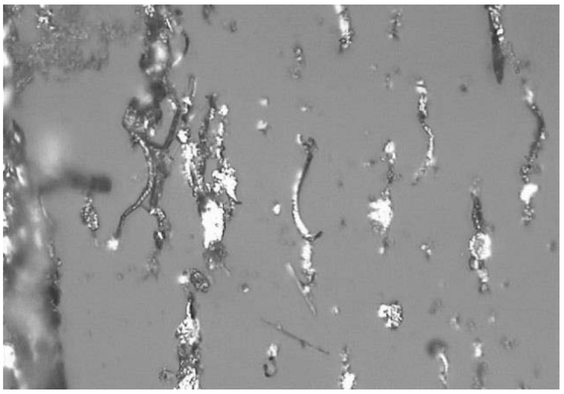


Figure 14-37 Wear particle analysis often reveals gear damage before vibration does.

Planetary (epicyclic) gearboxes

As discussed earlier, planetary gearboxes are used in a variety of applications, including wind turbines. Thanks to their complicated design, it can be very difficult to determine the forcing frequencies, and it can be difficult to take measurements that will pick up the vibration from the internal gears and bearings.

In many epicyclic gearing systems, one of these three basic components is held stationary; one of the two remaining components is an input, providing power to the system, while the last component is an output, receiving power from the system. The ratio of input rotation to output rotation is dependent upon the number of teeth in each gear, and upon which component is held stationary.

There are three configurations:

- Planetary: Ring stationary
- Star: Carrier stationary
- Solar: Sun stationary

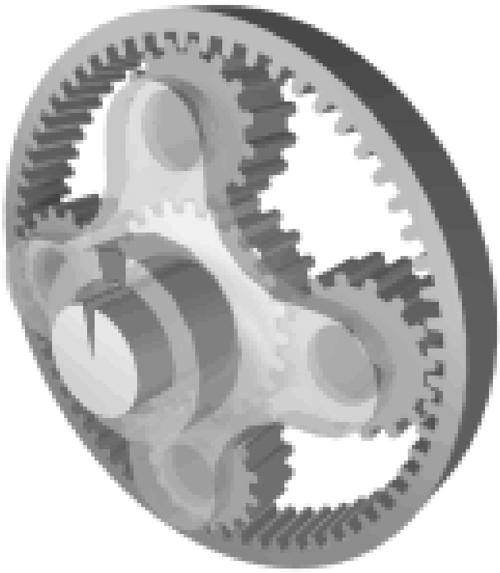


Figure 14-38 Planetary gear

The following equation must be true or else the gears cannot physically fit together.

$$N_{sun} + 2N_{planet} = N_{ring}$$

The equation that relates the speeds and gear teeth together is as follows:

$$\frac{N_{in}}{N_{out}} = \frac{S_{out} - S_{car}}{S_{in} - S_{car}}$$

Where 'S' is the **s**peed in the **in**put shaft, **out**put shaft or **car**rier, and 'N' is the number of teeth in the **in**put and **out**put gear.

Stationary carrier

One situation is when the planetary carrier is held stationary, and the sun gear is used as input. In this case, the planetary gears simply rotate about their own axes at a rate determined by the number of teeth in each gear. If the sun gear has *S* teeth, and each planet gear has *P* teeth, then the ratio is equal to -*S*/*P*. For instance, if the sun gear has 24 teeth, and each planet has 16 teeth, then the ratio is -24/16, or -3/2; this means that one clockwise turn of the sun gear produces 1.5 *counterclockwise* turns of the planet gears.

This rotation of the planet gears can in turn drive the annulus, in a corresponding ratio. If the annulus has A teeth, then the annulus will rotate by P/A turns for each turn of the planet gears. For instance, if the annulus has 64 teeth, and the planets 16, one clockwise turn of a planet gear results in 16/64, or 1/4 clockwise turns of the annulus.

Therefore:

One turn of the sun gear results in -*S*/*P* turns of the planets

One turn of a planet gear results in *P*/*A* turns of the annulus

So, with the planetary carrier locked, one turn of the sun gear results in -S/A turns of the annulus.

Rotating carrier

The annulus may also be held fixed, with input provided to the planetary gear carrier; output rotation is then produced from the sun gear. This configuration will produce an increase in gear ratio, equal to 1+*A*/*S*

If the annulus is held stationary and the sun gear is used as the input, the planet carrier will be the output. The gear ratio in this case will be 1/(1+*A*/*S*). This is the lowest gear ratio attainable with an epicyclic gear train.

The bottom line is that you have to consider each mesh as if it is a normal mesh, however when the planets are involved you have to take into account their movement. If the carrier is rotating, either as the input or output, then the frequency calculation is adjusted for that movement. You can take the formula above and immediately cancel one portion because either the carrier, annulus/ring or sun will be stationary. You will need to know the input speed or output speed, and then you solve for the speed of the output or input. And once you know each speed you can compute the gearmesh frequencies.

The general formula is:

$$(2+\frac{N_s}{N_p})S_r + \frac{N_s}{N_p}S_s + 2(\frac{N_s}{N_p}+1)S_c = 0$$

As with any gearbox, if you know the number of teeth and the rotating speed you can compute the gearmesh frequency. Planetary gearboxes must be designed such that the number of teeth on the ring will be equal to the number of teeth on the sun plus twice the number of teeth on the planet. Therefore the gearmesh frequency calculated when you consider the mesh between the plant and sun must be the same as the gearmesh between the planet and the ring. We can then compute the sideband spacing frequency based on the rotating speed (where applicable) of the planet, sun and ring/annulus.

When we look at a spectrum from a planetary gearbox, we will look for the traditional gearmesh frequency and sidebands, and we will look at twice the gearmesh frequency and three times the gearmesh frequency. We will also look at a frequency which is equal to the number of planets times gearmesh. So if there were four planets, we will multiply the gearmesh by 4.

The general formula for the gearmesh frequencies is as follows:

$$GM = N_s(S_c - S_s) = N_p(S_p - S_c) = N_r(S_r - S_c)$$

Where:

'S' denotes Speed
'N' denotes Number of teeth
's' denotes sun
'p' denotes planets
'r' *denotes ring*

Different configurations

We have focused on a common gearbox configuration. There are other configurations, and trains of planetary gearboxes are possible. Therefore it is important to learn about your configuration and ensure that it follows the same rules.

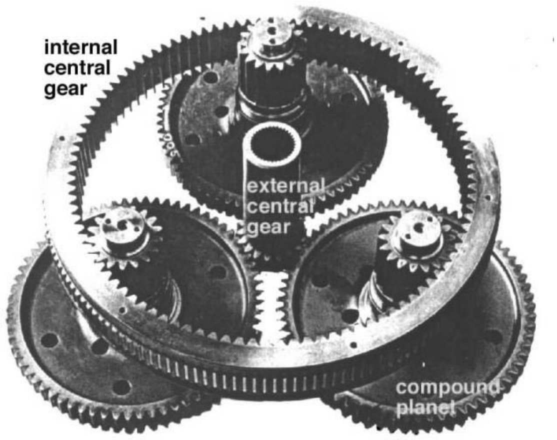


Figure 14-39 Planetary gearbox with different configuration

Monitoring gearboxes

A discussion of gearbox analysis is incomplete without a discussion of correct measurement techniques. We have to be sure to locate the sensor in an appropriate location; we must mount it so that the high frequencies can be measured; and we must set the Fmax correctly to cover the frequencies of interest. It is also necessary to collect time waveforms in addition to spectra, and to utilize demodulation or enveloping techniques. And finally, because of the number of gears and bearings, it may be necessary to perform time synchronous averaging.

It is essential that the transducer location is selected carefully. You are attempting to measure the vibration that results from mating gears inside the gear case. The vibration must travel along the gear shaft, through the bearings, and through any gaskets, etc. A complex gearbox may have a number of gears and you may need to take measurements from a number of locations.

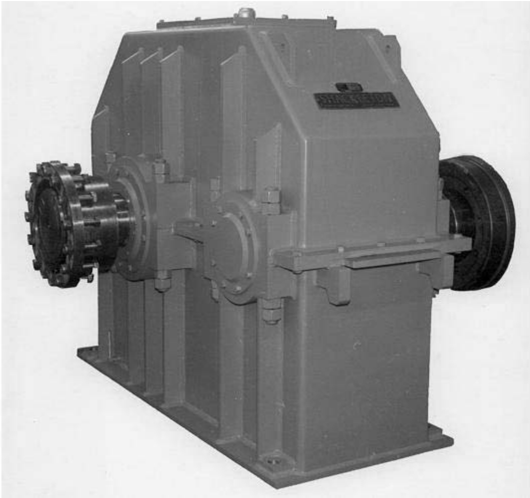


Figure 14-40 Gearbox case

If you are aware of where internal structural members (like webbing) are located, then you may place the sensor where the member meets the case. There should be a good mechanical transmission path.

Covers and walls may resonate, and are therefore ***not*** good locations for the sensor.

Gearboxes are typically important enough that you should take measurements at every possible location and axis. However, in addition to trying to find the location closest to bearings and solid structural members, you should also consider the direction of the primary forces.

A spur gear will tend to generate radial forces; so the sensor would be placed in an axis that is at right angles to the axis of the shaft.

Figure 14-41 Spur and helical gears

However a helical gear (or herringbone or any other gear with a bevel angle) will generate axial thrust; so the sensor should be mounted so as to measure axially.

You must also consider how you are mounting the accelerometer on the case. If you are hoping to detect high frequencies, then a handheld mount will not be acceptable, and a two-pole magnet may not be acceptable. Take great care with your measurements and you will get the best results.

Figure 14-42 Accelerometer mounting options

The measurement we take must provide coverage of all the frequencies of interest. Whether the gearbox generates high frequencies or low frequencies (reduction gearbox), it is essential that the Fmax is set to cover three times the highest gearmesh frequency.

Because we are looking for sidebands of the driver and driver rpm, the resolution chosen must allow you to resolve between these peaks.

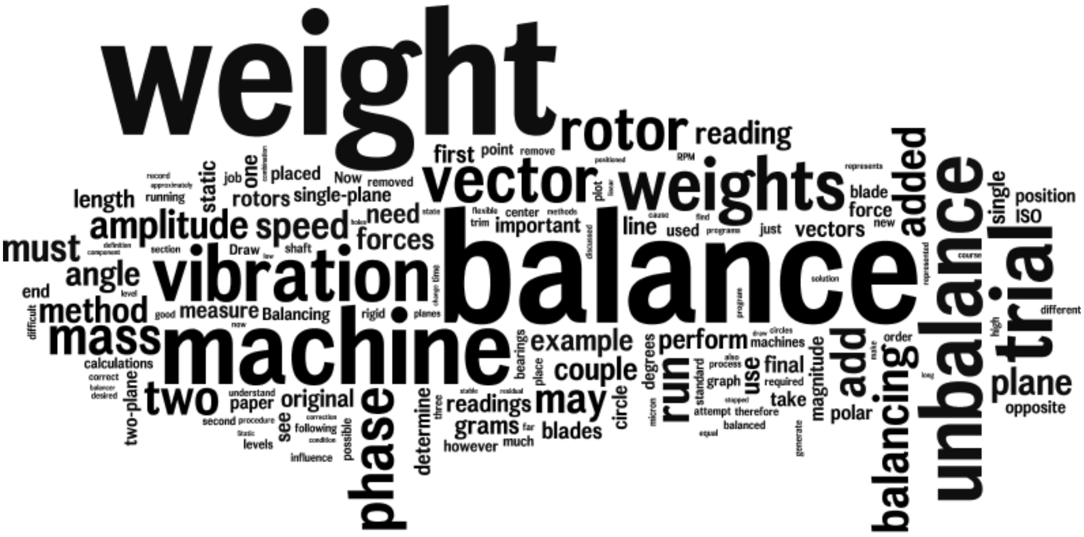
We can determine the resolution by dividing the frequency range (Fmax) by the number of lines of resolution – and then multiplying by 1.5 because we are using a Hanning window. For example if we have an Fmax of 2200 Hz, and 800 lines, each line will be separated by (2200/800) = 2.75 Hz, but each unique peak will require 4.125 Hz. In order to see a gap between peak (i.e. to resolve them), we need peaks to be approximately 50% further apart – so peaks can be a minimum of 6 Hz apart.

It is worthwhile considering the resolution requirements, or else the sideband peaks will merge and blur into one.

Time waveform analysis is an essential diagnostic tool. It is also helpful if you have a way to listen to the collected waveforms.

The time waveform shows the instantaneous vibration that results from each mesh. If there is variation in the way each tooth meshes together, due to eccentricity, misalignment, damaged teeth, or for any other reason, you will see these variations in the time waveform data. Events that do not occur in every cycle will be “washed out” of the time waveform. Spectrum averaging will remove any patterns that made it into the spectrum. The sidebands analysis is useful, however there is information in the time waveform that is lost in the FFT transform.

Therefore the time waveform must be off sufficiently high resolution to be able to see this data, and of sufficient duration to catch all events.



Chapter 15

Balancing Rotating Machinery

Objectives:

- Understand the important of balancing
- Learn about single-plane, two-plane and four-run-no-phase balancing
- Understand the difference between rigid rotor and flexible rotor balancing
- Understand the balance grade

Balancing rotating machinery

Balancing rotating machinery is an important part of any reliability program. Balanced machinery run more smoothly and therefore run longer without incident. Modern vibration analyzers and purpose-built balancing instruments have made it easier to balance machines in-situ – that is, without having to remove the rotors and ship them off to a balancing machine.

Like any instrument of this type, it is always important to understand how the instrument works and how to correctly prepare the job and deal with any unusual readings that you may observe. In an ideal world you can take the vibration and phase readings, follow each of the steps, and the machine will be precision balanced. But life is rarely that simple…

The goals of this chapter

This section is intended to provide an introduction to balancing. The aim is not to teach you *everything* you need to know so that you can balance a machine – there is simply not enough time on a detailed course like this. Instead you should:

- Understand in detail how single plane balancing works.
- Understand the difference between one plane and two-plane balancing.
- Understand the four-run-no-phase technique.
- Understand the difference between balancing rigid rotors and flexible rotors.
- Become familiar with the relevant ISO standards.

With this base knowledge you should find it far easier to learn the detailed steps required to perform the balance, and you will have a greater appreciation of what is involved if your request that a machine be balanced.

What is balancing?

First, what is balancing? The ISO standard 1940-1973 (E) definition is:

"*Balancing is the process of attempting to improve the mass distribution of a body so that it rotates in its bearings without unbalance centrifuge forces.*"

Those "unbalance centrifuge forces" are destructive. The forces slowly destroy bearings, and can generate vibration that can damage other equipment and processes. All machines should be precision balanced and aligned to reduce these destructive forces – and therefore increase the life of the machine and improve reliability.

The ISO standard 1940-1:2003 definition is:

"*Balancing: procedure by which the mass distribution of a rotor is checked and, if necessary, adjusted to ensure that the residual unbalance or the vibration in the journals and/or forces on the* *bearings at a frequency corresponding to service speed are within specified limits."* [Origin: ISO 1925:2001, definition 4.1]

The goal and the intent are the same – reducing vibration levels to within specified limits. Later in this section we will discuss the ISO standards for balancing and the recommended vibration levels for different applications.

There are two more definitions from ISO 1940-1:2003 that are important:

"Unbalance: condition which exists in a rotor when vibration force or motion is imparted to its bearings as a result or centrifugal forces." [Origin: ISO 1925:2001, definition 3.1]

"Residual unbalance: unbalance of any kind that remains after balancing." [Origin: ISO 1925:2001, definition 3.0]

The ISO standard 1925:2001 is "Mechanical vibration – Balancing – Vocabulary".

Preparing for the balance job - a word of warning

Given the ease of use of balancing software provided with modern analyzers it is conceivable that balancing technicians may jump straight into a balance job without thinking about the machine dynamics – the interplay between the unbalance forces, the rotor and the bearings. Balancing is a difficult, time consuming, and potentially dangerous task. It is essential that you take all precautions before and during a balance job. Follow all lock-out tag-out procedures, and do not make any assumptions about the safety systems installed at the site.

Safety first!

There are four main sources of danger that you should be aware of. They are:

- The machine starting unexpectedly
- Balance weights flying off the rotor
- The general dangers inherent in working in a hazardous environment
- Becoming entangled – you or your equipment – in the machine

Figure 15-1 Follow all lock-out and tag-out procedures

ISO 20806:2009 "Mechanical vibration – criteria and safeguards for the in-situ balancing of medium and large rotors" is a useful standard to be acquainted with. It also includes information on safety

Another 'warning' is to not to underestimate balancing. If everything goes smoothly you may complete a balance job in a couple of hours. However there is a lot that can go wrong. You need to think about what you are doing and not make any assumptions. Balancing programs can make the measurements and calculations easier, however resonances, foundation problems, thermal growth and more can make the balance job quite challenging.

Is the machine out of balance?

Many people fail at the first hurdle. They measure high vibration and assume the machine is out-of-balance. So they attempt to balance the machine. Yet, to their surprise, the balance weights recommended do not reduce the vibration levels as much as expected.

So you should verify that the machine is out of balance before you spend time trying to attempt to balance it! It may sound obvious, but a great deal of time has been lost by people trying to balance a machine that is misaligned, or that has some other fault that presents similar vibration symptoms to unbalance.

It could be that the machine has more than one fault condition. Correct those other fault conditions before you attempt to balance the machine.

Some conditions can make it very difficult to balance a machine. For example, a natural frequency close to the running speed will cause the phase readings to be unsteady. If the speed, phase or amplitude are not steady, then it will be difficult to successfully balance the machine.

Can the machine be balanced?

Before you consider performing an in-situ balance job, you must first consider the following checklist. In order to balance the machine you must check that:

- You can start and stop the machine.
    - The machine will need to be stopped and started in order to add trial weights and the final balance weights.
- It is possible to add balance weights (you will need access to the rotor, fan blades, etc.).
    - Many machines have balance rings, or places on the rotor where balance weights can be added. You must check how the weights will be added (or removed) and that you have suitable balance weights.
- It is possible to gain access to the machine.
    - You will need to be able to access the balance ring or rotor (wherever the balance weights will be added).
- It is possible to control the speed of the machine.
    - It is important that the speed of the machine remains constant during the tests and from one test to the next (i.e. between the original run and each of the trial, final and trim runs).
- The speed, amplitude and phase must be steady during the tests.
    - The speed of the machine, and the amplitude and phase readings, must be constant during the tests. This may mean that you must run the machine for many minutes (sometimes longer than one hour) in order for the machine to be running in a stable, repeatable condition. Thermal transients and load changes must be considered before attempting to balance the machine.
- It is possible to take a phase reading (you will need access to the shaft).
    - You will need to acquire phase readings during the test. This step will be discussed in greater detail, however you must ensure that you can either add reflective tape to the shaft and mount a laser or photo-tachometer, or mount a non-contact probe to detect a keyway or other such physical irregularity.
- You will need the required instrumentation and a balance program/method.
    - Of course, you will also need the instrumentation to measure the amplitude and phase, a calculator or balance program (or polar plotting paper and protractor), and some scales to weigh the balance masses.

Vectors and polar plots

If you wish to perform a single-plane balance with a polar plot, then you must understand vectors. If you wish to understand the balancing process and the effect of adding trial weights, then it helps a great deal if you understand vectors.

Understanding vectors, and knowing how to add and subtract vectors may involve learning some new concepts, but I am sure that this section will make it very clear to you.

Vectors are a combination of an angle and a "scalar" amount. The "scalar" amount could be wind speed, current flow in an ocean, or in the vibration world, a vibration amplitude at the running speed of the machine. For example, if a ship was sailing at 10 knots in a northerly direction, then we can represent that as a vector: 10 knts @ 0°. Readings such as 2 micron @ 45° and 0.5 in/sec @ 125° are vibration vectors.

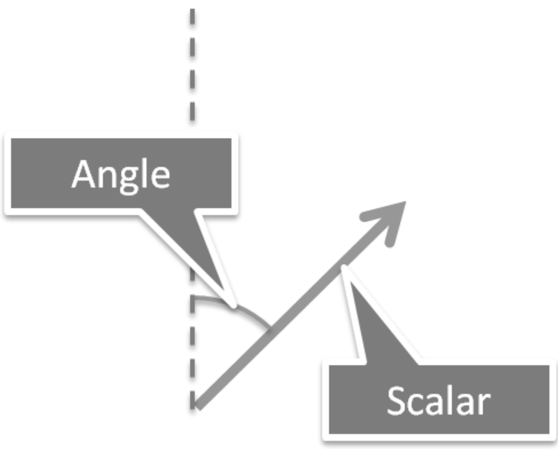


Figure 15-2

Vectors are normally represented on a circular plot called a "polar plot" (Figure 15-3).

The scalar value (amplitude) is represented by the length of the line (arrow), and the angle is drawn with 0° at the top.

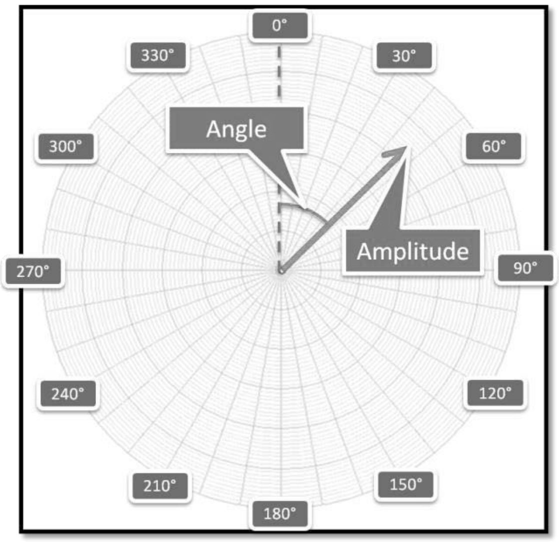


Figure 15-3

The markings on the polar plot make it possible to read angles and lengths, although it can be helpful to also have a ruler and a protractor; especially when it comes to vector addition and subtraction.

The radiating circles represent the amplitude.

Figure 15-4 shows a simple example. We made the radius of the plot equal to 6 mils which means every ring represents 1 mil (units of displacement). The ruler will help us to measure the length of the vectors.

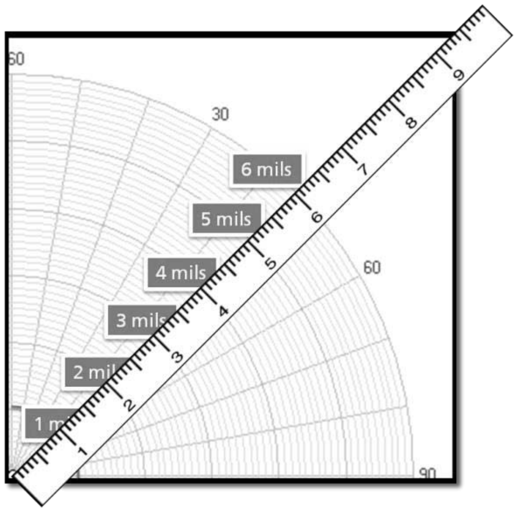


Figure 15-4

If we measured a vibration of 5 mils at 45° then we would draw the vector as shown in Figure 15-5.

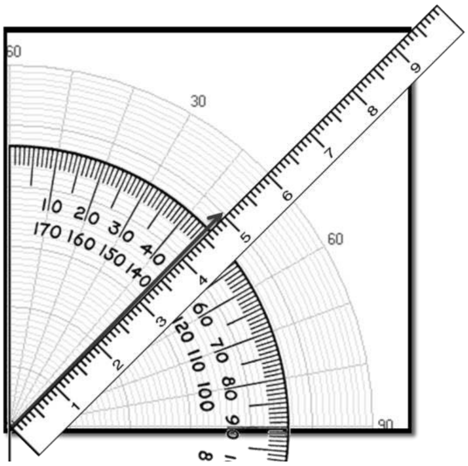


Figure 15-5

To make sure we really understand vectors, let's go through a simple example.

In Figure 15-6 we have a ship traveling at its maximum speed of 10 knots towards the east. The ship is moving due to its engines. There is no wind and no current.

Therefore the ship travels east at 10 knots. As a vector we would describe this as: **10 knts @ 90°**

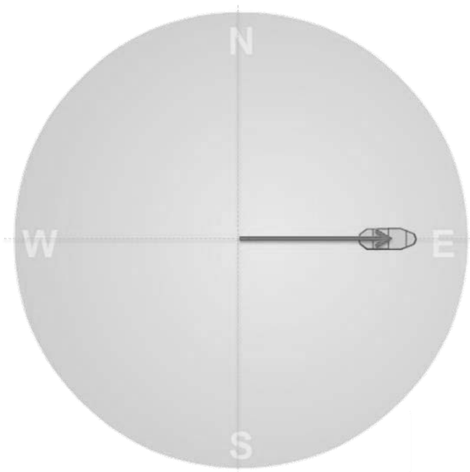


Figure 15-6

Unfortunately for the ship, a wind blows up from the south which has the effect of pushing the ship to the north. The wind is strong enough such that if it turned its engines off the ship would travel at 10 knots to the north.

But if the ship turns on its engines and tries to sail to the east, it would find that it does not travel to the east; instead it travels to the north-east.

We can therefore represent the actual movement of the ship with a different vector, one that points at 45° (Figure 15-7).

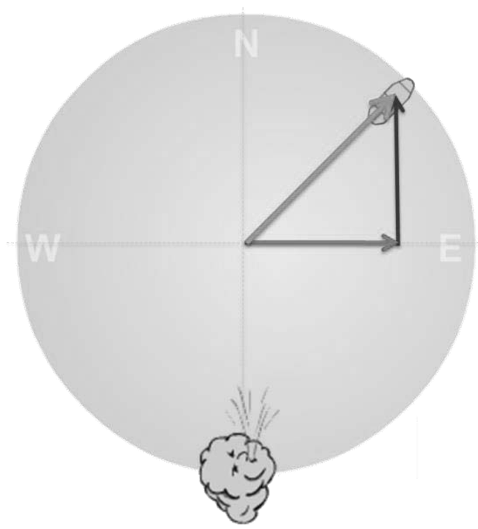


Figure 15-7

If we were to measure the length of that vector we would find that it was 14.14 knots in length. So the good news for the ship is that it is traveling faster than before. The bad news is that it is going in the wrong direction…

We can determine the angle and length mathematically or we can plot the vectors on a polar plot and either use the scale provided or use a ruler and protractor.

Adding vectors

What we have actually just done is to add two vectors. Let's look at that a little more closely.

Adding vectors is a case of moving the tail of one vector to the tip of the other vector.

It does not matter which vector is added first. No matter how you add them the result will be the same (Figure 15-8 center and right).

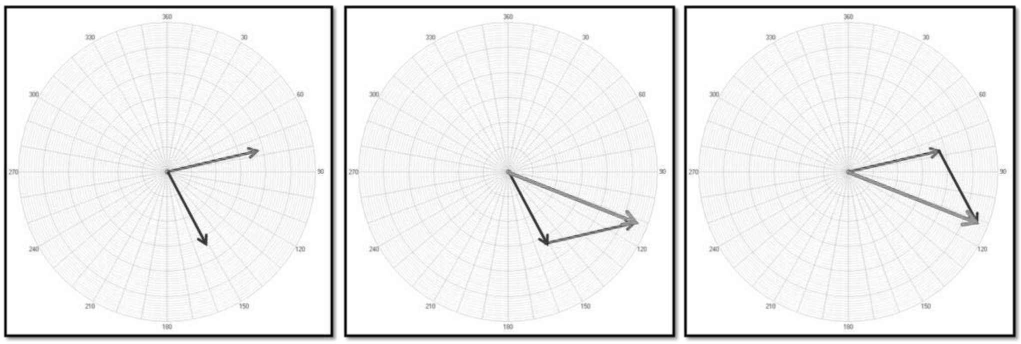


Figure 15-8

And it does not matter how many vectors must be added; the process is the same (Figure 15-9).

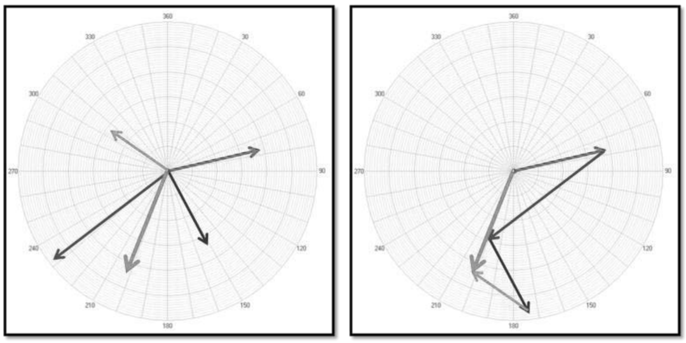


Figure 15-9

Vectors can "easily" be added mathematically. It is a case of breaking each vector into its "X" and "Y" components, and then adding the "X" components and adding the "Y" components.

The length of the vector "Z" can be calculated using Pythagoras' theorem. And the angle can be calculated using ATAN ($\tan^{-1}$), as shown in Figure 15-10).

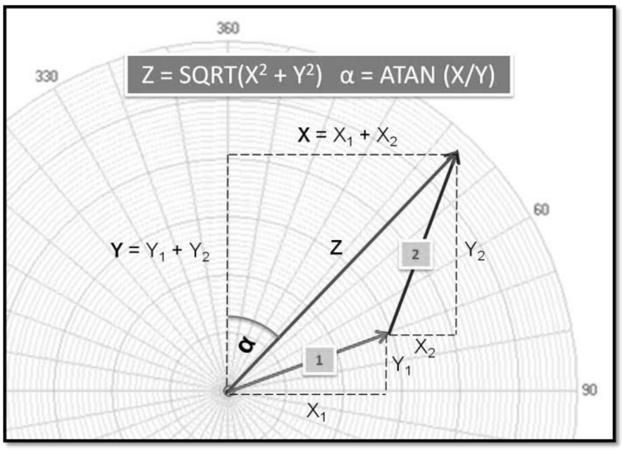
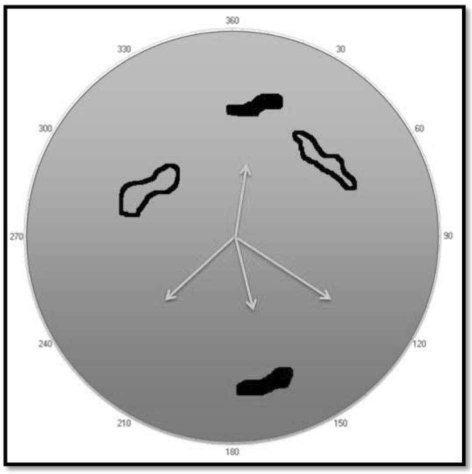


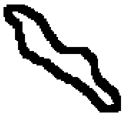
Figure 15-10

It is worth making a very important point right now. If you consider an out-of-balance rotor, the source of imbalance is probably not just one source at precisely one location (like a bolt that has been attached to the rotor in the wrong location).

More than likely there are a number of sources of erosion and material porosity, dirt/grim build-up, and so on.

Each source of mass (or loss of mass) can be represented by a vector (Figure 15-11).





Porosity or erosion: this is a *loss* of mass so the vector points *away* from these points.

Build-up of dirt/grime that cannot be cleaned off. They add mass.

Figure 15-11

And just as before, we can add the vectors together. The red vector is the final vector that we would measure as the unbalance (Figure 15-12).

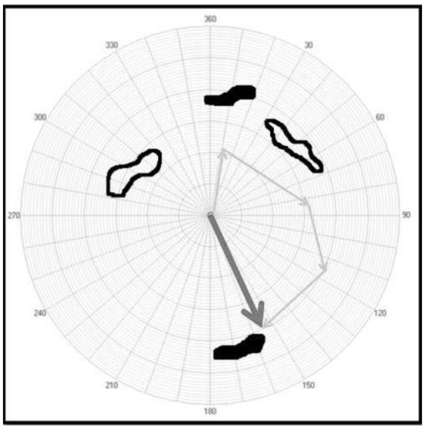


Figure 15-12

Please note that you would not need to add these vectors; this just happens naturally. The point is that each individual source of unbalance contributes to the final unbalance that causes the machine to vibrate.

There is one interesting situation to consider. What would happen if we added the two vectors in Figure 15-13?

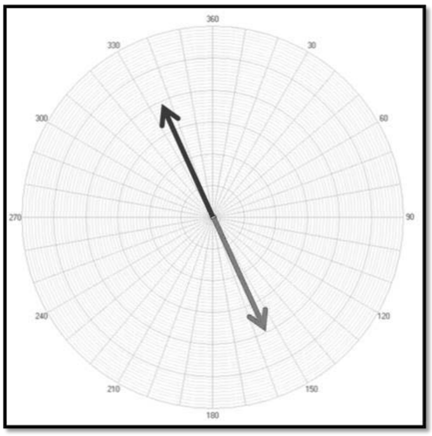


Figure 15-13

Yes, that is correct. Because they have the same length, and their angles are opposite, the result is a vector with zero length!

Why is this relevant? Because that is exactly what we are trying to do when we balance the rotor. If the red vector represents the sum of the sources of unbalance (porosity, erosion, etc.), then we will add a weight that would create the blue vector – it will have an equal and opposite effect.

Subtracting vectors

Subtracting vectors often causes a great deal of confusion. It is a necessary step when performing single-plane balancing with vectors. The confusion is unnecessary.

To subtract vector “A” from vector “B”, you simply turn vector “A” around 180° and then add it!

If we have vectors “A” and “B” and we want “A-B” then we flip “B” around by 180° and then add its tail to the tip of vector “A” (Figure 15-14).

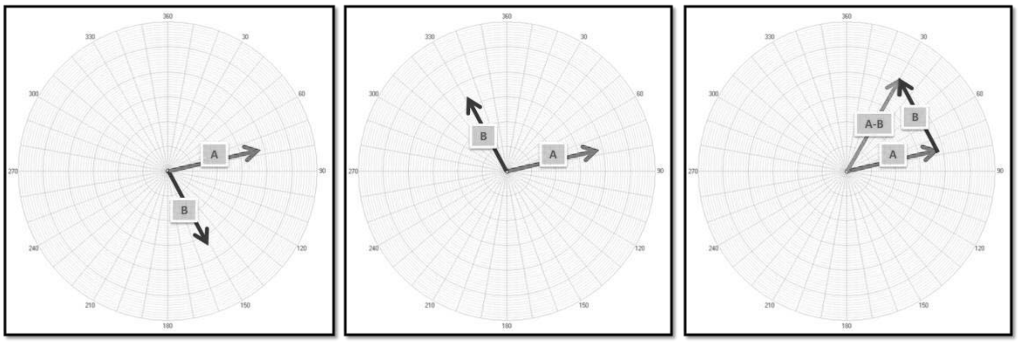


Figure 15-14

Note that subtraction is not like addition: “A-B” is different to “B-A” (sorry if that sounds obvious).

The result is the same vector as last time, rotated 180° (Figure 15-15).

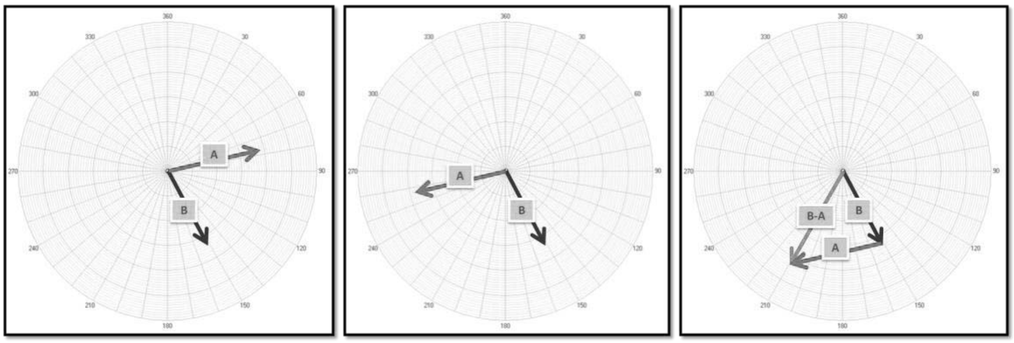


Figure 15-15

Single-plane balancing

In order to perform a single plane balance we need to determine where to place the balance correction weight, and how much it must weigh. Note that more than one weight may need to be added if it is not possible to add the weight at the desired angle (because there is not a blade or balance hole at that angle.

Summary of the single plane method

First we take a vibration measurement on a bearing and record the magnitude and phase at the running speed of the machine (therefore we need a once-per-revolution tachometer reference). We document this information as the "As-found" condition of the machine prior to balancing it.

Next we add a "trial weight" to the rotor, and measure at the same point and record the magnitude and phase at the running speed. We do this to see how the rotor was "influenced" by the addition of the weight. The magnitude and phase should both change by 15% or 15 degrees or more.

Now we know how the magnitude and phase of the vibration at the monitoring point has changed based on the addition of the trial weights. We can then perform a calculation (or sketch it out with vectors on polar graph paper) to determine where we should add weights (and how much they should weigh) to cancel the effect of the out of balance force.

Remember, we do not know why the machine/rotor is out-of-balance – we can't see a weight on the rotor. It is out-of-balance because of poor design/manufacturing/assembly, wear/erosion on blades/vanes, build-up on blades/vanes, incorrect key length, and distortion/bends.

We then add the correction weight and take another measurement to see if the unbalance is within tolerance. Often the unbalance will not be within tolerance on the first attempt, so we use the new readings to determine where a "trim" weight should be added in order to bring the machine into tolerance.

Once the machine has been balanced satisfactorily, the results should be documented.

The procedure just described is used regardless of whether you use a data collector, balance computer, or graph paper. We take initial readings, add trial weights to see how the balance state was "influenced", and then we determine where the final weights should be placed. (Please note that there are other ways to balance a machine, but they are not covered in this course.)

Using vectors

Although you will probably use a balance calculator or your data collector to perform the balancing task, it is important to understand the single-plane vector method. If you understand this procedure, you will be able to cope when your balance job does not go smoothly, and you will find it much easier to understand the two-plane and "influence coefficients" method used by most balance calculators and data collectors.

We will use polar graph paper to record the vibration amplitude (magnitude) and phase. The radial rings represent the magnitude (in this example each ring is 1 mil), and the radial lines are angle (degrees). 360/0 degrees is at the top.

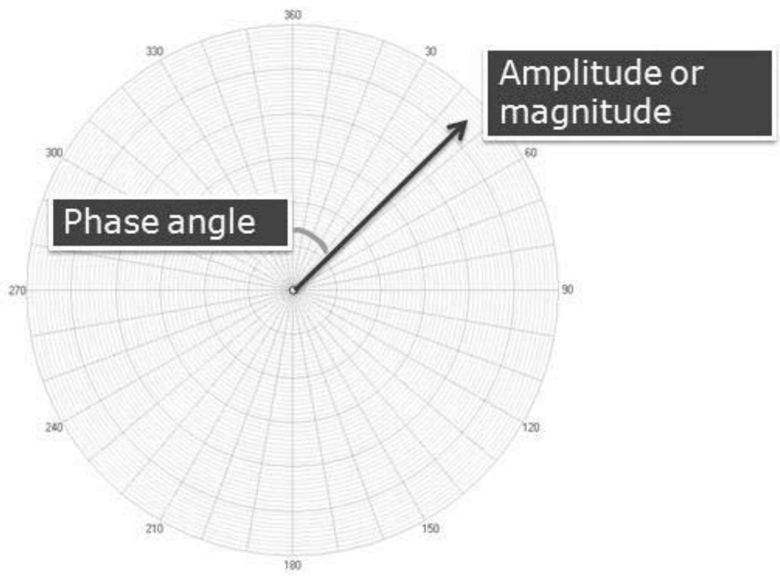


Figure 15-16

When presented this way, the magnitude/phase pairs are called vectors. We draw a line the length of the magnitude at an angle of the phase reading. We can use the format regardless of whether we measure the vibration in units of displacement, velocity or acceleration. The only rule is that we must use the same units (and the same sensors mounted in the same locations) for the entire test.

Measurement setup

The first step is to set up your measuring equipment (data collector) so that you can take a magnitude (i.e. vibration amplitude reading) and phase reading at the running (turning) speed of the machine. The sensor will be positioned so as to measure the highest amplitude, which is normally in the horizontal direction.

Figure 15-17

Original balance run

With the machine is running at normal speed, the magnitude, phase and RPM are recorded. This is called the "original reading". The magnitude and phase are plotted on the polar graph paper. In this example, the magnitude was 5 mils, and the phase reading was 45 degrees. It is labeled "O" for original unbalance.

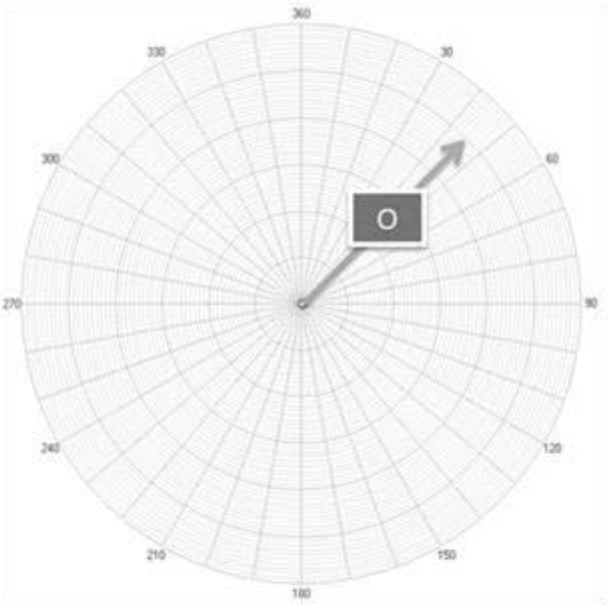


Figure 15-18

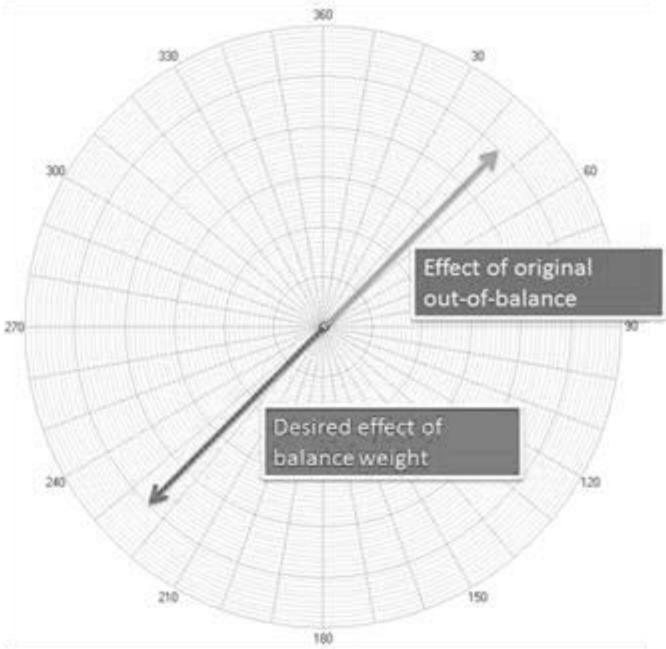


Figure 15-19

Add the trial weight

Then the machine must be stopped and we add a "trial weight". In a moment we will discuss how to calculate the mass of the trial weight, but we are trying to influence the state of balance of the rotor.

For example, if we add 5 grams of weight and we see that the vibration amplitude reduces by 30% and phase angle changes by 30°, we can then calculate how much weight is required, and the location of the weight, so that the unbalance force is cancelled altogether.

If we add a trial weight that is too small, the vibration amplitude and phase will not change very much, and any calculation we perform will be very inaccurate. On the other hand, if the weight is too large, and we happen to place it where the unbalance is greatest, we could create a damaging and potentially unsafe situation.

Balance weights come in a large variety of shapes and sizes. Some are designed so that they are easily added for the “trial run”, but they must then be affixed permanently at the conclusion of the balance job.

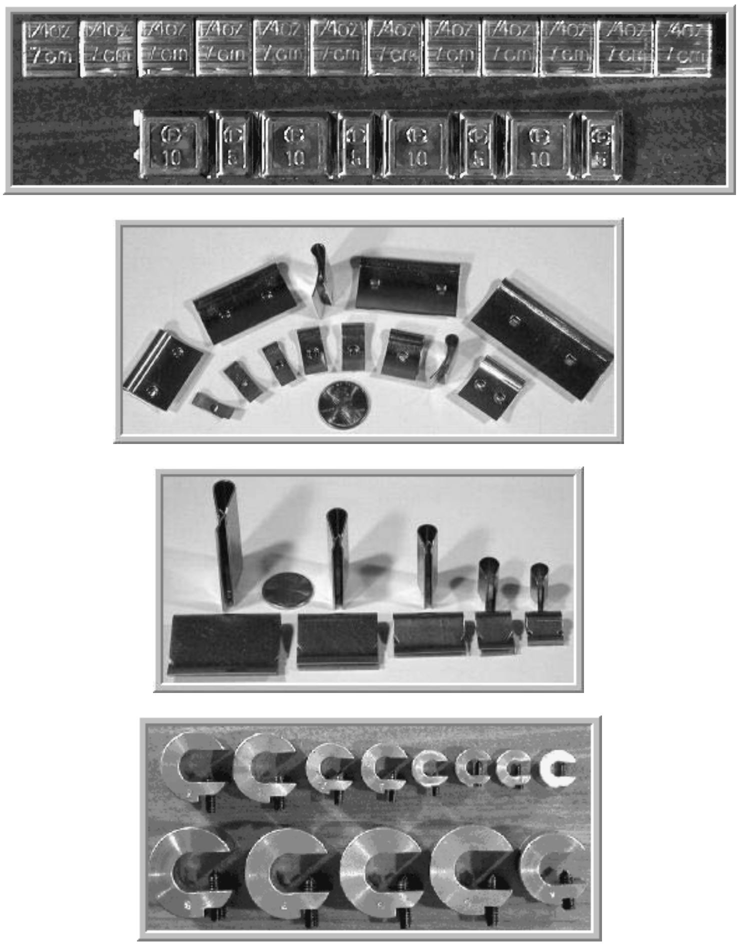


Figure 15-20

It is also very important to have a good set of scales so that you can weigh the trial and final balance weights.

Figure 15-21

Selection of the trial weight is very important. As a general guideline (Wowk), add a weight that produces a force of 10% of the rotor weight. This criterion was originally developed for flexible rotors where the forces may be amplified as the rotor passes through the first critical speed. For rigid rotors, you could safely use 2-3 times this mass.

To calculate the forces produced by the trial weight, you can use these formulas

Imperial:

$$F = 1.77\ x\ W_t\ x\ R\ x\ \left(\frac{RPM}{1000}\right)^2$$

W_t=Trial weight (oz)

R=Radius of trial weight (inches)

Metric:

$$F = 0.01\ x\ W_t\ x\ R\ x\ \left(\frac{RPM}{1000}\right)^2$$

W_t=Trial weight (grams)

R=Radius of trial weight (cm)

As noted, this force should be equal to 10% of the rotor mass, therefore, the "F" in the above equations should be equal to the rotor mass (in pounds or kg respectively) x 0.1.

The following equation can also be used to calculate the mass of the trial weight:

$$W_t = 56{,}375\,\frac{W_R}{N^2 r}$$

W_T=trial weight (oz)

W_R=Static weight of the rotor (lb)

N=Speed of the rotor (RPM)

r=Radius of trial weight (inches)

For shafts rotating in the range 1200 to 3600 RPM, this equation can be simplified to:

$$W_t = 0.004\frac{W_R}{r}$$

In metric this simplified to:

$$W_t = 30\frac{W_R}{r}$$

W_T=trial weight (oz)

W_R=Static weight of the rotor (lb)

r=Radius of trial weight (inches)

For example, if the machine had a 90 kg rotor, and the balance weights were positioned at a radius of 350 mm, the trial weight would be 30 x 90 / 350 = 7.7 grams

Selecting the position for the trial weight

Many people simply select an arbitrary position for the trial weight; often in line with the tachometer reference. However we can be smarter than that. We can position the weight opposite where we believe the heavy spot is located. If you are using a proximity probe and the proximity probe is positioned in-line with the phase reference (optical-tach, laser-tach or keyphasor), then we can simply place the trial weight opposite the unbalance vector. That is, if we measure a phase angle of 45° we should place the weight at 225° (45+180). If the rotor is rigid, and there are no significant phase lags due to the mechanical system or electronics, then we are bound to reduce the vibration on the first shot. Not only will this help our calculations, but it will ensure a safe trial run – if we happen to place the trial weight in-line with the out-of-balance weight (for example at 45°), then we would cause the vibration levels to rise considerably, and possibly to dangerous levels.

Trial run

With the trial weight carefully installed on the machine, we run it up to the same speed, wait for the reading to settle, and then again record the magnitude and phase at the running speed.

If the magnitude has not changed by 30% or more, and if the angle has not changed by 30 degrees or more, then we will have to stop the machine and replace the trial weight with a larger weight.

Let's assume the new reading was 4 mils and 130 degrees. We add this to the polar plot. This vector represents the vibration level due to the original unbalance and the trial weight, so it is labeled "O+T".

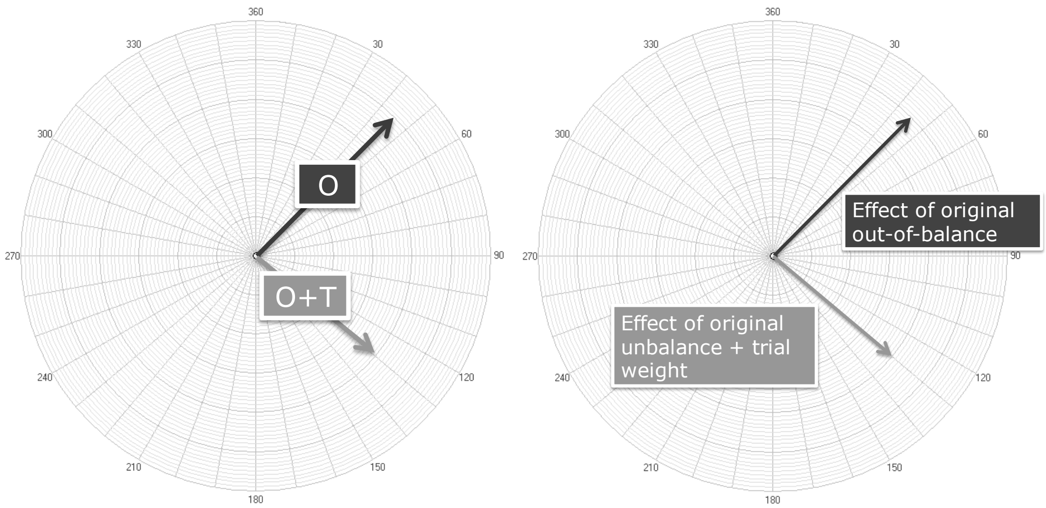


Figure 15-22

Now we draw a line from the end of the original "O" vector to the end of the "O+T" vector. This represents a **subtraction of the two vectors**, leaving just the influence of the trial weight alone, so we mark this vector "T".

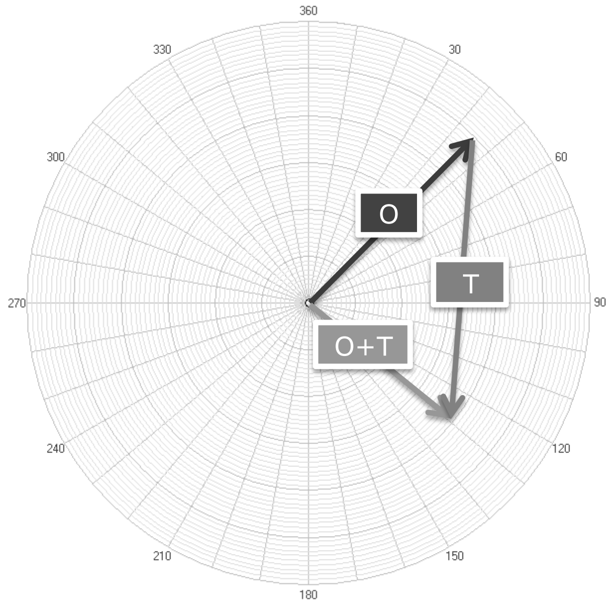


Figure 15-23

The new vector represents the vibration due to the trial weight alone. If we could balance this machine perfectly and then add the trial weight and measure the vibration, we would measure the same amplitude and phase as represented by this vector.

Now we have to measure the length of the “T” line from the graph paper, and the angle between the trial “T” vector and the original “O” vector. The length of the line is 6 cm (or major divisions), so that represents 6 mils. Using a protractor you can measure the angle as 40 degrees CCW (counter-clockwise).

So the solution requires the new mass to be added to the machine at a position 40 degrees counter-clockwise from where you placed the trial weight. But what should it weigh?

The weight is calculated as a ratio of the original vibration level “O” to the vibration due to the influence of the trial weight – our new vector “T” – multiplied by the mass of the trial weight. It is therefore (5 x 5/6) = 4.1 grams.

If the trial weight is removed

So, we need to remove the trial weight and add a new 4.1 gram weight 40° from where we placed the trial weight.

But the big question is; in which direction do we measure the 40°; clockwise or counter-clockwise? This is an area where a lot of balance jobs go wrong.

In modern vibration analyzers that use a photo-tach, ***positive*** phase angles are measured ***against*** the direction of rotation.

Therefore:

If the rotor is rotating **clockwise**, move **opposite** the O+T shift
If the rotor is rotating **counter-clockwise**, move **with** the O+T shift

In our example in Figure 15-24, the shaft is turning counter-clockwise, so we have to place the weight 40° against rotation, i.e. clockwise.

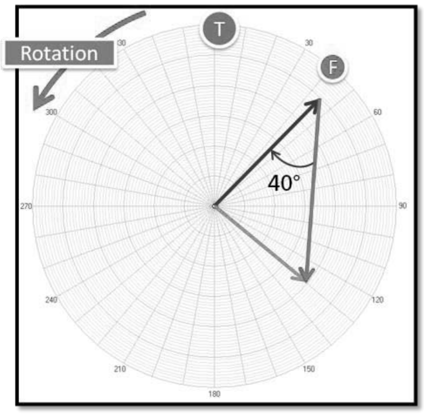


Figure 15-24

Note that if you use a strobe to measure phase, it reports positive phase angles in the direction of rotation.

Therefore:

If the rotor is rotating **clockwise**, move **with** the O+T shift
If the rotor is rotating **counter-clockwise**, move **opposite** the O+T shift

Trial weight left on

Everything we have discussed assumes that the trial weight will be removed before the final weight is added. But that may not suit your needs. You may like to leave the trial weight in place and therefore add a different final balance weight to the rotor.

Let's see how we can do that!

Originally we wanted a weight to counteract the "O" vector.

If we leave the trial weight on the rotor we must now add a final weight that will counteract the "O+T" vector. The desired effect of adding a weight is to have the *opposite* effect of the original unbalance PLUS the unbalance due to the trial weight. We need less weight and it needs to be placed in a different location.

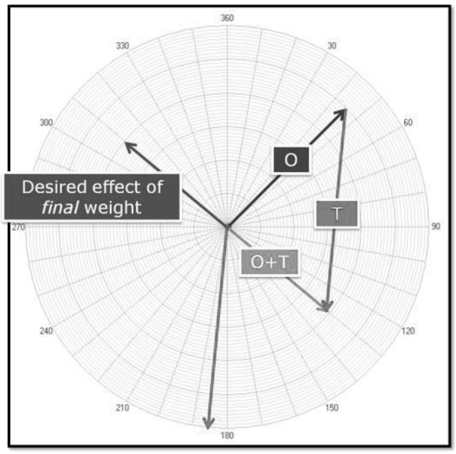


Figure 15-25

Measure the angle and measure the lengths of the O+T and T vectors (Figure 15-26) and compute the ratio.

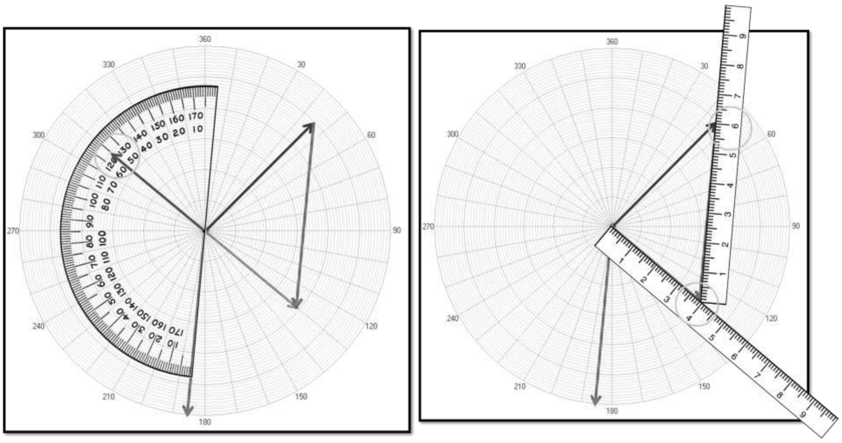


Figure 15-26

The solution is to add the weight 125° around from the location of the trial weight, *against* the direction of rotation.

The mass of the weight is:

Trial mass x (O+T)/T

5 x 4/6 = 3.3 grams

This assumes that the trial weight of 5 grams is not removed.

Residual unbalance

At this point we know what the unbalance is. We have just been told to add 3.3 grams. If we are adding that weight at a radius of 12 inches, or 300 mm, then our unbalance is 39.6 gr-in or 990 gr-mm.

This is $\mathbf{U_{res}}$ – it is the residual unbalance. We could now look up the standards and determine if that is permissible.

We will discuss this more in a bit.

Trim balance

Now we would add the 4.1 grams (we removed the trial weight) at 40° and run the machine again. We are hoping that the vibration afterwards would be very low. But we do not always find that... We often have vibration that is still a little too high.

In this case we measured 3 mil @ 320° (Figure 15-27).

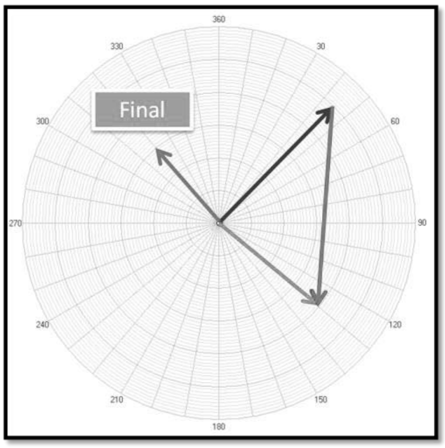


Figure 15-27

We need to perform another step to further reduce the vibration.

This step is called the *trim run*.

If you consider what we have just done, it is very similar to the trial weight process. We measured a vibration and phase, added a weight, and then measured the vibration again. Like the situation we discussed where we left the trial weight on the rotor, in this case we are going to leave the final weight on the rotor.

So now we have to use the vectors to determine a new weight and location.

Now we have re-labeled the plot (Figure 15-28). The “O” original vector is still the original. But the “trial” is actually the vibration measured with the trial weight removed and the final weight of 4.1 grams added.

Note: we will leave the final weight on the rotor, so we are trying to balance out “O+T1” not “O”.

We will label the trim run “T1” (our first trim run) so we have “O+T1” and the vector “T1” is just the effect of adding the 4.1 gram weight.

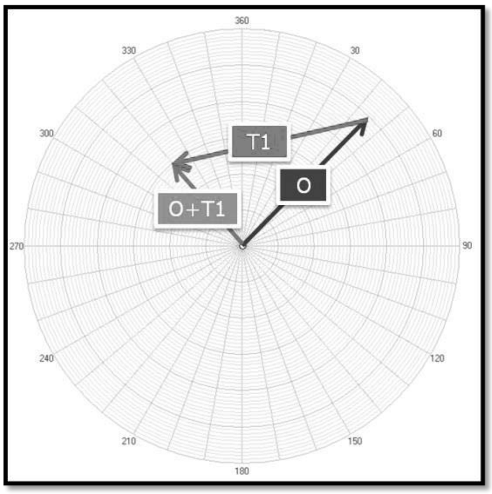


Figure 15-28

Now we are trying to counteract the unbalance force due to the final weight we added to the rotor (Figure 15-29).

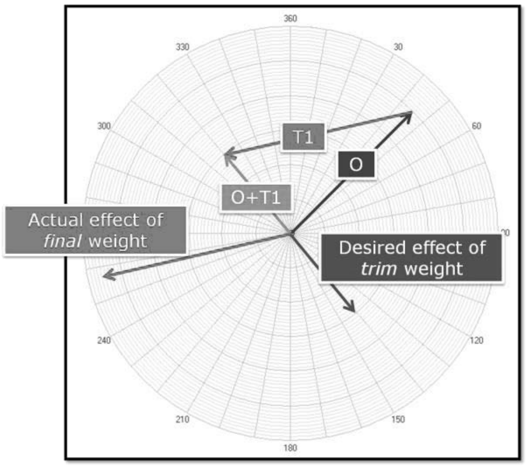


Figure 15-29

You can see that the red vector is too long so we need less weight, and it needs to be moved in the opposite direction by approximately 120°

If we measure the angle we get 116° (Figure 15-30). Measure the lengths of the O+T1 vector (which we know is 3 mils) and T1 vectors and compute the ratio – it tells us how much weight should now be added.

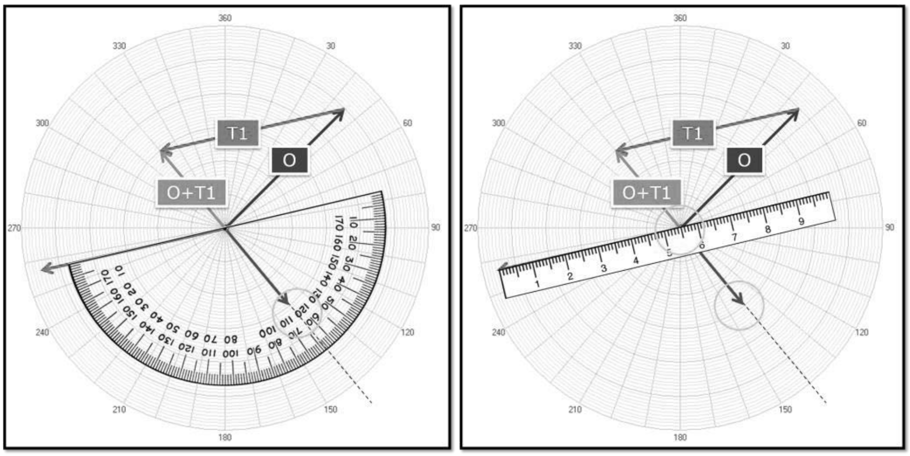


Figure 15-30

The solution is to add the balance weight 116° from where you placed the "final" weight *in the direction of rotation.*

The weight is the ratio:

Final mass x (O+T1)/T1

4.1 x 3.0/5.5 = **1.8 grams**

Note that "4.1" was the final weight we added to the machine; "3.0" was the measured vibration in mils after we added the final weight, and we computed the effect "T1" as 5.5 mils.

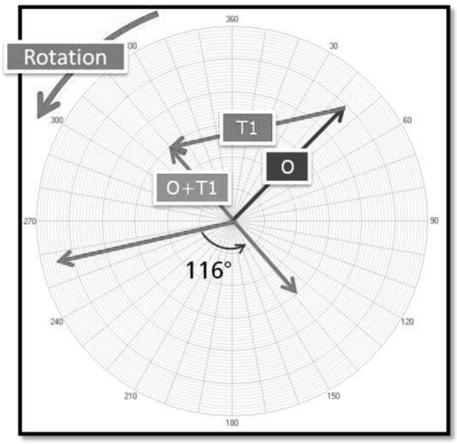


Figure 15-31

Once again we can calculate U_{res} and check if it is less than U_{per} based on G2.5 (or G1.0 or whatever you choose to use).

Of course, we can repeat this process again and again until the vibration amplitude, or U_{res} is low enough.

If you have to repeat it more than a couple of times the rule of thumb is to remove all of the weights and start from scratch. You should also consider other possible causes of your inability to balance the machine. These causes can include the presence of other mechanical faults such as looseness, resonance, misalignment, soft foot etc. Other possibilities include problems with the measurement equipment, phase readings, phase lag etc.

Splitting weights

When you calculate the final weight for the balance job, it must be added to the rotor. In some cases the weight can be added at any angle on the rotor. However some rotors have holes where weights are affixed or blades where weights can be attached. In this situation it may not be possible to add the weight at the precise angle required. The solution is to add two weights to the rotor (on available blades or holes) such that the combination of both weights has the same net effect as a single weight at the desired angle.

For example if the balance solution was to add 10 grams at 75°, yet we have blades at 60° and 120°, we would need to split the weight between the two blades so that the vectorial summation of the two weights is equal to 10 grams at 75°.

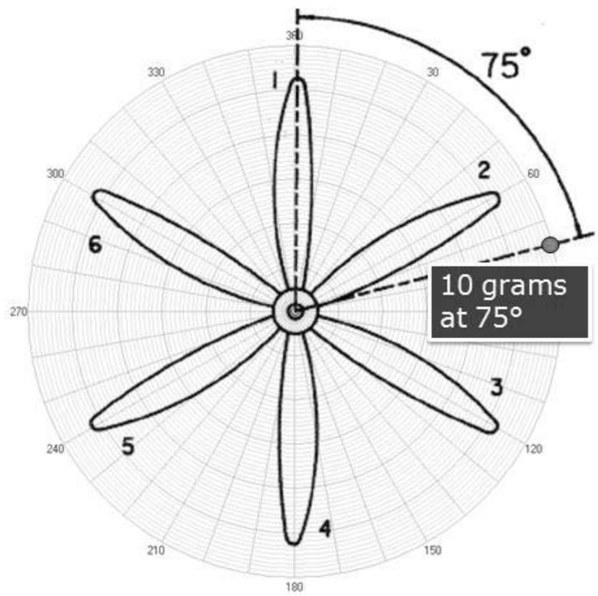


Figure 15-32

Note: The assumption is that the final weights will be placed on the blades at the same radius that the trial weights were added.

The following procedure may be used to determine the correct mass of the weights:

Draw the positions of the blades and desired balance weight on the polar plot

For this example I have drawn the blades, but you simply need to indicate where the two blades (or balance holes) are located. (We have to resolve the 10 gram weight at 75° into two equivalent vectors at the angles of the two blades.)

Add the vector that indicates the desired weight. In our case the mass is 10 grams and the angle is 75°. The weight is drawn at the limit of the polar plot.

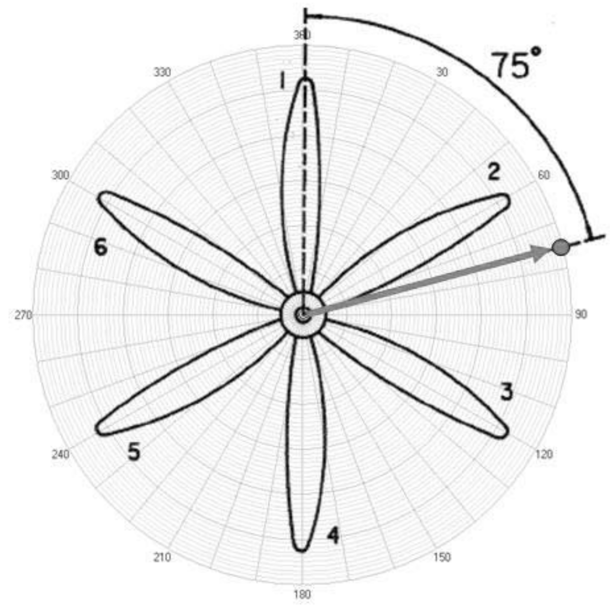


Figure 15-33

Draw the equivalence vectors for one of the blades

Draw a line parallel with blade 3 so that it goes through the point representing the balance mass. In this example we drew a line through blade 3 and a parallel line that goes through the weight.

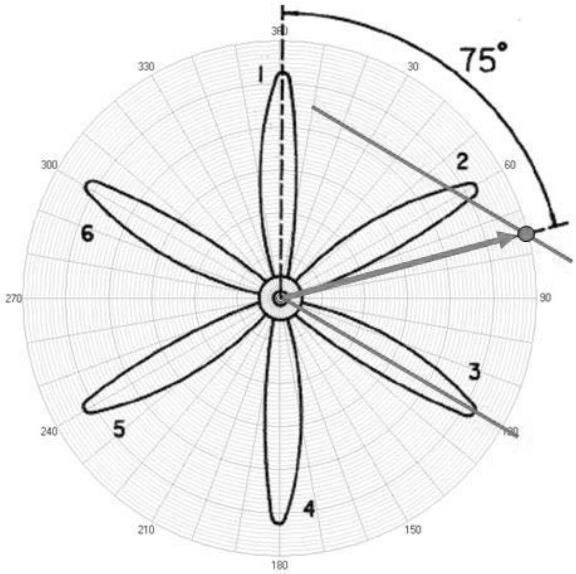


Figure 15-34

Draw the equivalence vectors for the other blade

Now we will repeat the process for the second blade. We will draw one line through the blade and a second line parallel with the blade that intersects the balance weight.

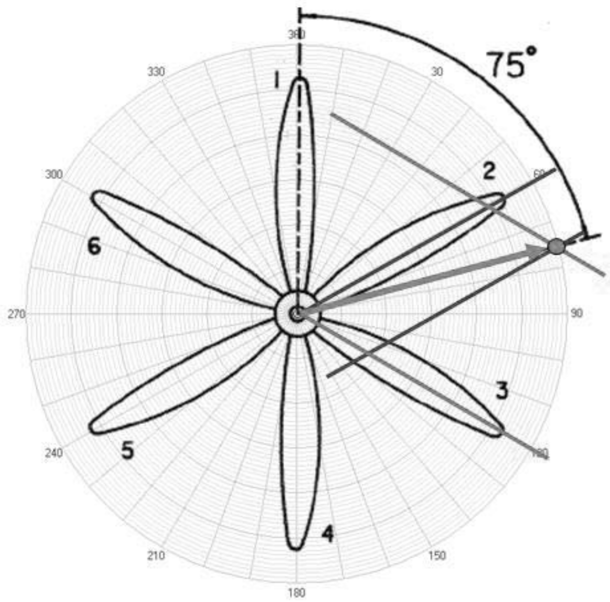


Figure 15-35

Add vectors through the blades to the intersection of the equivalence vectors

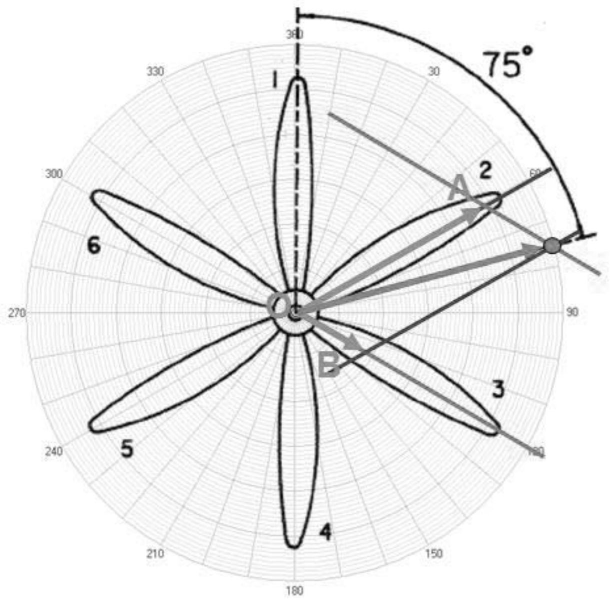


Figure 15-36

Determine the mass of the equivalence weights

Now we can compare the length of the two vectors to the length of the vector representing the desired balance mass. You can see that the vector "OA" is longer than the vector "OB", therefore more mass will be placed on blade 2.

In our example, the vector OA is approximately 70% the length of the length of the mass vector, so the mass on blade 2 will be 7 grams. And the vector OB is approximately 37% the length of the balance mass vector, so the mass on blade 3 should be 3.7 grams.

When 7 grams is added to blade 2, and 3.7 grams is added to blade 3, this fan will be balanced; we will have added the equivalent of 10 grams at 75°.

It is very important that you are not confused with the last step. Although we have drawn the OA and OB vectors along the blades, this does not indicate where on the blades the balance weights will be added. They must be added at the same radius as the trial weights. The vectors are simply used to help us compute what the mass of the weights should be.

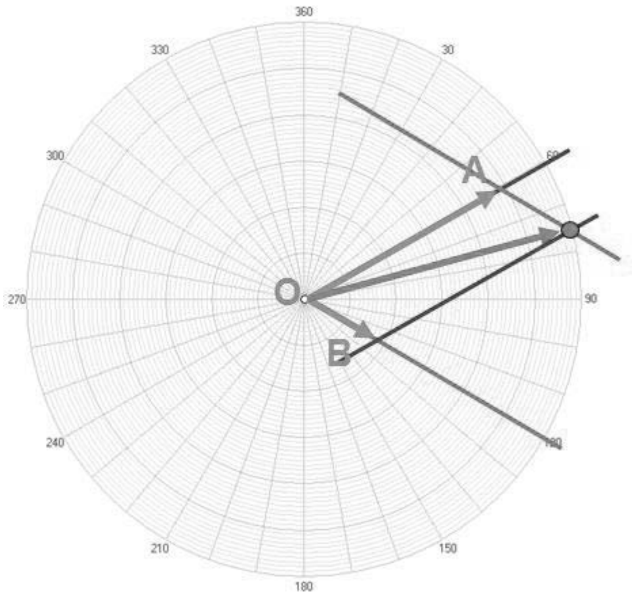


Figure 15-37 Find the two vectors that sum to the desired vector

Combining weights

When adding weights to a machine you may find that there are already one or more balance weights on the machine: attached to blades, bolted into holes, etc. There will be situations where it will become necessary to consolidate the existing weights with the weight(s) that must now be added to the machine.

The vector method of summing weights is the reverse of splitting weights which was covered in the previous pages. Instead of having a single weight and splitting it, start with two weights and sum them.

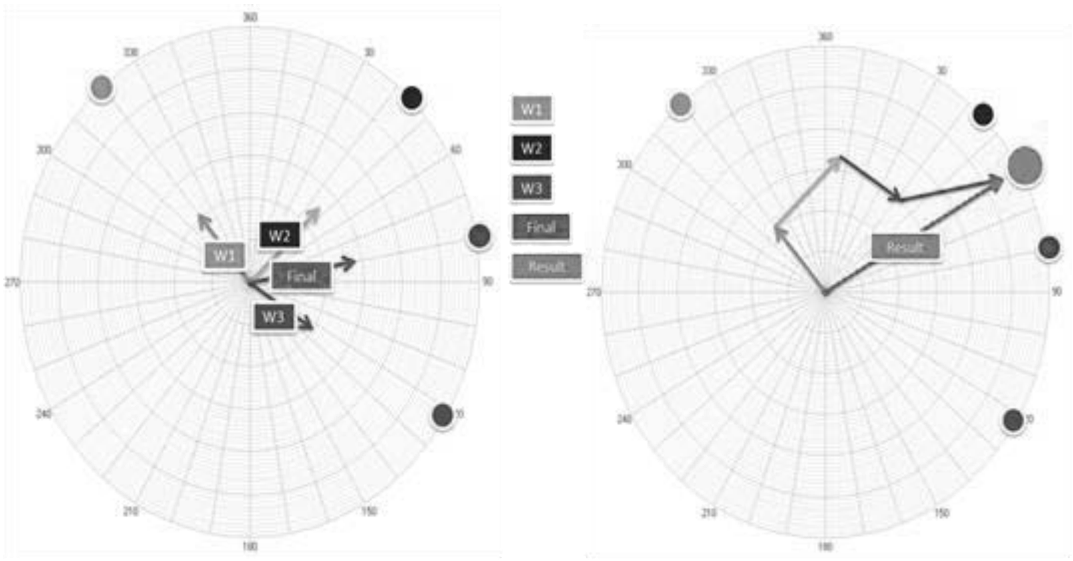


Figure 15-38

Draw a vector representation of each weight. Draw parallel lines across the tip of the opposite vector. Where these two new parallel lines intersect is the point of the Sum Vector.

Software programs can sum multiple weights at a time and are very useful tools.

There are times when adding correction weights is not permissible for various reasons. But mass can be removed. Most balancing programs have the option in the setup to choose whether to add weights or remove mass. Selecting the 'remove mass' option (if available) forces the analyzer to display the locations where weight should be removed instead of added. There are tables that provide information about weight removed for various diameter and depth of holes on various gauges of metal.

Four-Run No-Phase Balancing

The four run no phase balance method can be used to successfully balance a machine. This method does not require the balancer to collect a phase reading. This is important for two reasons; sometimes collecting a phase reading can be difficult, and secondly, single-plane and two-plane balancing methods require a stable phase reading which may not be possible. Resonances, beating with neighboring machines, shaft or foundation looseness, unstable foundations, instrument problems and other reasons can cause the phase readings to become unstable.

Also, the single plane method assumes that the machine is "linear", that is, that the machine will respond to 10 grams added at 70° in linear proportion to 20 grams added at 70°. For a variety of reasons, this may not be true. Another "flaw" in the single-plane and two-plane balance methods is that we are required to add a trial weight to the rotor and measure the influence of the weight, and from that, infer how the rotor will respond to a different weight at a different location. While this relates back to the linearity of the rotor and machine, it also relies on the balancer adding a trial weight of an appropriate size – if it is too small the balance calculations will be compromised. If the weight is too high the vibration levels may in turn be very high, and it may cause the machine to respond in a non-linear fashion.

As you can tell from the name of this method, it is necessary to perform four runs of the machine in order to determine the optimum balance weight. You may see this as a disincentive to use this method. However when you consider that the single plane method will use two runs as a minimum (original run, trial run and final run), and it is likely that you will need to perform a trim run, you can see that the four-run method is not as bad as it may seem.

It should also be noted that this method is still widely used today as it achieves very good results, and can be used when the single-plane or two-plane methods cannot be used, for the reasons discussed previously.

Balancing with the four-run method

The first step is to acquire an "original" vibration reading from the machine. A phase reading is not required, however it is important to take a reading at the turning speed frequency.

This reading should be plotted on polar graph paper. You should select a scale that will provide space for the additional readings that will be acquired. Note that the angle of the vector is not important; the circle is important. The circle is centered at the zero-point of the polar graph paper.

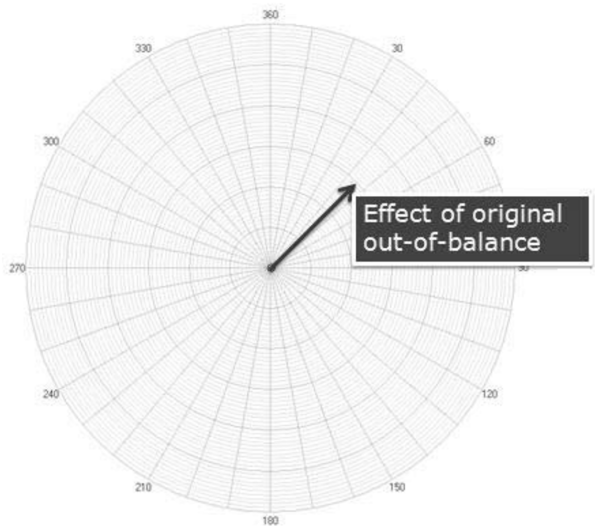


Figure 15-39 The original vector

It does not matter which units you use; acceleration, velocity, or displacement – just as long as you use the same units for all of the readings you are about to collect.

A circle should then be drawn with a radius equal to the vibration reading. 120° markings should be added to the circle. Note that it is possible to use different angles than 120°, however we will use 120° for now.

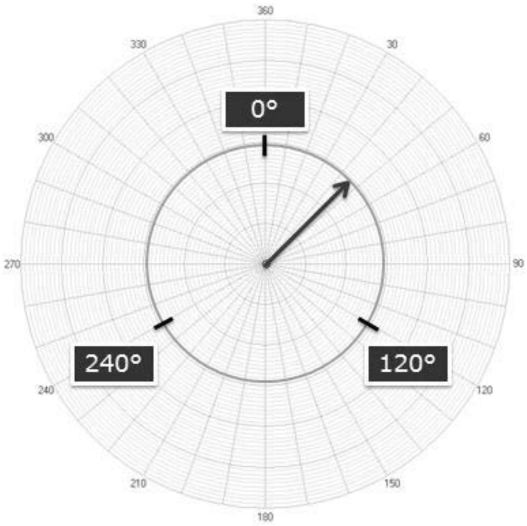


Figure 15-40 Draw the circle and mark out the 120° divisions

Now we must begin the three trial runs. A trial weight should be added at 0° on the rotor and the machine started and run until the vibration readings are stable. Please note that the mass of the trial weight should be selected using the same methods and calculations as described in the single-plane section. Once the amplitude reading has been taken it must be added to our polar plot. The center of the circle is placed at the 0° mark on the original circle.

For the sake of this example, we will assume that the vibration level was now higher than during the original run. I have drawn this circle so that it goes off the scale. You can either select a different grid spacing or just plot the portion of the circle that fits on the graph paper.

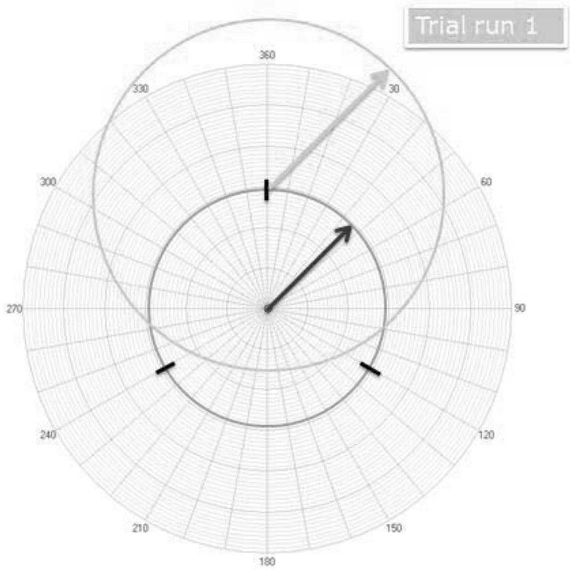


Figure 15-41 Draw the circle for the first balance run

Now you should stop the machine and move the same trial weight to a point 120 deg around from the first point. With that done, start the machine, wait for it to be running in a stable state, and then take another vibration reading.

By the way, if the reading is not stable because of beating for example, try and get an average value of vibration to use for the balance reading.

Plot this vibration amplitude on to the graph paper, but this time center the circle at the 120° intersection point. For the sake of this example, this reading had a lower amplitude of vibration.

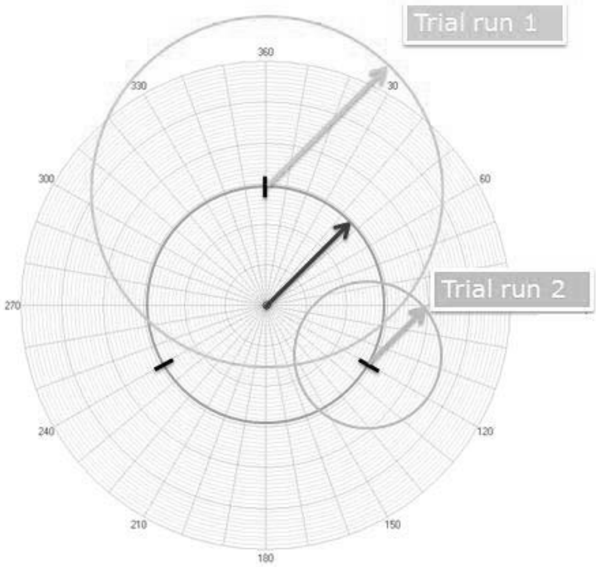


Figure 15-42 Draw the circle for the second run

And finally we repeat the process, however this time we move the trial weight to the 240° point and draw the circle centered at the 240° intersection point. For the sake of this example, the third reading had a very high vibration amplitude. It is clear that the final balance weight should be placed closer to the 120° position than the 240° position.

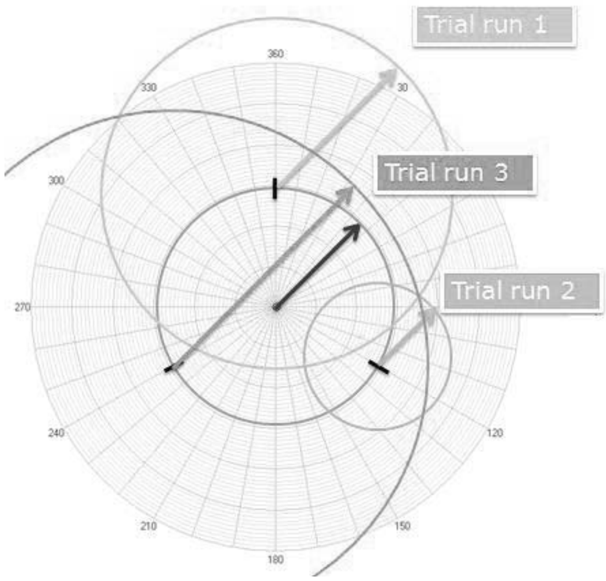


Figure 15-43 Draw the circle for the third run

Now that all three runs have been performed, and the three circles have been plotted accordingly, the mass and the position (angle) balance weight can be determined by looking for the intersection of the three circles.

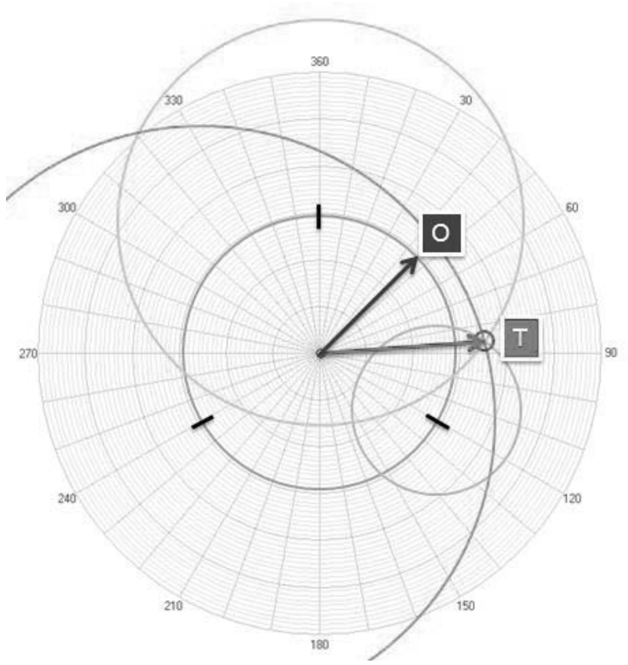


Figure 15-44 Find the intersection between the three circles

The final balance weight should be placed 86° from the 0° weight position, and the mass of the weight can be calculated in the same way as the single-plane weight – as a ratio of the length of the original vector to the solution vector, multiplied by the mass of the trial weight. In our example, let's assume that the trial weight weighed 10 grams, and the length of the solution vector was 3.6 vibration units compared with the original vibration value of 3. The mass should be: 10 x 3/3.6 = 8.33 grams.

$$Balance\ weight = \frac{O}{T}\ x\ Trial\ weight$$

What if the circles do not intersect?

In our example the three circles intersected perfectly. This will not always happen. Therefore you must find the geometric center of the region they create, as illustrated below.

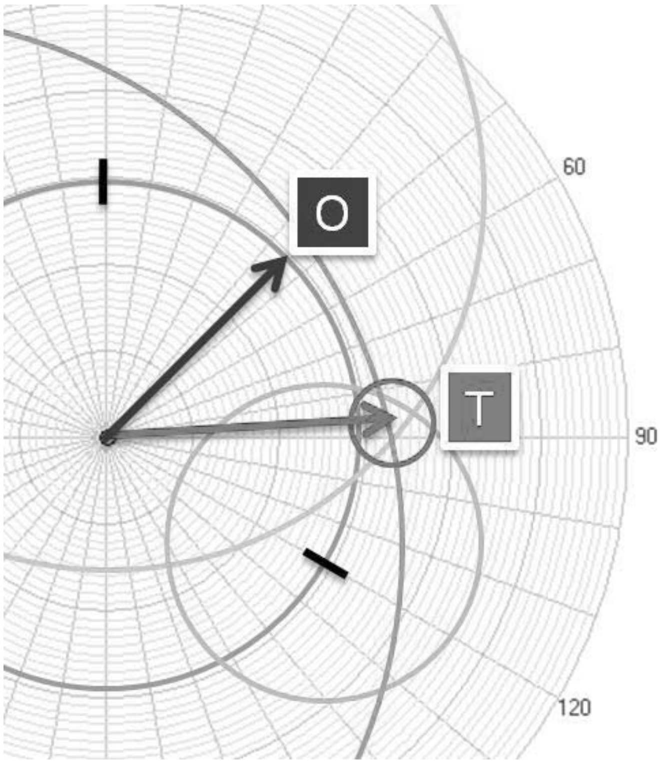


Figure 15-45 Find the geometric center

The following is another example where the circles do not intersect. The geometric center must still be identified and used to determine the mass and angle for the balance weight.

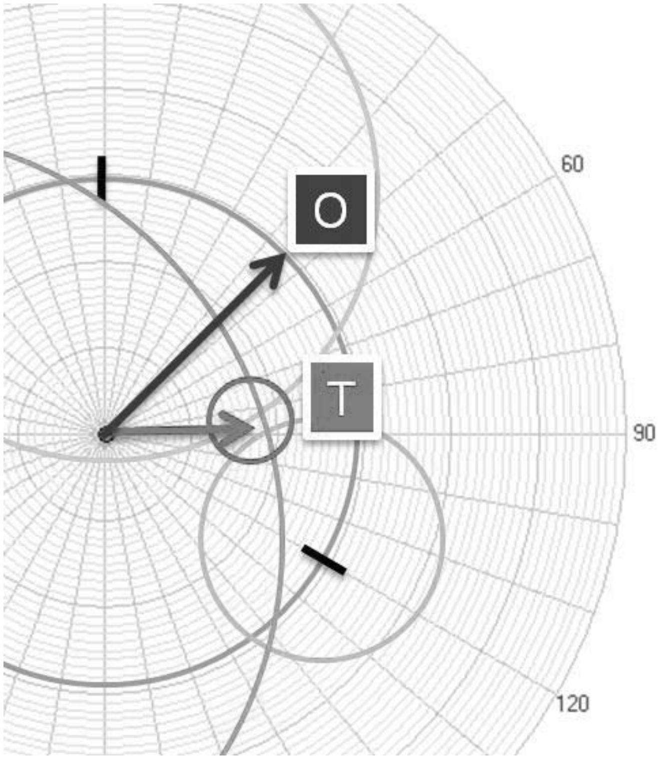


Figure 15-46 Find the geometric center

Two-Plane Balancing

Two plane balancing is required when there is dynamic unbalance (i.e. a static component and a couple). Although it is possible to perform a two-plane balance with vectors, it is not often performed that way and we will not discuss that method. Instead we will assume that you have a balancing system or a vibration analyzer with a balancing program.

It is beyond the scope of this course to go into much detail on two plane balancing, but we do feel it is important that you know what it is and when it is necessary.

The first issue we must quickly discuss is whether you *need* to perform a two-plane balance. There are basically four ways to determine whether a two plane balance is required:

1. You can allow the ISO standards to guide you.
2. You can decide based on a rule of thumb related to the ratio of the length of the rotor to the diameter of the rotor.
3. You can attempt a single plane balance, and if that is unsuccessful, attempt a two-plane balance
4. You can use phase readings to determine how the phase at the two bearings compare.

ISO Standards

The table from the ISO standard 19499:2007 provides guidance; not only to answer the single-plane/two-plane question, but also whether you need to treat the rotor as rigid or flexible. The different rotor configurations are presented with guidance related to the speed of operation and the first natural frequency (Figure 15-47).

Table 2 — Guidelines for balancing procedures

Guidance given in this table should be used with care as the ratio of rotor speed to first flexural critical speed *in situ* is only a typical value. This will be highly dependent on the dynamic characteristics of the bearings and their supports, initial unbalance, the balance quality required and the detailed design of the rotor.

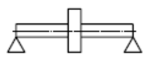
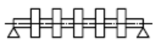
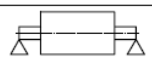
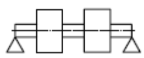
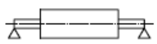
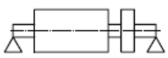
Schematic rotor representation	Rotor description		Rotor speed/first flexural critical speed *in situ*	Balancing process and International Standard
	Any rotor configuration		< 0,30	Single- or two-plane balancing ISO 1940-1
	Single rigid disc installed on an elastic shaft whose unbalance is negligible	Disc perpendicular to shaft axis	All values	Single-plane balancing ISO 1940-1
		Disc with axial runout	All values	Two-plane balancing ISO 1940-1
	Two rigid discs installed on an elastic shaft whose unbalance is negligible	Both discs perpendicular to shaft axis	All values	Two-plane balancing ISO 1940-1
		One or both disc(s) with axial runout	< 0,7	Two-plane balancing ISO 1940-1
			⩾ 0,7	Multiple-speed balancing ISO 11342:1998, procedure G
	More than two rigid discs installed on an elastic shaft whose unbalance is negligible		< 0,7	Two-plane balancing ISO 1940-1
			⩾ 0,7	Multiple-speed balancing ISO 11342:1998, procedure G
	Single rigid section installed on an elastic shaft whose unbalance is negligible		All values	Two-plane balancing ISO 1940-1

Table 2 (*continued*)

Guidance given in this table should be used with care as the ratio of rotor speed to first flexural critical speed *in situ* is only a typical value. This will be highly dependent on the dynamic characteristics of the bearings and their supports, the initial unbalance, the balance quality and the detailed design of the rotor.

Schematic rotor representation	Rotor description		Rotor speed/first flexural critical speed *in situ*	Balancing process and International Standard
	Two or more rigid sections installed on an elastic shaft whose unbalance is negligible		< 0,7	Two-plane balancing ISO 1940-1
			⩾ 0,7	Multiple-speed balancing ISO 11342:1998, procedure G
	Cylindrical roll	Assumed unbalance distribution according to machining tolerances (generally uniform unbalance distribution)	< 0,6	Two-plane balancing ISO 1940-1
			⩾ 0,6	Two-plane balancing in two optimum correction planes ISO 11342:1998, procedure F
		No known unbalance distribution	< 0,6	Two-plane balancing ISO 1940-1
			⩾ 0,6	Multiple-speed balancing ISO 11342:1998, procedure G
	Rotor including roll(s) and rigid sections and/or discs Integral rotors		< 0,6	Two-plane balancing ISO 1940-1
			⩾ 0,6	Multiple-speed balancing ISO 11342:1998, procedure G
	Any other configuration		No factor available	ISO 11342

NOTE 1 Factors stated are based on typical balance tolerances with limited initial unbalance.

NOTE 2 These guidelines apply to assembled rotors where the individual components have not been balanced prior to assembly.

NOTE 3 In the fourth column of the table, it is the first flexural critical speed of the rotor that is considered, not the rigid-body critical speeds.

NOTE 4 There are examples where the first critical speed observed is associated with rigid-body modes of the rotor. This should not be confused with the first flexural.

Figure 15-47

Rule of thumb

The general rule of thumb for determining whether a two plane balance is required is presented in Table 15-1. It appears in MIL-STD-167A (2005) and in numerous text books.

Rotor characteristics	Speed (RPM)	Type of balance
L/D ≤ 0.5	0 - 1000	Single-plane
	> 1000	Two-plane
L/D > 0.5	0 - 150	Single-plane
	> 150	Two plane

Table 15-1 (The Length and Diameter of the rotor are exclusive of the shaft)

While this method provides a guide, there are many situations where it will not work. It is best to use phase readings to determine if there is a couple component.

Two plane balancing is far more challenging than single-plane balancing. On the face of it, two-plane balancing may seem marginally more difficult as it only requires an additional trial run. While the calculations may be far more difficult, most analyzers can perform them quite easily.

Why is two plane balancing more difficult? The rotor must be rigid. There must be just two bearing supports, and they must be flexible, but not resonant. And the system must be linear. (Wowk)

In the single-plane balancing section we provided a rule of thumb regarding when two-plane balancing must be used instead of single-plane balancing. In reality, the issue is simpler. If there is a strong couple effect, or a strong cross-effect between the two bearings, then you cannot use the single-plane method. As long as the machine meets the requirement above, you may use the two-plane method.

It is therefore recommended that you attempt single-plane balance first. There are times when one end of the rotor has much more amplitude than the other end. Some technicians choose to perform a single plane balance on the rough end first and then do the opposite end. Sometimes bringing the rough end into specs reduces the amplitude at the opposite end enough that it does not have to be balanced.

A benefit of the two plane balance is that all planes of the machine are corrected in one process.

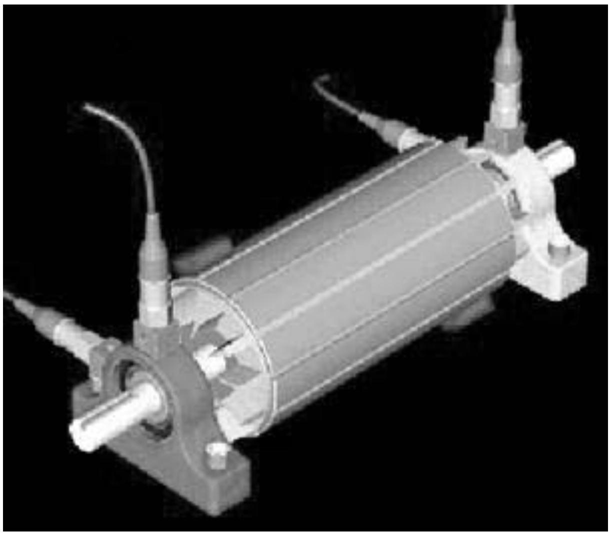


Figure 15-48 A rotor set up for two plane balancing with four sensors

At least one sensor is required for each plane. The sensor is typically mounted in the horizontal plane because there is likely to be greatest movement in the horizontal axis. Most software packages can perform calculations for up to two sensors at each plane.

Here is a summary of the two-plane method:

- The machine is run at normal speed
- Amplitude and phase is recorded at both points
- The machine is stopped. A trial weight is added to one plane.
  The machine is run and readings are taken at both points.
- The machine is stopped. The first trial weight is removed. A trial weight is added to the other plane.
- The machine is run and readings are taken at both points.

The original run

After following all of the safety procedures discussed previously, you will take vibration readings on each bearing. These readings indicate the original unbalance state.

It is important to remember to double-check that the machine is out-of-balance and not misaligned or suffering some other fault condition that generated high vibration.

Trial run one

Now the machine is stopped and a trial weight is added to **one** of the balance planes. You must add enough weight to the shaft so that, ideally, the amplitude and phase will both change sufficiently for the balance program (or your vector plots if you wish to do it that way) to be able to accurately determine what influence that mass at that location had on the balance of the shaft. You would hope for a minimum of 15% change in amplitude and/or 15° change in phase angle.

You should therefore perform the trial weight calculations that have been discussed in the single-plane section.

The trial weight can be placed at any angular position on the shaft. You will need to record that information in the balance program. The simplest approach is to place the weight at 0°. Alternatively you can place the weight opposite the high spot.

Trial run two

Now you must remove the first trial weight (some balance programs allow you to leave it on) and add a weight to the second plane.

As just discussed, the balance program will measure the amplitude and phase and determine how much that mass influenced the balance state.

Balance calculation

The two-plane balance programs use an "influence coefficient" method. It is designed to determine the optimum position for a weight on the two balance planes that will minimize the vibration levels. Some programs allow you to optimize the results over different speeds, and they allow you to give greater preference to one end of the machine or the other.

The bottom line is that you will be given a mass and angle for each plane. You must then install those weights, after removing the trial weights, and run the machine again to see if the vibration amplitude is sufficiently low (see the balance grades section).

Trim run

If the amplitudes are not low enough, the balance program will suggest new masses and positions for trim weights. You typically have the choice of either leaving the "final" weights on the machine or removing them before adding the trim weights.

If the system is linear then this method should give you good results.

Balancing overhung rotors

Balancing an overhung rotor presents a unique challenge. Conventional single and two plane methods do not work because of the "rocking" motion of the rotor. We will present two methods for balancing an overhung rotor.

In order to perform the two techniques we must identify the two measurement points ("A" and "B"), and the two balance planes ("1" and "2") as illustrated in Figure 15-49.

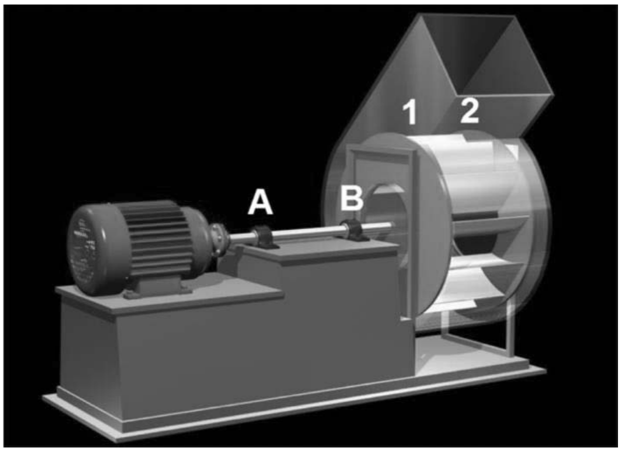


Figure 15-49

The basic theory behind these methods is that the unbalance will comprise of a static component and a couple component. It is believed that the static component can be corrected by adding weight to balance plane "1" and that it will affect the vibration at bearing "B" more than bearing "A".

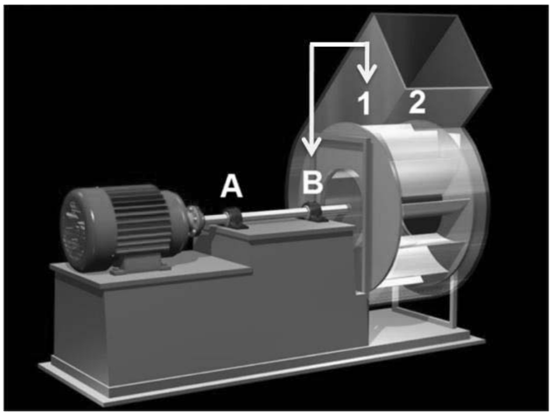


Figure 15-50

It is also believed that the couple component will affect the bearing closest to the motor (bearing "A"), and it is believed that the couple effect can be corrected by adding a weight to plane "2" and a weight of the same mass, but opposite angle, on plane "1".

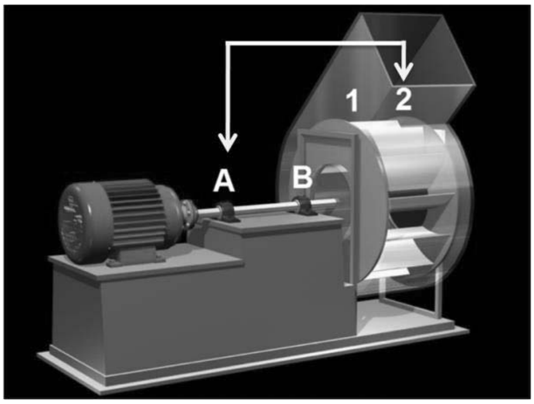


Figure 15-51

Single plane method

The first method we will use will attempt to balance the static component first, and then the couple component, but using a single-plane balancing technique. It is assumed that you understand the single plane balancing method.

STEP ONE: We will begin by performing a standard, single plane balance using the bearing closest to the rotor "B" for the measurement location, and the inboard balance (correction) plane, plane 1, for the weights.

STEP TWO: Now we must correct the couple component. We will again perform a single plane balance, *but with a difference*. The vibration sensor will be mounted on bearing "A" and we will use the second correction plane (plane 2) for trial weights and final weights.

However, whenever a trial weight or final weight is attached to plane 2, a weight of equal mass will be attached to plane 1, but at the opposite angle.

STEP THREE: When the couple has been corrected, it is important to check the static component. Place the sensor at bearing "B" again and check the vibration. If the amplitude (or residual unbalance U_{res}) is still too high, you should repeat STEP ONE.

Two plane method

STEP ONE: We will start like any other two-plane balance job. The two accelerometers will be placed at bearings "A" and "B". We will run the machine and record the "original" run.

STEP TWO: Then, as normal, we will place a trial weight on plane "1" and measure the vibration at bearings "A" and "B".

STEP THREE: But now we do something different. We will remove the trial weight from plane "1" and move it to plane "2". However, we will also place a weight on plane "1" which is equal in mass, but located 180° opposite the trial weight on plane "2".

STEP FOUR: The balance program will come up with a solution for weights that must be placed on plane 1 and 2. The weight on plane 1 should be added as reported. However, the recommended weight should be added to plane 2, and another weight with the same mass should be placed on plane 1 at the opposite angle.

After all of these steps are complete the rotor should be balanced; the static and couple components should be corrected.

Balancing machines with flexible rotors

Thus far we have discussed how to balance rigid rotors. Rigid rotors turn at speeds that are well below their critical speed. There will be no deflection in the shaft.

Flexible rotors do operate close to or above the critical speed. They will deflect according to their first, second and/or third flexural or bending modes. Depending upon the stiffness of the bearings, it may also move according to the first and second rigid body modes.

Flexible rotors include paper machine rolls, spindles and long shafts, and high-speed turbines, multi-stage pumps, generators, and multi-stage compressors. (Wowk)

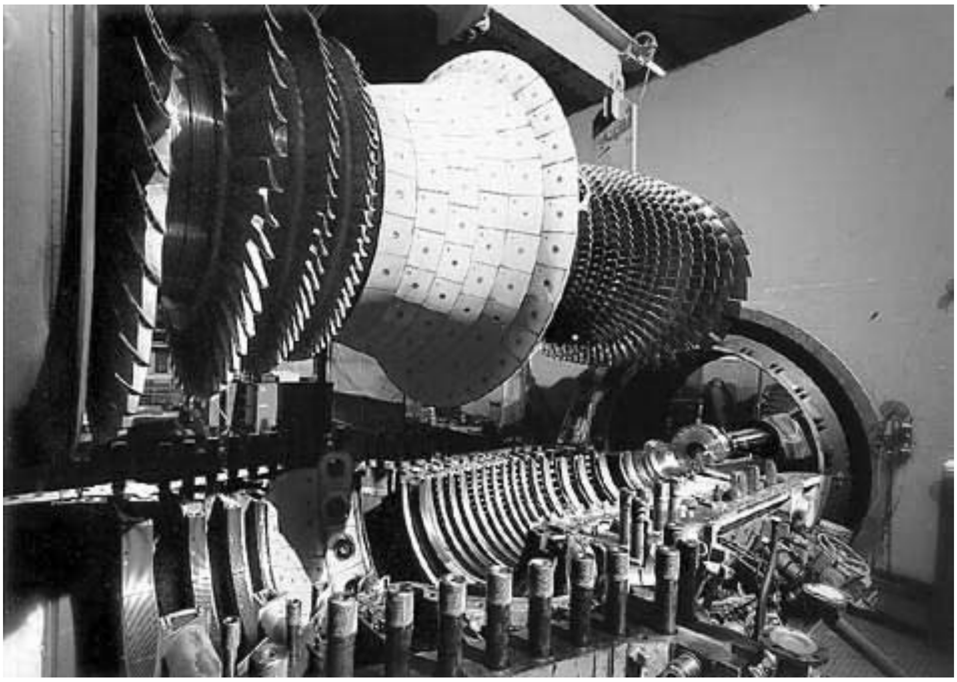


Figure 15-52

The challenge of the balancer is to place balance weights that will minimize the bending moments in the shaft. The correction weights must be placed in the same axial plane as the unbalance, otherwise additional bending will occur. In long shafts as may be found in a paper machine, it can be difficult to know where the unbalance exists. With rotors found in turbines and compressors, the unbalance will exist in the fan disks and thus they can be balanced more easily.

Whereas rigid rotors will generate high vibration at the bearings, flexible rotors generate far lower vibration at the bearings, and instead generate greater movement/bending in the shaft; which may cause it to fatigue and fail.

Performing the balance is beyond the scope of this course. Suffice to say that rigid-rotor methods can be used to remove gross unbalance, however additional methods are available including modal balancing, influence coefficient, and multi-speed, multi-plane method.

Balance standards

When balancing a machine the obvious question can arise; are we there yet? When you add the final weights, the vibration amplitude will (should) be reduced – but it will never be entirely eliminated. But is the amplitude low enough? You can continually perform trim runs, but again, how do you know when to stop?

We will look at four methods, with a focus on the fourth option:

1. Using the “lights-out” method when using a stroboscope
2. Using generic values of unbalance
3. Using standards based on vibration amplitude
4. Using standards based on residual unbalance

“Lights-out” balancing threshold

When balancing a rotor, the 1X peak in the spectrum will be reduced in amplitude. When using a strobe that is triggered by the accelerometer, the trigger will no longer work when the amplitude is low (compared to other sources of vibration). Therefore some people consider the balance job is done when they can no longer trigger their strobe. Of course, there is no repeatability to this method; it depends so much on the individual circumstances.

Generic unbalance specification

Some people will also use a generic value of unbalance in oz-in, gr-in or gr-mm to specify the permissible unbalance. However the speed of the rotor must also be included in such a specification in order to be meaningful. As discussed in the balancing theory section, the force generated is proportional to the speed squared. A generic value of unbalance cannot be used for all rotors.

Balance standards

Balance standards exist and are commonly used to specify a balance job. Using a balance standard does provide repeatability and accountability. Manufacturers and corporations may also have their own criteria for balance quality, but we will focus on the published standards.

There are two types of standards: standards based on the final amplitude of vibration, and standards that specify the "permissible residual unbalance". The residual unbalance is the unbalance that remains once the balance job is complete.

Amplitude limits

It is not normally recommended that standards based on amplitude level be used for determining balance quality. Although ISO 14694 is designed specifically for fans, the ISO 10816 and ISO 7919 standards are designed for assessing overall condition. They are normally used initially to determine the necessity for balancing, but not to determine whether the balance is good enough. This section is included because these standards are sometimes used for this purpose.

The ISO standards presented in this section focus on amplitude of vibration readings taken from machines of different types (fans, steam turbines, gas turbines, reciprocating machines, and others), measurements of different type (machines tested with accelerometers and machines tested with non-contact displacement probes) and different mounting (flexible and rigid).

We will now take a look at these standards.

ISO 7919

ISO 7919: Mechanical vibration of **non-reciprocating** machines – **Measurements on rotating shafts** and evaluation criteria

This standard provides guidelines for measurement and evaluation criteria for a variety of machine types:

Part 1 – General Guidelines
Part 2 - Land-based steam turbines and generators in excess of 50MW with normal operating speeds of 1500 r/min, 1800 r/min, 3000 r/min and 3600 r/min
Part 3 - Coupled industrial machine
Part 4 - Gas turbine sets
Part 5 - Machine sets in hydraulic power generating and pumping plants

ISO 10816

ISO 10816: Mechanical vibration – Evaluation of machine vibration by measurements on **non-rotating parts**

This standard provides guidelines for measurement and evaluation criteria for a variety of machine types:

Part 1 - General guidelines
Part 2 - Land-based steam turbines and generators in excess of 50 MW with normal operating speeds of 1500 r/min, 1800 r/min, 3000 r/min and 3600 r/min
Part 3 - Industrial machines with nominal power above 15kW and nominal speeds between 120 r/min and 15000 r/min when measured in situ
Part 4 - Gas turbine sets with fluid film bearings
Part 5 - Machine sets in hydraulic power generating and pumping plants
Part 6 - Reciprocating machines with power ratings above 100 kW

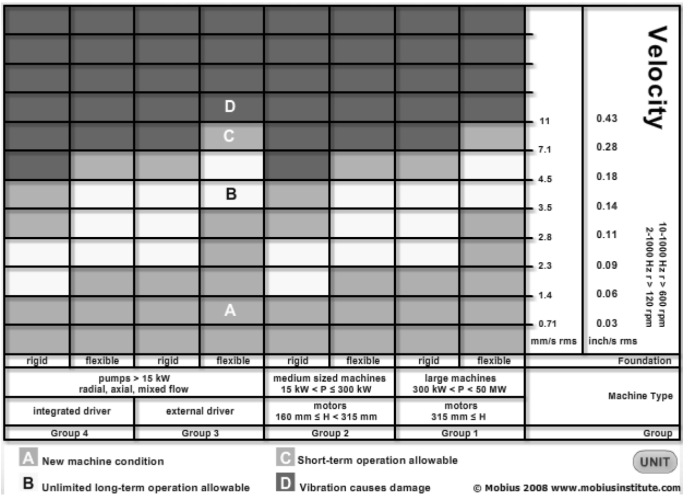


Figure 15-53 Velocity limits for industrial machines

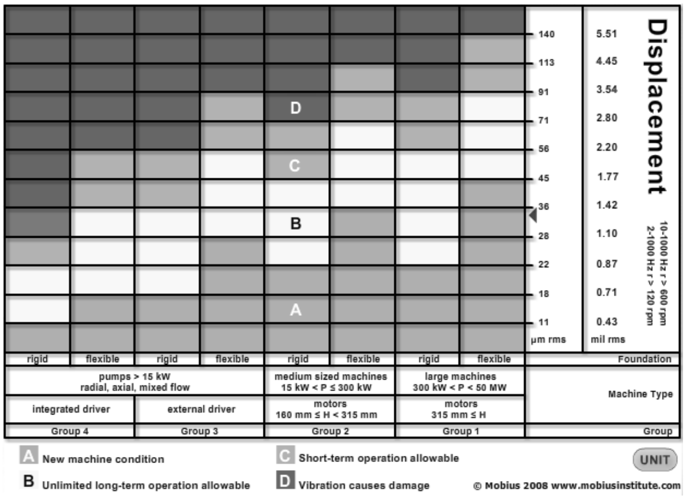


Figure 15-54 Displacement limits for industrial machines

ISO 14694: 2003

ISO 14694: *Industrial fans – Specifications for balance quality and vibration levels*

This standard is specifically designed for fans. It describes the measurement locations, vibration amplitudes, and satisfactory/alarm/shutdown limits. We are not allowed to reproduce the standard, but we can share three key tables in Figure 15-55 to Figure 15-57.

Table 1 — Fan-application categories

Application	Examples	Limits of driver power kW	Fan-application category, BV
Residential	Ceiling fans, attic fans, window AC	⩽ 0,15	BV-1
		> 0,15	BV-2
HVAC and agricultural	Building ventilation and air conditioning; commercial systems	⩽ 3,7	BV-2
		> 3,7	BV-3
Industrial process and power generation, etc.	Baghouse, scrubber, mine, conveying, boilers, combustion air, pollution control, wind tunnels	⩽ 300	BV-3
		> 300	See ISO 10816-3
Transportation and marine	Locomotive, trucks, automobiles	⩽ 15	BV-3
		> 15	BV-4
Transit/tunnel	Subway emergency ventilation, tunnel fans, garage ventilation, Tunnel Jet Fans	⩽ 75	BV-3
		> 75	BV-4
		none	BV-4
Petrochemical process	Hazardous gases, process fans	⩽ 37	BV-3
		> 37	BV-4
Computer-chip manufacture	Clean rooms	none	BV-5

NOTE 1 This standard is limited to fans below approximately 300 kW. For fans above this power refer to ISO 10816-3. However, a commercially available standard electric motor may be rated at up to 355 kW (following an R20 series as specified in ISO 10816-1). Such fans will be accepted in accordance with this International Standard.

NOTE 2 This Table does not apply to the large diameter (typically 2 800 mm to 12 500 mm diameter) lightweight low-speed axial flow fans used in air-cooled heat exchangers, cooling towers, etc. The balance quality requirements for these fans shall be G 16 and the fan-application category shall be BV-3.

Figure 15-55

Table 4 — Vibration-levels limit for test in manufacturer's work-shop

Fan application category	Rigidly mounted mm/s		Flexibly mounted mm/s	
	Peak	r.m.s.	Peak	r.m.s.
BV-1	12,7	9,0	15,2	11,2
BV-2	5,1	3,5	7,6	5,6
BV-3	3,8	2,8	5,1	3,5
BV-4	2,5	1,8	3,8	2,8
BV-5	2,0	1,4	2,5	1,8

NOTE 1 Refer to Annex A for conversion of velocity units to displacement or acceleration units for filter-in readings.

NOTE 2 The r.m.s. values given in this Table are preferred. They are rounded to a R20 series as specified in ISO 10816-1. Peak values are widely used in North America. Being made up of a number of sinusoidal wave forms, these do not necessarily have an exact mathematical relationship with the r.m.s. values. They may also depend to some extent on the instrument used.

NOTE 3 The values in this Table refer to the design duty of the fan and its design rotational speed and with any inlet guide vanes “full-open”. Values at partial load conditions should be agreed between the manufacturer and user, but should not exceed 1,6 times the values given.

Figure 15-56

Table 5 — Seismic vibration limits for tests conducted *in situ*

Condition	Fan-application category	Rigidly mounted mm/s		Flexibly mounted mm/s	
		Peak	r.m.s.	Peak	r.m.s.
Start-up	BV-1	14,0	10	15,2	11,2
	BV-2	7,6	5,6	12,7	9,0
	BV-3	6,4	4,5	8,8	6,3
	BV-4	4,1	2,8	6,4	4,5
	BV-5	2,5	1,8	4,1	2,8
Alarm	BV-1	15,2	10,6	19,1	14,0
	BV-2	12,7	9,0	19,1	14,0
	BV-3	10,2	7,1	16,5	11,8
	BV-4	6,4	4,5	10,2	7,1
	BV-5	5,7	4,0	7,6	5,6
Shutdown	BV-1	Note 1	Note 1	Note 1	Note 1
	BV-2	Note 1	Note 1	Note 1	Note 1
	BV-3	12,7	9,0	17,8	12,5
	BV-4	10,2	7,1	15,2	11,2
	BV-5	7,6	5,6	10,2	7,1

NOTE 1 Shutdown levels for fans in fan-application grades BV-1 and BV-2 should be established based on historical data.

NOTE 2 The r.m.s. values given in this Table are preferred. They are rounded to a R20 series as specified in ISO 10816-1. Peak values are widely used in North America. Being made up of a number of sinusoidal wave forms, these do not necessarily have an exact mathematical relationship with the r.m.s. values. They may also depend to some extent on the instrument used.

Figure 15-57

Residual unbalance

Now we will look at the standards that define balance quality in terms of "residual permissible" unbalance.

When you take readings on a machine after you have added a trial weight, you will know the amount of unbalance "U" in oz-in, gr-in, or gr-mm. You will then add weight(s) and take a new reading, and compute the unbalance "U" once again. At that point you could perform a trim balance to further reduce the unbalance.

At any point in the process there is therefore a "residual unbalance", U_{res}, which is the amount of unbalance that still remains.

The following standards provide a guide as to how much residual unbalance is *permissible* (or allowed). This is named U_{per}.

The standards are:

1. ISO 1940-1:2003 (and ANSI S2.19-1989)
2. American Petroleum Institute: API (standard for centrifugal pumps and compressors)

3. MIL-STD-167-1A (2005) Mechanical Vibrations of Shipboard Equipment [supersedes MIL-STD-167-1 1974 (SHIPS)]

The first task, therefore, is to determine the residual unbalance that exists and then refer to the standard to determine if that amount of unbalance is permissible. Therefore it is *not* possible to simply walk up to a machine, take a vibration reading and determine if the vibration level is acceptable (unless you wish to use the standards presented in the previous section).

Quick review:

The quantity "Unbalance" is discussed in detail in the "Balancing theory" section. In that section it describes that we can describe the unbalance in terms of the force generated by an unbalance mass [m] at a set radius [r] on a shaft rotating at a given speed (ω or RPM), AND it can be described in terms of the eccentricity [e] created due to a rotor of mass [M] rotating.

The following information is from that section and will provide a reminder.

The unbalance "U" and centrifugal force "F" can be computed based on the mass of the unbalance [m] operating at the radius [r]:

$$U = m\,r$$

$$F = m\,r\,\omega^2$$

$$\omega \approx RPM / 10$$

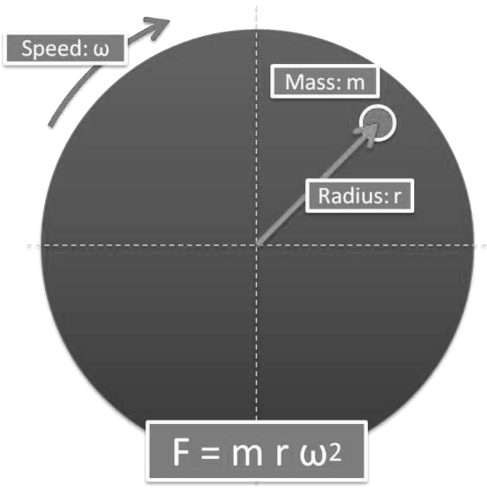


Figure 15-58

The unbalance "U" and centrifugal force "F" can also be computed based on the mass of the rotor [M] and the eccentricity [e] due to the unbalance:

$$U = Me$$

$$F = Me\,\omega^2$$

$$\omega \approx RPM / 10$$

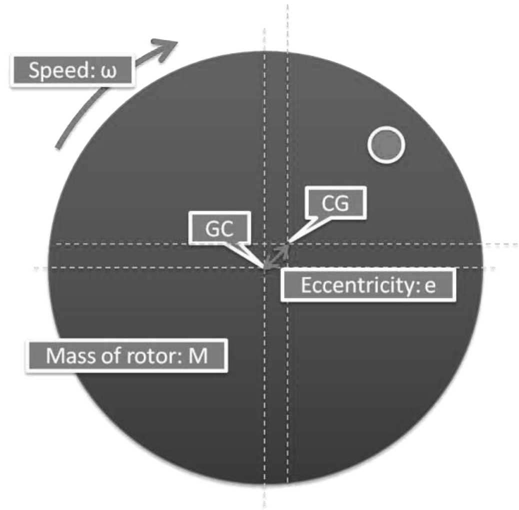


Figure 15-59

In the next section we will look at how we can use the measured/calculated residual unbalance to determine if the balance job has achieved a high enough level of quality.

Residual unbalance: ISO 1940

ISO 1940-1:2003 and ANSI S2.19-1989 provide a means of determining the permissible residual unbalance based on the type of machine. It uses quality grade numbers: G1.0, G2.5, G6.3, etc. In Figure 15-60 is an excerpt from the standard. It shows that G6.3 covers many of the industrial machines that are found in industrial plants, however these days, G6.3 is not acceptable. G2.5 is more common, and we would argue that G1.0 should be the goal.

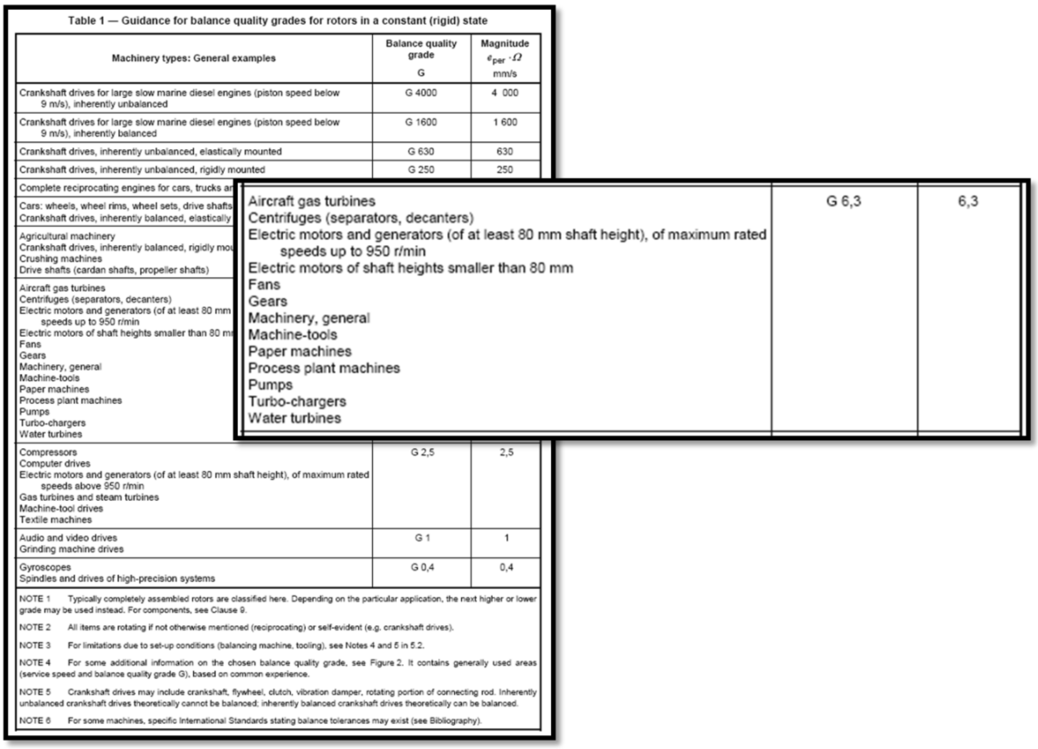


Table 1 — Guidance for balance quality grades for rotors in a constant (rigid) state

Machinery types: General examples	Balance quality grade G	Magnitude $e_{per} \cdot \Omega$ mm/s
Crankshaft drives for large slow marine diesel engines (piston speed below 9 m/s), inherently unbalanced	G 4000	4 000
Crankshaft drives for large slow marine diesel engines (piston speed below 9 m/s), inherently balanced	G 1600	1 600
Crankshaft drives, inherently unbalanced, elastically mounted	G 630	630
Crankshaft drives, inherently unbalanced, rigidly mounted	G 250	250
Complete reciprocating engines for cars, trucks ar	[illegible]	[illegible]
Cars: wheels, wheel rims, wheel sets, drive shafts Crankshaft drives, inherently balanced, elastically	[illegible]	[illegible]
Agricultural machinery Crankshaft drives, inherently balanced, rigidly mou Crushing machines Drive shafts (cardan shafts, propeller shafts)	[illegible]	[illegible]
Aircraft gas turbines Centrifuges (separators, decanters) Electric motors and generators (of at least 80 mm speeds up to 950 r/min Electric motors of shaft heights smaller than 80 mr Fans Gears Machinery, general Machine-tools Paper machines Process plant machines Pumps Turbo-chargers Water turbines	[illegible]	[illegible]
Compressors Computer drives Electric motors and generators (of at least 80 mm shaft height), of maximum rated speeds above 950 r/min Gas turbines and steam turbines Machine-tool drives Textile machines	G 2,5	2,5
Audio and video drives Grinding machine drives	G 1	1
Gyroscopes Spindles and drives of high-precision systems	G 0,4	0,4

NOTE 1 Typically completely assembled rotors are classified here. Depending on the particular application, the next higher or lower grade may be used instead. For components, see Clause 9.

NOTE 2 All items are rotating if not otherwise mentioned (reciprocating) or self-evident (e.g. crankshaft drives).

NOTE 3 For limitations due to set-up conditions (balancing machine, tooling), see Notes 4 and 5 in 5.2.

NOTE 4 For some additional information on the chosen balance quality grade, see Figure 2. It contains generally used areas (service speed and balance quality grade G), based on common experience.

NOTE 5 Crankshaft drives may include crankshaft, flywheel, clutch, vibration damper, rotating portion of connecting rod. Inherently unbalanced crankshaft drives theoretically cannot be balanced; inherently balanced crankshaft drives theoretically can be balanced.

NOTE 6 For some machines, specific International Standards stating balance tolerances may exist (see Bibliography).

Aircraft gas turbines Centrifuges (separators, decanters) Electric motors and generators (of at least 80 mm shaft height), of maximum rated speeds up to 950 r/min Electric motors of shaft heights smaller than 80 mm Fans Gears Machinery, general Machine-tools Paper machines Process plant machines Pumps Turbo-chargers Water turbines	G 6,3	6,3

Figure 15-60

Note: The author believes that a great deal of confusion has been generated by the ISO 1940 standard because of the units used in the magnitude column: mm/s. For example, if the balance grade G2.5 were used, it would be easy to assume that it meant the vibration amplitude would be 2.5 mm/s (0.1 in/s); which is quite high.

In actual fact, the "mm/s" refers to the product of eccentricity [e] in units of "mm" and angular velocity (typically ω but listed here as Ω) which has units of "rad/sec" which makes it "mm/s".

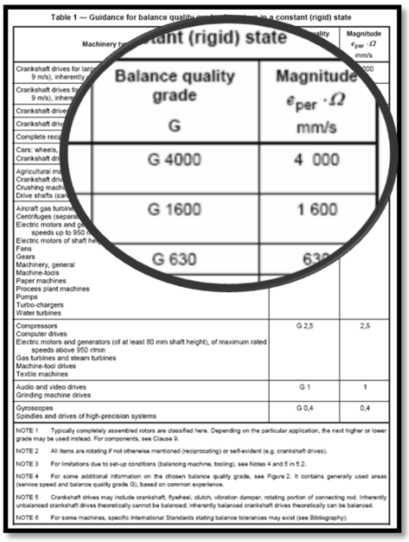


Figure 15-61

If you work through all of the equations, it turns out that if the rotor was operating in space then it would generate 2.5 mm/s (0.1 in/s). But I assume you are not working in space…

There are two ways to apply these standards.

First, the standard provides a chart for "easy" look-up. You can search along the bottom of the chart (see the next slide) for the RPM of the machine in question. You then search upwards for the G grade number you wish to apply, and then you move across to the y-axis for the "Permissible residual specific unbalance" (g-mm/kg) per correction plane. Knowing this value, and the rotor mass, we can compute how much unbalance is permissible in each correction (balance) plane.

In the example in Figure 15-62 the speed is approximately 1800 RPM and we are aiming for G2.5.

The unbalance permissible is 20 g-mm per kg of rotor.

For a rotor of 10 kg we can allow 200 g-mm of unbalance.

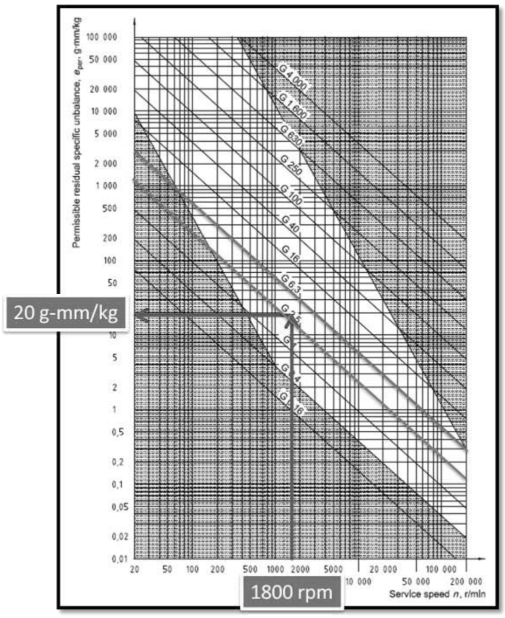


Figure 15-62

The **metric** version of the chart in Figure 15-63 is a little more clear.

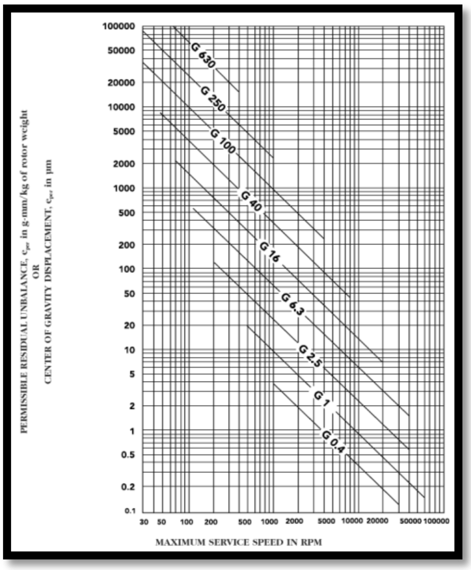


Figure 15-63 Chart from Dennis H Shreve, Commtest, "Balance Quality of Rigid Rotors"

Figure 15-64 shows the **imperial** version of the chart.

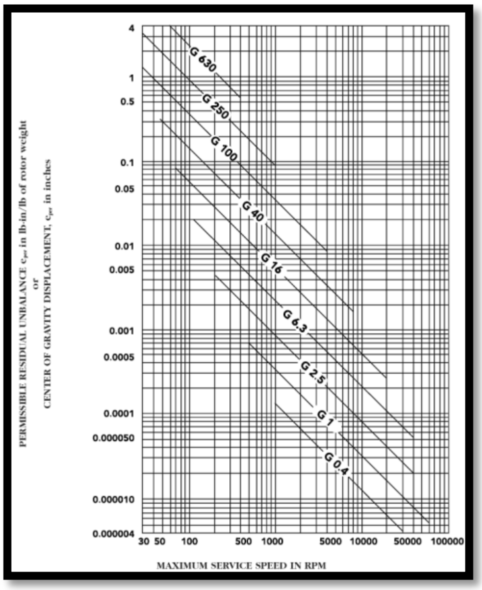


Figure 15-64 Chart from Dennis H Shreve, Commtest, "Balance Quality of Rigid Rotors"

You can also calculate the permissible residual unbalance. The ISO and ANSI standard provides the detail, but you can simply enter the balance grade G (1.0, 2.5, 6.3) to compute the permissible unbalance U_{per}:

Imperial:

$$U_{per} = 6 \times G \times W / N \qquad \text{oz-in}$$

Unbalance [U_{per}]: **oz-in** | Rotor weight [W]: lbs | Speed [N]: RPM

Combination:

$$U_{per} = 170 \times G \times W / N \qquad \text{gr-in}$$

Unbalance [U_{per}]: **gr-in** | Rotor weight [W]: lbs | Speed [N]: RPM

Metric:

$$U_{per} = 9549 \times G \times W / N \qquad \text{gr-mm}$$

Unbalance [U_{per}]: **gr-mm** | Rotor weight [W]: kgs | Speed [N]: RPM

Allocating U_{per}[4]

Please note that the permissible unbalance must be allocated to each balance plane. For a single plane balance, then the permissible unbalance you read from the chart, or calculate, is applied to the single plane. If the rotor was center hung, with an equal amount of weight on each bearing, then the unbalance allowance would be halved for each balance plane. However, if that is not the case then the unbalance must be allocated to each plane accordingly.

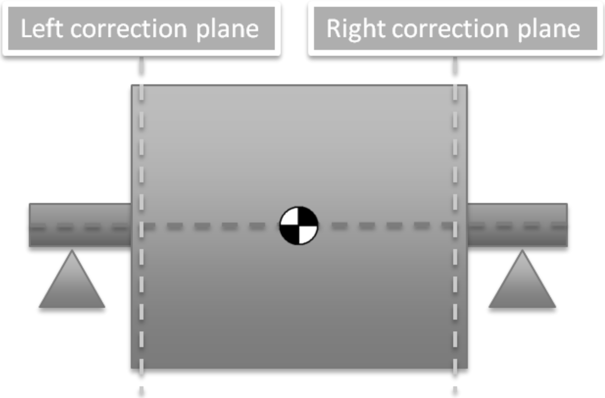


Figure 15-65

Just to make sure this is clear, let's just go through those points again. If we have a single plane balance (we have measured vibration at one point and there is one correction plane), then the U_{per} will be applied to that plane. Based on the RPM, and the G balance grade (e.g. 2.5) and the rotor weight, you will determine the oz-in, gr-in, or gr-mm value that is permissible.

When you go through the balance process and take a final reading the balance calculations will offer a trim balance weight that can be added, for example, 0.5 grams. If the radius at which you must place that weight is 10 inches, then the residual unbalance is 5 gr-in. If that value is lower than the value calculated (or extracted from the chart), then you can pack up your gear because the machine is balanced.

And to continue with the explanation, if you performed a two plane balance, and were looking at the trim balance recommendations for each plane (i.e. a mass at plane one and a different mass at plane two) then you can determine the residual unbalance U_{res} at each correction plane if you know the radius where the weights will be added.

If you know the mass of the rotor, and the speed (maximum continuous service speed), and the G balance grade, then you can determine the permissible unbalance U_{per} for the rotor.

[4] *Equations and diagrams from Dennis H Shreve, Commtest, "Balance Quality of Rigid Rotors"*

However, the U_{per} *must be halved* (the weight is spread evenly between the two bearings), and the residual unbalance U_{res} calculated for each plane must be compared to the revised U_{per} target.

If the weight of the rotor is not evenly distributed between the two bearings, then the permissible unbalance U_{per} would be adjusted in proportion to the weight ratio. If one bearing took 40% of the weight and the other bearing took 60% of the weight, then you would take the U_{per} value and adjust it accordingly (0.4 x U_{per} assigned to one bearing and 0.6 x U_{per} assigned to the other bearing).

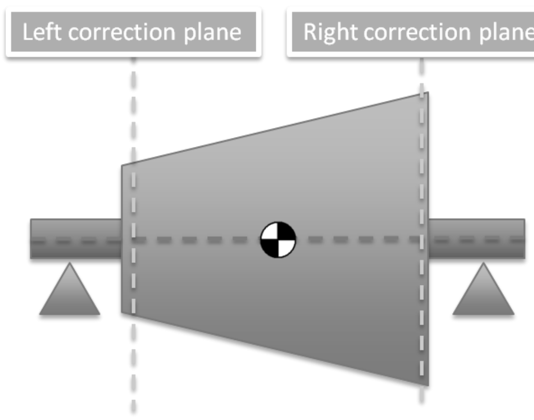
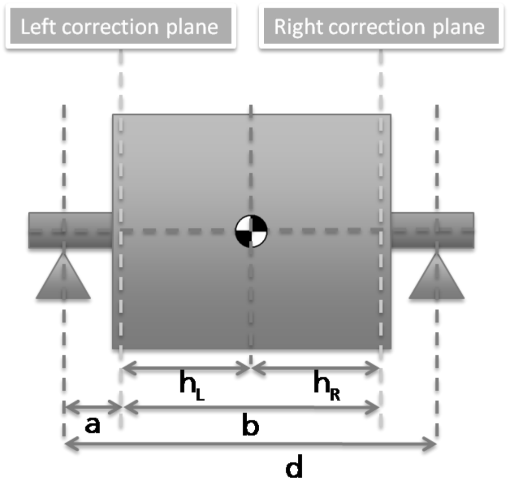


Figure 15-66

OK, let's look at a few rotor shapes and see how the permissible unbalance should be allocated. The diagrams and a number of the equations came from a paper by Dennis H Shreve of Commtest (previously of IRD), "*Balance Quality of Rigid Rotors*", and from ISO 1940-1.

Symmetrical

If the rotor is symmetrical, and the correction planes are within the bearings, and distance "b" is greater than 1/3 "d", and the correction planes are equidistant from the center of gravity, then the unbalance U_{per} can be split between the two balance planes.

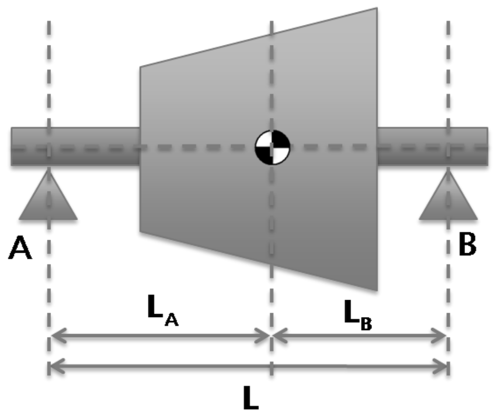


$$U_{per\ left} = U_{per\ right} = \frac{U_{per}}{2}$$

Figure 15-67

Non-symmetrical

From ISO 1940 we are given the following equations for a center hung rotor. The challenge is to know where the center of mass (gravity) is located. As described, the permissible unbalance U_{per} is then distributed between the two balance planes.

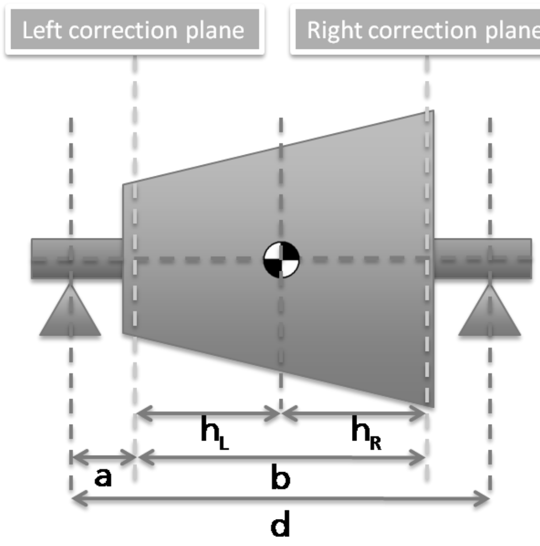


$$U_{per\ A} = \frac{L_B \times U_{per}}{L}$$

$$U_{per\ B} = \frac{L_A \times U_{per}}{L}$$

Figure 15-68

When the correction planes are not equidistant from the center of gravity then the permissible unbalance U_{per} must be allocated to the two balance planes in proportion to the distance.



$$U_{per\ left} = \frac{h_R \times U_{per}}{d}$$

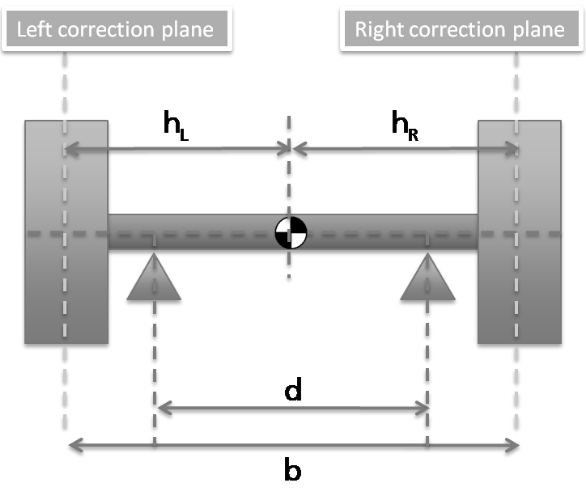
$$U_{per\ right} = \frac{h_L \times U_{per}}{d}$$

Figure 15-69

Having said that, there are limits in the standard according to how much can be allocated to each correction plane. With a center hung rotor, no *more* than 70% can be allocated to any one plane, and therefore, no less than 30% can be allocated to the other plane.

"Dumb bell"

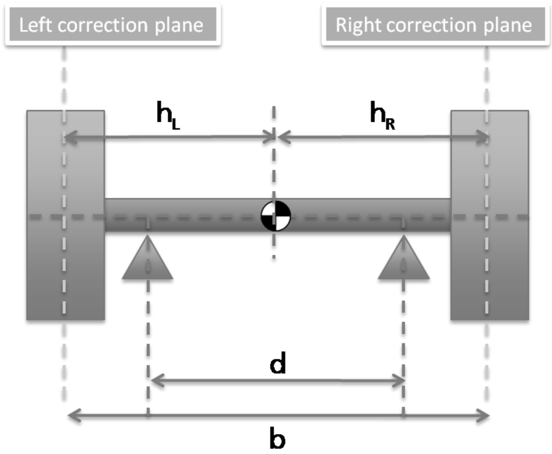
Special rules exist when the correction planes are located outside the bearings. If the correction planes are equidistant from the center of gravity the permissible unbalance is split between the two planes.



$$U_{per\ left} = U_{per\ right} = U_{per} \times \frac{d}{b}$$

Figure 15-70

When the correction planes are not equidistant from the center of gravity, the allocation between the left and right planes are as follows:



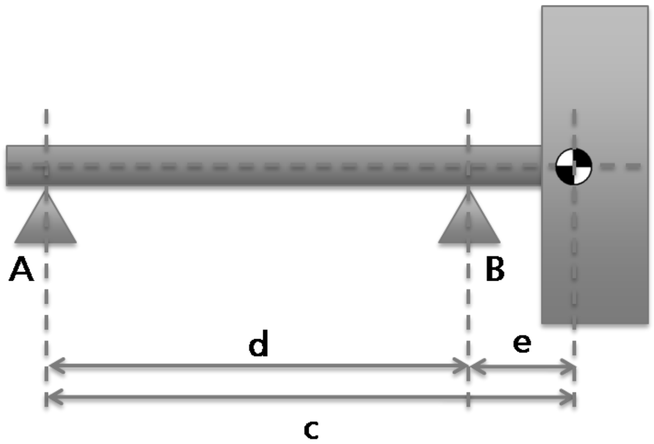
$$U_{per\ left} = \frac{h_R \times U_{per}}{b}$$

$$U_{per\ right} = \frac{h_L \times U_{per}}{b}$$

Figure 15-71

Overhung

From ISO 1940, if the location of the center of mass is known for an overhung rotor, then the permissible unbalance for each of the bearings can be allocated as follows.



$$U_{per\ A} = \frac{e \times U_{per}}{c}$$

$$U_{per\ B} = \frac{d \times U_{per}}{c}$$

Figure 15-72

And, per the standard, for overhung rotors, no *more* than 130% can be allocated to any one plane, and no less than 30% can be allocated to the other plane.

Alternatively, you may define two balance planes on the overhung rotor, and optionally a third static plane (if the left plane is not to be used).

If the distance between the balance planes "d" is less than 1/3 the distance between bearings, i.e. b <1/3 d, and the bearings are equally able to carry the dynamic load, then the permissible unbalance can be allocated between the static and couple planes.

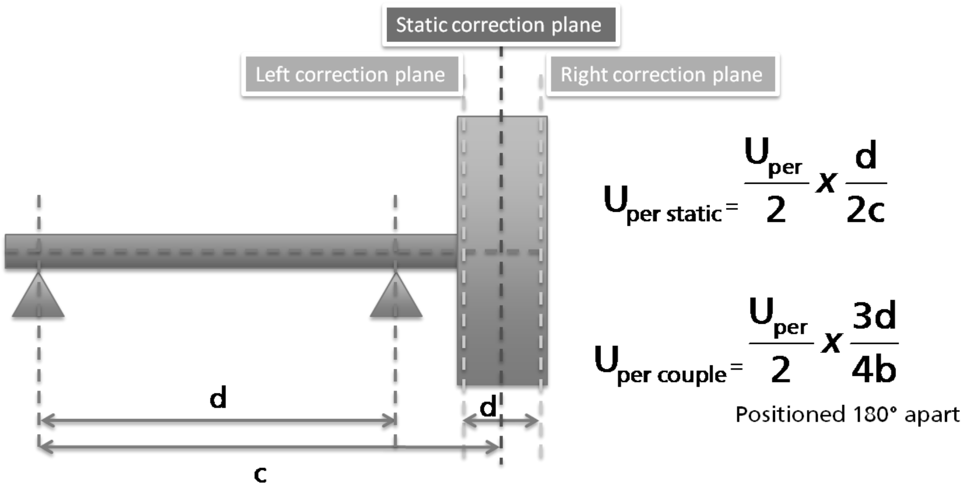


Figure 15-73

The couple corrections are placed 180° apart in their respective planes. The static correction may be made in a third plane.

To clarify the concept of breaking the solution into the static and couple components, please refer to the section on "Static-couple balancing".

Example

Let's go through an example (from the standard).

Rotor mass [m] = 3600 kg

Speed [N] = 3000 r/min

L_A = 1500 mm

L_B = 900 mm

L= 2400 mm

Balance grade [G] = 2.5 mm/s

We can perform the calculation:

U_{per} = 9549 x 2.5 x 3600 / 3000 = 28647 g-mm

$$\boxed{U_{per} = 9549 \text{ x } G \text{ x } W / N \quad \text{gr-mm}}$$

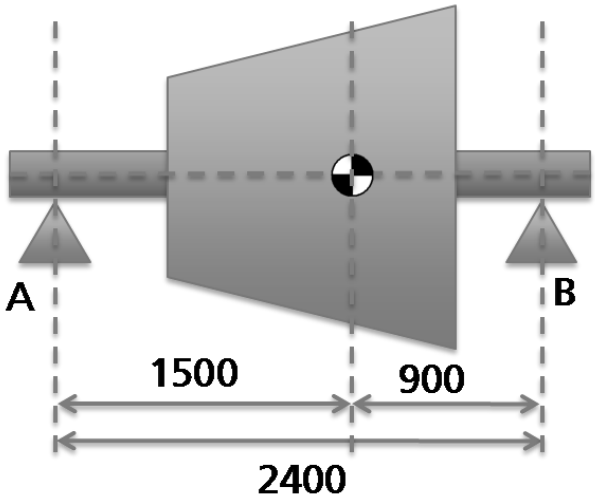


Figure 15-74

$U_{per\ A}$ = 28647 x 900 / 2400 = 10742 gr-mm

$U_{per\ B}$ = 28647 x 1500 / 2400 = 17904 gr-mm

Check that $U_{per\ A}$ and $U_{per\ B}$ are within limits: 0.7 x U_{per} (20052 gr-mm) and 0.3 x $U_{per\ A}$ (8594 gr-mm) – *it is OK*.

Now it is possible to compare these values to the results from the two-plane balance performed on the machine. If the residual balance computed for the A and B planes are less than 10742 gr-mm and 17904 gr-mm respectively, then the balance is within tolerance.

Chapter 16

Shaft Alignment

Objectives:

- Understand the importance of precision shaft alignment
- Understand the limitations of dial-indicator alignment
- Understand the importance of pre-alignment checks and soft foot corrections
- Understand reverse-dial and rim-face alignment
- Learn how to move the machine
- Understand thermal growth

Introduction

This chapter of the training course will help you to understand misalignment - understand why it is important to align your machines, and understand how to align your machines.

If you can achieve these goals, and you can put it into practice, your machines will run more smoothly, and your plant will operate more profitably.

There is a lot to learn, but we will take it one small step at a time.

Why is misalignment so important?

We are all under great pressure to increase uptime, reduce costs, and improve product quality. No matter what your role in the organization we can all contribute to these goals. One way you can help is to perform precision shaft alignment on your rotating machinery.

The fact is that misalignment is the root cause of the majority of machine breakdowns: bearing failures, and damaged seals, shafts and couplings. It is widely believed that 50% of machine failures are due to misalignment.

Figure 16-1 The majority of bearing failures are caused by misalignment

Poor lubrication practices, imbalance, resonances and other factors also contribute to machinery failure, but it is widely agreed that misalignment is the major cause.

Bearing damage

Rolling element (anti-friction) bearings are precision components designed to operate with clean lubricant, reasonably constant temperature, and axial and radial forces/loads within design guidelines. When a machine is misaligned, the pre-load and dynamic forces are raised considerably.

Figure 16-2 Misalignment places considerable load on bearings

If you increase the load on a bearing by just 20%, its life is halved. If you double the load on a bearing, you reduce the life to one seventh of its design life. When machines are misaligned, the load on bearings is increased considerably.

When a bearing fails, production may stop and there can be secondary damage, i.e. damage to other parts of the machine due to the bearing failure. The downtime, parts, and labor are very expensive - and avoidable.

Seal damage

Seals are high cost items, often costing up to a third of the total pump cost. Misalignment also causes seals to fail prematurely due to the increased load. Seals do not tolerate misalignment: face rubbing, elevated temperatures, and ingress of contaminants quickly damage expensive components. The life of the seal can be reduced to 30-50% of design life.

Figure 16-3 Seals will fail prematurely due to misalignment

The result is lubricant leakage and other lubrication problems, and in many cases, total seal failure with little or no warning. When the seal fails, production may stop. The seal and bearing will have to be replaced. The total cost of parts, labor and downtime can make this a very expensive failure.

Coupling damage

The impact of misalignment on couplings varies greatly according to the type of coupling used. The forces and friction are detrimental to all couplings, however the signs of damage, and the nature of the damage can vary.

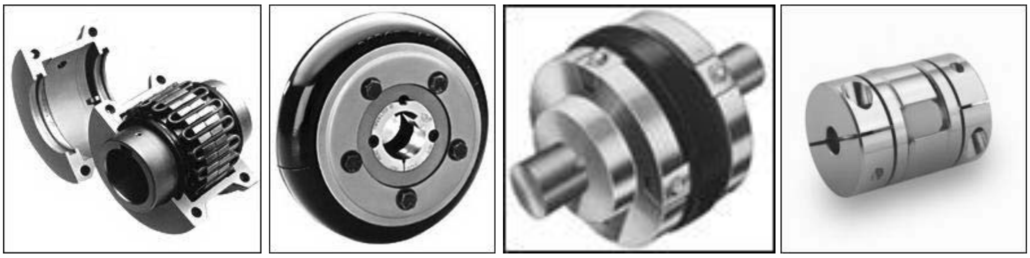


Figure 16-4 Misalignment has different effects on different types of couplings

The rubber or plastic sleeve between the two hubs in flexible couplings can wear - in fact you will often find a small pile of rubber or plastic under the coupling when you remove the coupling guard.

There is a common misconception: "*I use flexible couplings so I do not need to worry about precision alignment*". It is very important to understand that the life of the coupling will be reduced if misaligned, and more importantly, bearings, seals and shafts will still be under increased load and thus likely to fail prematurely.

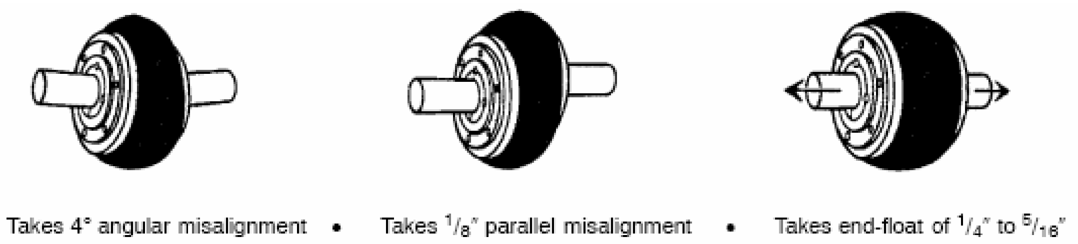


Figure 16-5 Misalignment will reduce the life of flexible couplings

In gear couplings, misalignment results in increased wear on the mating teeth. Under severe misalignment conditions, the load on the teeth will be concentrated to the end of the gear tooth flank. Misalignment can also cause lubrication problems, resulting in metal to metal contact and therefore greatly increased wear.

Figure 16-6 Misalignment can increase the wear on gear couplings

If you eliminate misalignment, the machine will provide greater service. Bearings, shafts, seals and couplings will last longer. Unexpected breakdowns cause secondary damage to machines, and the downtime and repair can cost a small fortune.

Vibration

The rotational forces that result from misalignment generate vibration. We have seen how these forces can damage the coupling, seals, shaft and bearing, but the vibration can damage other components - even machines located within close proximity.

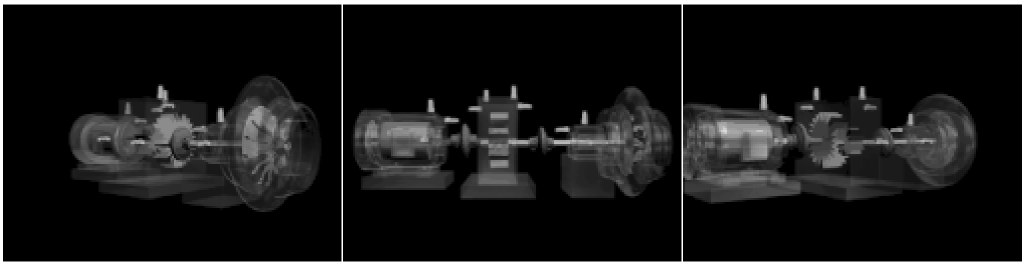


Figure 16-7 Misalignment generates vibration

Have you ever heard of a standby machine that was started, only to quickly fail? This can be due to "brinelling". The bearings in the standby machine are subjected to the vibration of local machines, and that vibration either creates localized wear, or it results in "plastic deformation" of the bearing surface.

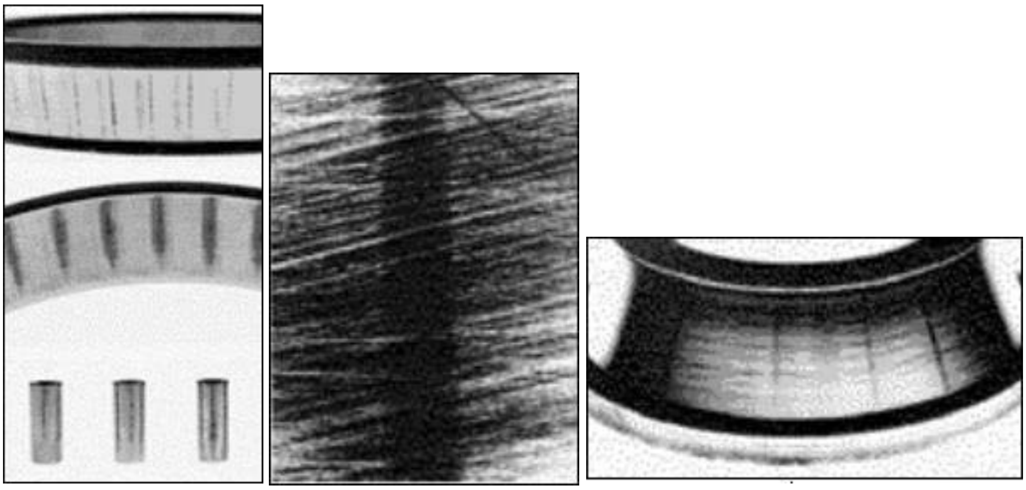


Figure 16-8 Brinelling – bearings are subject to vibration from surrounding machines

Energy consumption

Misaligned machines can also consume more energy. Although studies have produced varying results, it is generally considered that a misaligned machine will consume up to 15% more energy. Large studies have documented savings of between 3% and 8%.

Consider the following equations. If we can reduce the consumption of a 30 HP (22 kW) 460 Volt motor from 36 amps to 32 amps, and we are paying 0.06 per kW Hour then the savings are approximately $1,500.

kW = (Volts x Amps x pf x 1.732)/1000

Annual savings = 8400 x kW x kWH cost

Example kW = (460 x 4 x 0.92 x 1.732)/1000 = 2.931

Example savings = 8400 x 2.931 x 0.06 = $1477

Product quality

Misalignment can also result in reduced product quality in many industries. By reducing vibration levels, and aligning rolls and other items, product quality can be maximized.

Downtime and production capacity

The biggest issue of all is production capacity. If you can increase the reliability of rotating machinery, downtime will be reduced. If you increase the uptime, your plant can increase production which has a major impact on the bottom line.

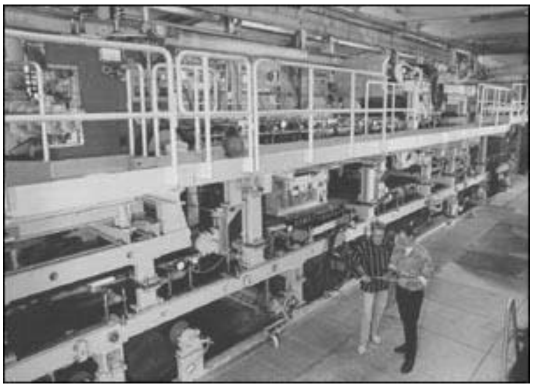


Figure 16-9 Reliability = increased uptime

It is therefore in everyone's best interest to correctly align machines. The extra time and effort required will increase profitability - and job security.

Detecting misalignment

There are two ways to determine if a machine is misaligned: you can look at the maintenance records and alignment records to see if the machine was misaligned, and you can monitor various physical parameters such as vibration and temperature.

First, if a machine has not been "precision" aligned, that is, it was aligned "by eye" or only using a straightedge, then it is very likely that it will be misaligned.

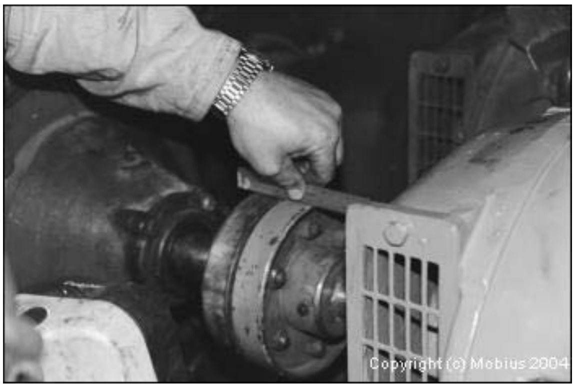


Figure 16-10 Machines will likely be misaligned if precision alignment was not performed

Later we will discuss the alignment procedures, and we will discuss tolerances (which tell us how good the alignment must be to achieve maximum life), but for now it is safe to say that if you are not correctly using dial indicators or a laser alignment system, the machine will not be within tolerance.

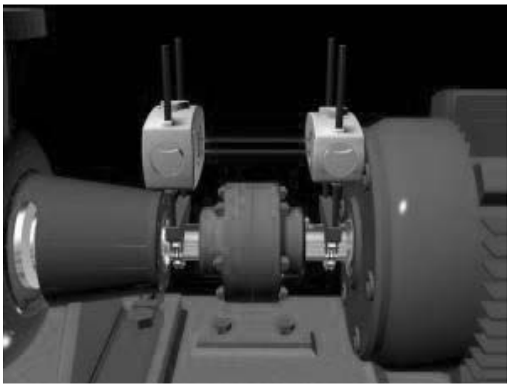
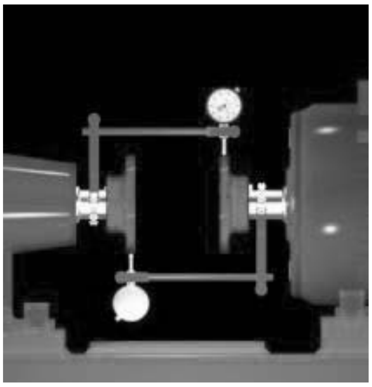



Figure 16-11 Laser alignment (left) and dial indicator (right) systems

We have just described all of the reasons why you should perform shaft alignment: damaged bearings, shafts, seals, and couplings. If you have experienced any of these problems then it is probably due to misalignment.

Other physical signs include loose or broken hold-down bolts, loose shim packs or dowel pins, excessive oil leaks at bearing seals, and loose or broken coupling bolts.

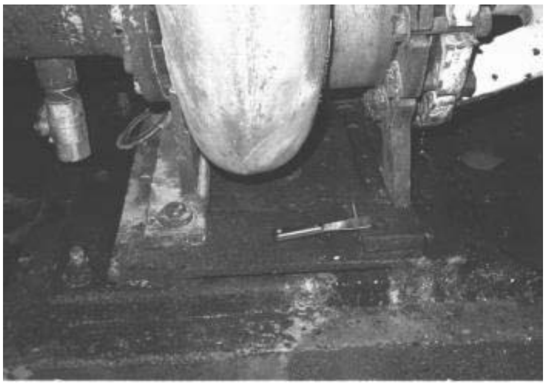


Figure 16-12 Misalignment can be detected via physical signs

You should always perform "root cause analysis" when a machine fails. Root cause analysis is used to determine why a machine failed. It is not good enough to simply repair it and put it back in service - what is to stop the same fault from developing again? Issues like lubrication, imbalance, resonance and misalignment should be considered when a machine fails. And thus when a machine is put back into service, it should be *precision* balanced and *precision* aligned, and the lubrication program should be maintained.

Detecting misalignment

Due to the nature of misalignment, there are at least two physical signs that shafts are not aligned: the machine will vibrate in characteristic ways, and the coupling may get hot which can be detected with non-contact temperature guns or thermal imaging cameras.

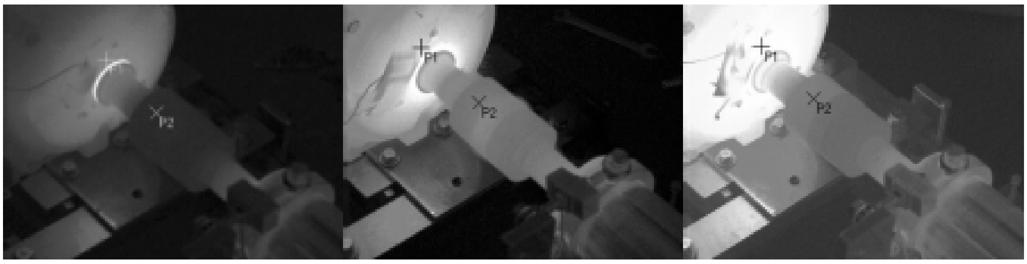


Figure 16-13 Thermal imaging showing heat buildup due to misalignment

Using vibration analysis to detect misalignment

Vibration analysis has been successfully used for many years to detect misalignment. Depending upon the nature and severity of the misalignment (and the nature of the coupling and size/speed of the machine), vibration measurements can help us detect misalignment and assess the severity.

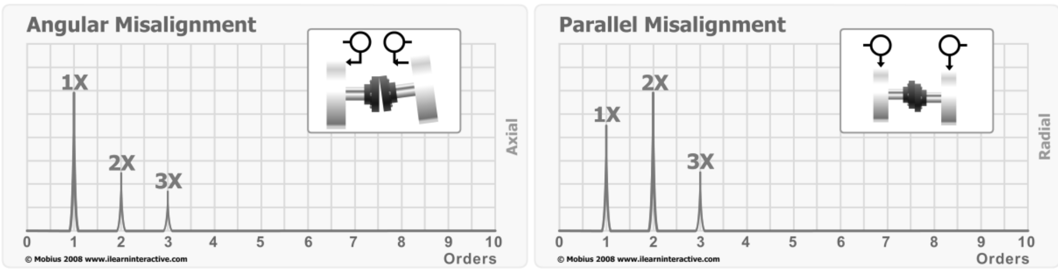


Figure 16-14 Vibration measurements help us detect misalignment

The specific methods used to detect misalignment are covered in the diagnostic section of Category II course, however there are a few quick comments to make:

- The combination of vibration spectra, time waveforms, **and phase readings** give the best indication of misalignment.
- Measurements should be taken in the axial and radial directions. One direction alone is insufficient.
- Simple measurements like rms "overall" readings cannot provide definitive evidence of misalignment.
- High frequency readings (like shock pulse, spike energy, HFD, Peak Vue) cannot be used to detect misalignment.
- Vibration analysis cannot always be used to detect misalignment. Sometimes the vibration levels and patterns do not change when a machine is misaligned.

What is misalignment?

We should take a few minutes to understand what misalignment means so that you can better understand why it is so destructive. A definition:

"Shafts are misaligned when their rotational centerlines are not collinear when the machines are operating under normal conditions."

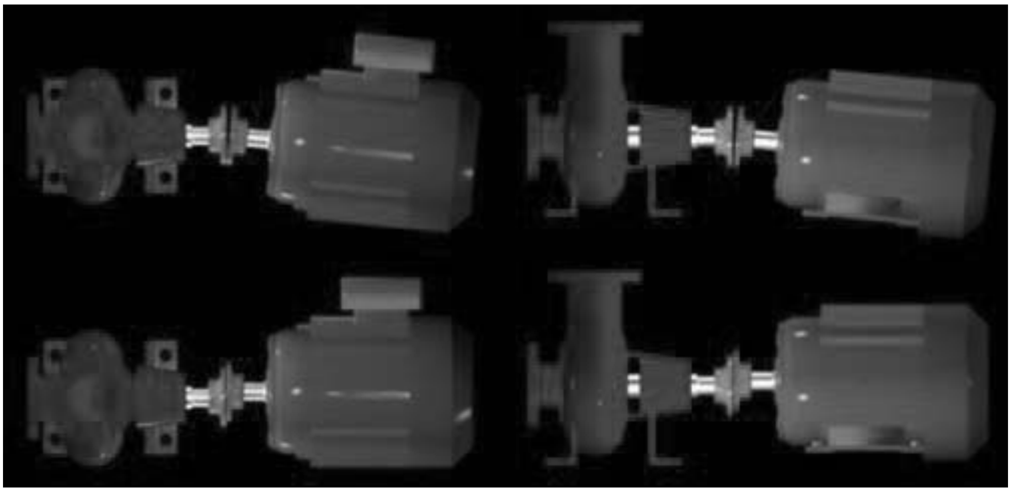


Figure 16-15 Different forms of misalignment

When two machine components, for example a motor and a pump, are assembled and coupled together, there is likely to be some misalignment - in fact, it is guaranteed. The pump will have been attached to its piping, and the motor connected to its conduit, and both will have been bolted to their baseplate.

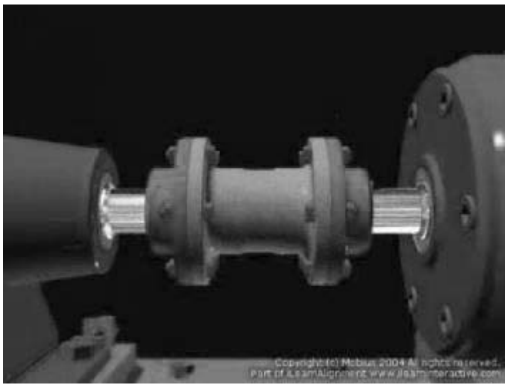


Figure 16-16 Motor and pump coupling

We hope there is no stress involved with the connection to piping and conduit - but there will be. We hope the feet of the machines and the baseplate are true and make perfect, flat contact, but they probably won't. And we hope that the two machines are positioned at the correct height and in a straight line, but that is also highly unlikely.

Sadly, all of these factors, and others, contribute to the fact that unless we take special precautions, machines will not be aligned correctly, and thus we will not get the desired life or performance out of the machine.

A closer look at misalignment

First, let's look more closely at the term misalignment. Our aim is to have the "rotational centerlines" of both shafts in line when the machine is operating. But what is the rotational centerline?

If the pump was uncoupled and you turned its shaft, it would rotate around a straight line. This is the pump's rotational centerline. When it is uncoupled, it is easy to turn the shaft (well, as easy as it will ever be). The same is true for the motor - its shaft also has a rotational centerline. Ideally, the motor and pump (Figure 16-17) will be mounted so that the two rotational centerlines are aligned perfectly - there will be no offset, and no angle between the shafts in either the vertical or horizontal direction. The coupling will bolt together easily, and there will be minimal stress on the bearings, seals or any other components when the shaft is turned. In this case the shafts are said to be "collinear".

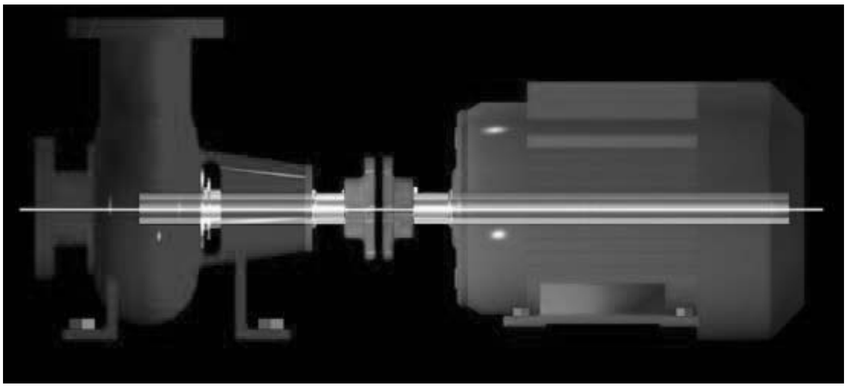


Figure 16-17 Collinear shafts – aligned rotational centerlines

But instead what happens is that the two components do not come together perfectly - perhaps one is a little higher than the other, and a little to one side. It might look okay to the naked eye, so the components are coupled together and the machine is run that way. It "looks" OK, and the machines runs, so what's the problem?

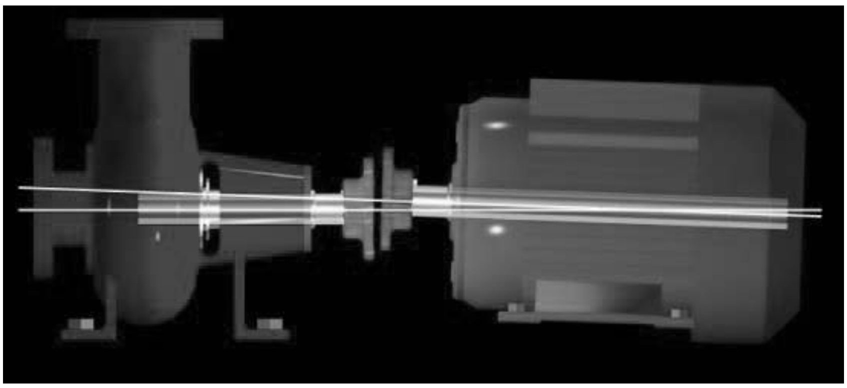


Figure 16-18 Rotational centerlines misaligned

The different forms of misalignment are shown in Figure 16-19.

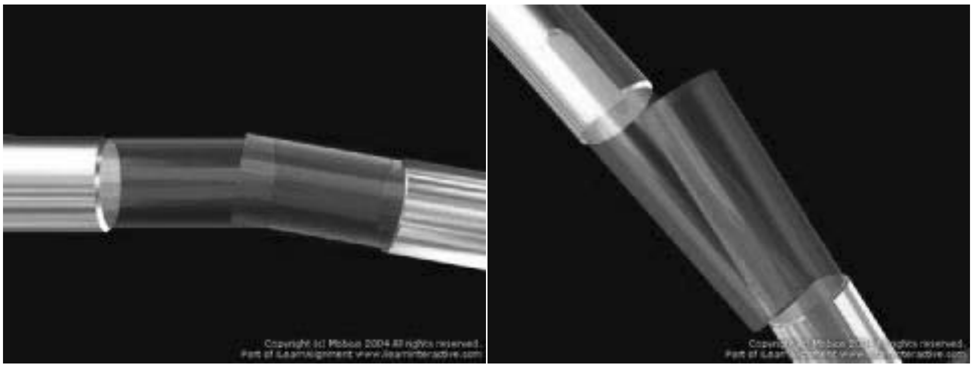


Figure 16-19 Vertical (left) and horizontal (right) offset and angular misalignment

You will see an offset and angle introduced in the vertical direction, and then you will see a different offset and angle introduced in the horizontal direction. The result is two shafts that are misaligned. In actual fact, when the components are coupled together, the coupling will be forced to "give" in order to accommodate the misalignment, and the two shafts may be forced to bend slightly (Figure 16-20).

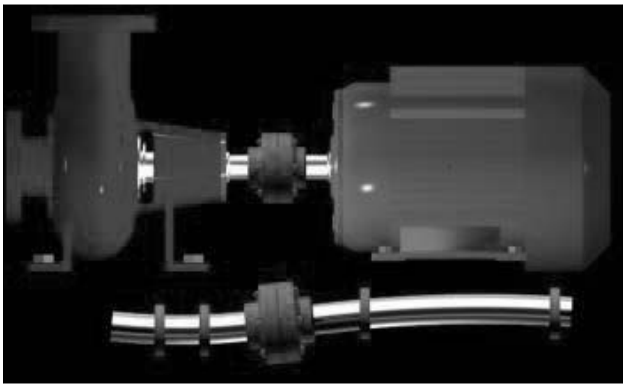


Figure 16-20 Bend in shafts due to misalignment (exaggerated)

Think about what is happening inside this machine - what is happening to the shaft, and the bearings, and the seals, and the coupling? With every rotation the shaft has to flex - it is forced to stay within the bearings. That puts strong radial and axial forces on the coupling, bearings, seals and shaft.

Offset and angular misalignment

Let's have a look at the different types of misalignment.

If one component is higher than the other, or to one side, there will be an offset. This is called "offset misalignment", or sometimes "parallel misalignment".

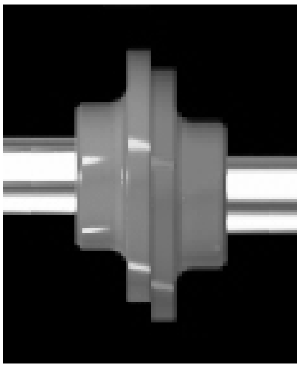


Figure 16-21 Offset or parallel misalignment

It is called "angular misalignment" (or "gap misalignment") when the two shaft centerlines meet at an angle.

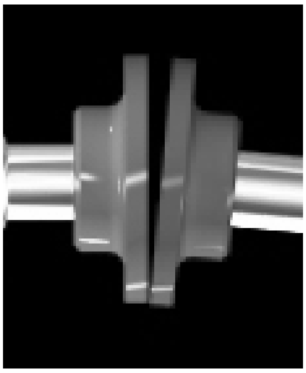


Figure 16-22 Angular or gap misalignment

In reality, shaft misalignment is a combination of both offset and angular - it will almost never be just angular or offset (parallel) misalignment.

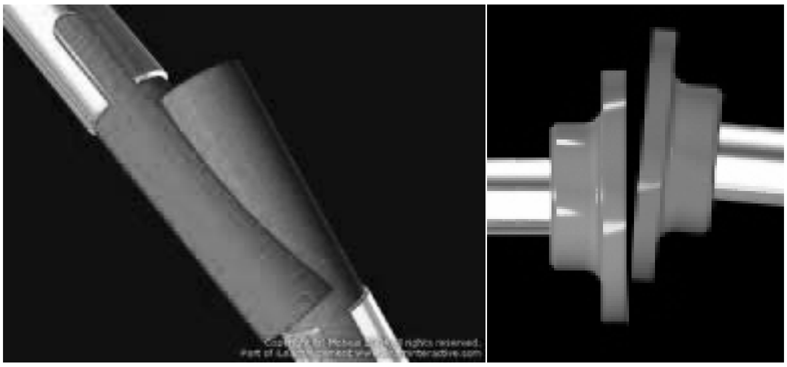


Figure 16-23 Offset and angular misalignment combined

It must be noted that the offset and angle exist both vertically and horizontally, to different degrees. For example, the motor shaft may be offset a little higher but very much to the left of the other shaft, and it may make a large angle vertically, but a small angle horizontally.

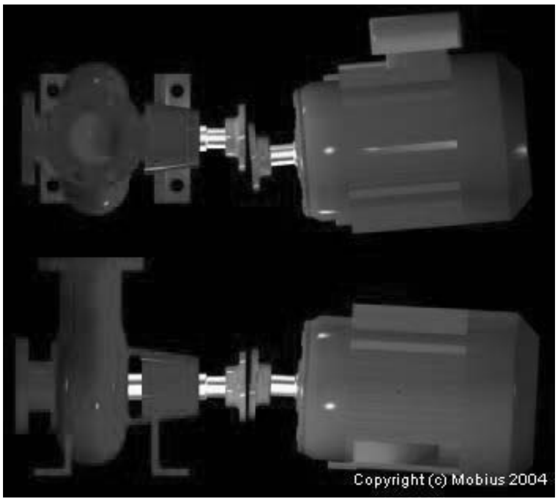


Figure 16-24 Vertical offset and angle (plan view, top) and horizontal offset and angle (side view, bottom)

Visualizing tolerance

You can think of the tolerances as a cone - the motor's shaft centerline can exist anywhere within the cone (Figure 16-25). The cone does not come to a point because we allow a certain amount of offset.

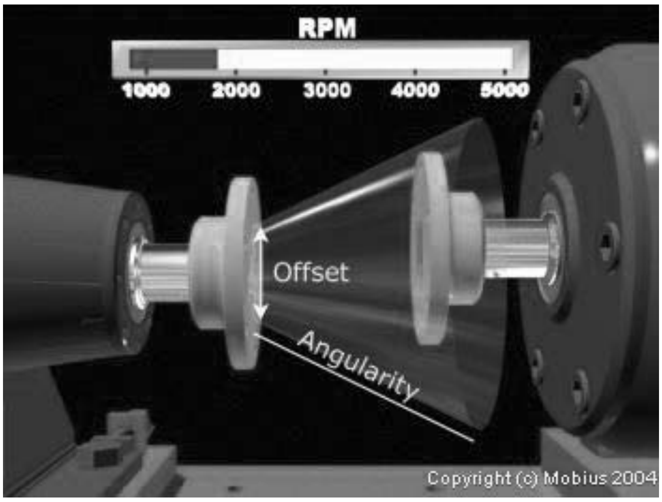


Figure 16-25 Tolerance 'cone' (exaggerated)

Tolerances and speed

While there are a number of factors to consider, the key issue when it comes to the relationship between the misalignment and the damage that can be done to the machine is the machine speed. If the speed is greater, the damage will be greater.

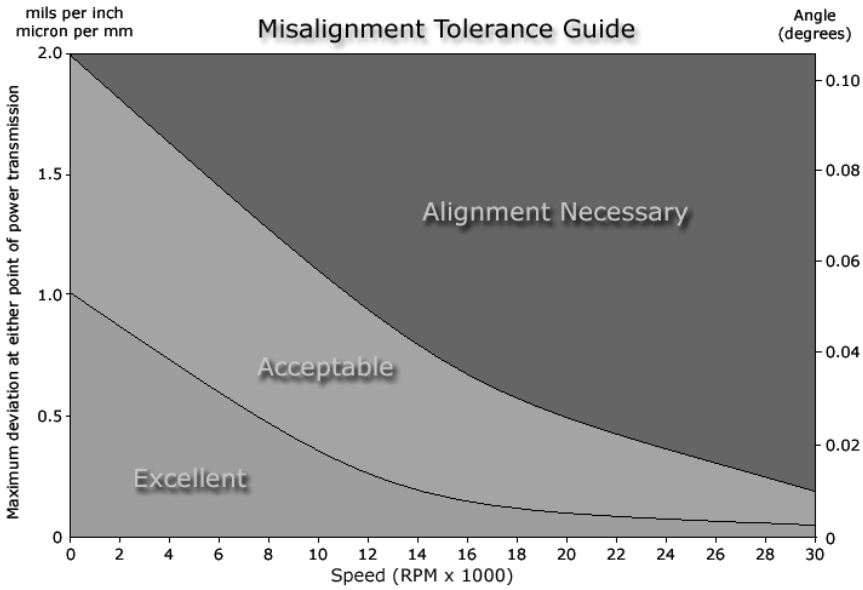


Figure 16-26 Misalignment tolerance guide (From: Shaft Alignment Handbook by John Piotrowski)

For this reason the tolerances need to be tighter on higher speed machines - that is, we allow a smaller offset and smaller angle between the shafts for higher speed machines.

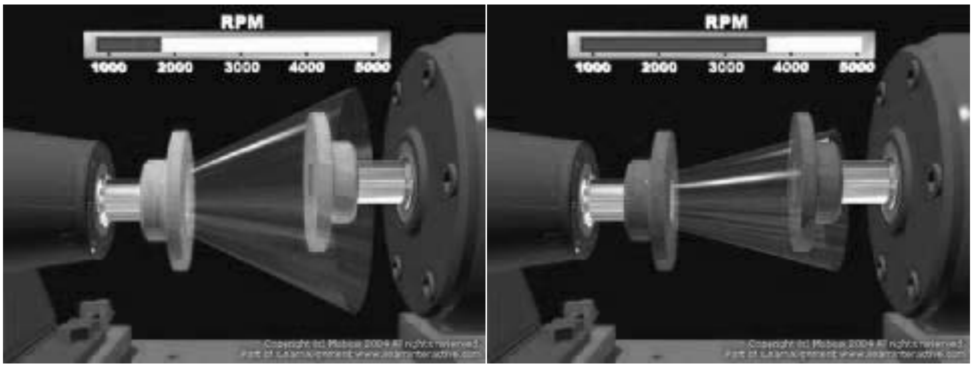


Figure 16-27 Smaller tolerance 'cone' on higher speed machine (right)

In Figure 16-27 we show that we can start with a certain amount of misalignment which is acceptable (within tolerance) for a lower speed machine (1800 RPM), but if the machine was running at 3600 RPM, the allowable offset and angularity is reduced, thus the machine is no longer in tolerance (it turns red).

You can see that the machine has the same amount of misalignment, but because it is running at a higher speed it is now out of tolerance.

Published tolerances

Tolerances are available from a number of sources. The tolerances from the PRUFTECHNIK company have been included here for your reference. You can see that the allowable offset and angularity is reduced for higher speed machines. You can also see that the tolerances have been set as "acceptable" and "excellent". You should always aim for "excellent" tolerance - your machine (and company balance sheet) will thank you for it.

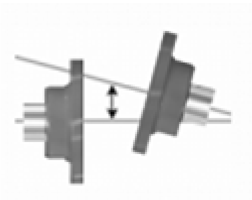
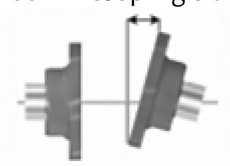
	RPM	inch (mils)		metric (mm)	
		Acceptable	Excellent	Acceptable	Excellent
Short “flexible” couplings					
	600	9.0	5.0		
Offset:	750			0.19	0.09
	900	6.0	3.0		
	1200	4.0	2.5		
	1500			0.09	0.06
	1800	3.0	2.0		
	3000			0.06	0.03
	3600	1.5	1.0		
	6000			0.03	0.02
	7200	1.0	0.5		
Angularity:					
Inch: Gap difference per 10 inch coupling diameter	600	15.0	10.0		
	750			0.13	0.09
Metric: Gap difference per 100mm coupling diameter	900	10.0	7.0		
	1200	8.0	5.0		
	1500			0.07	0.05
	1800	5.0	3.0		
	3000			0.04	0.03
	3600	3.0	2.0		
	6000			0.03	0.02
	7200	2.0	1.0		
Soft foot	Any	0.06		2	

Table 16-1 Tolerances from PRUFTECHNIK

Here is the tolerance table for spacer shafts.

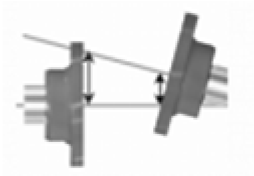
	RPM	inch (mils)		metric (mm)	
		Acceptable	Excellent	Acceptable	Excellent
Spacer shaft and membrane (disc) couplings: **Inch:** Offset per inch spacer shaft **Metric:** Offset per 100mm spacer shaft 	600	3.0	1.8		
	750			0.25	0.15
	900	2.0	1.2		
	1200	1.5	0.9		
	1500			0.12	0.07
	1800	1.0	0.6		
	3000			0.07	0.04
	3600	0.5	0.3		
	6000			0.03	0.02
	7200	0.03	0.02		
Soft foot	Any	0.06		2	

Table 16-2 Spacer shaft tolerances from PRUFTECHNIK

Dynamic movement

When we originally defined misalignment we said:

"Shafts are misaligned when their rotational centerlines are not collinear when the machines are operating under normal conditions."

Did you notice the phrase *"under normal operating conditions"*?

When you take alignment measurements the machine is normally cold, and it is certainly not operating. But when it is started, a number of things happen. Rotational forces, operating pressures, and increased temperature all cause the position of the two rotational centerlines to change! Your machine could go from a state of precision alignment to out-of-tolerance misalignment.

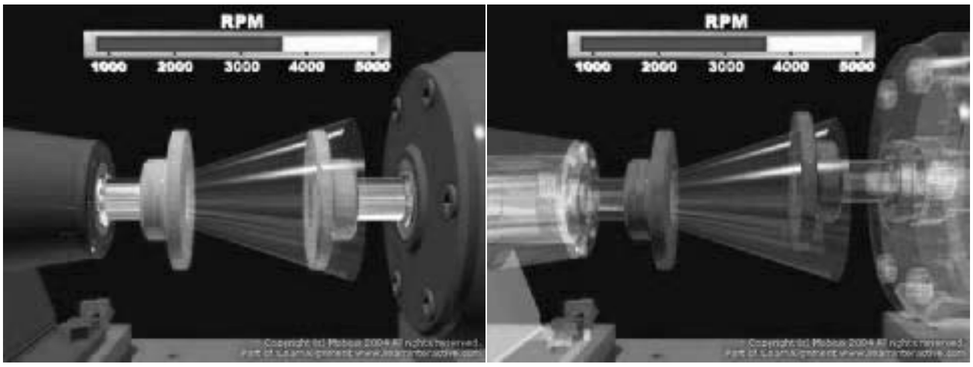


Figure 16-28Misalignment within tolerance when cold (left), but out of alignment when running (right)

In Figure 16-28, the machine has a certain amount of misalignment (exaggerated) which is within tolerance. But as the machine starts running, it heats up and moves out of alignment. All machines will undergo some change - they all increase in temperature, and thus the metal expands. They all experience rotational forces and most undergo changes due to pressure/flow of process fluids/gases. But most of the time we can afford to ignore this effect.

The simplest way to deal with this condition on the physically larger machines that undergo greater temperature changes is to compute the thermal growth and include that information into the alignment targets. Some manufacturers of affected machines will provide this data.

Figure 16-29 Compute thermal growth for larger machines

For example, using calculations that utilize the coefficient of expansion for different types of metals, you may determine that the shaft of the blower may lift 10 mils [0.25 mm], while the shaft of the motor will only lift 5 mils [0.125 mm]. You would therefore align the machine when it is cold so that blower shaft is 5 mils [0.125 mm] below the motor shaft. When it starts the two machines will gradually increase in temperature, and they will move into alignment.

It is also possible to align the machines soon after they have stopped when they are still hot. There are a number of issues to consider, but it is better than doing nothing.

It is also possible to attach laser heads to the bearings of the machine via special brackets so that you can measure exactly how much the two components will move. Measurements are taken when it is cold and again when it is hot. The movement is then factored into the cold alignment targets.

Pre-alignment tasks

One of the most important issues related to shaft alignment is what you do before you measure and correct the alignment. Your preparation of the alignment job is key to your success (and to your safety).

If possible, you should begin your work before the machine comes off-line. You should make sure all of your alignment equipment is ready to go, and that the batteries in the laser alignment equipment (if applicable) are ready for a day's work.

You should also review the maintenance records. Depending upon the nature of the alignment job (i.e. why are you doing an alignment, and what type of machine is being aligned), it is a very good idea to look at:

1. The maintenance records, to see why it should require alignment (unless it is simply the replacement of a motor).
2. Make sure that you have a good selection of shim sizes. They should be clean and straight.
3. Previous notes from the last time you performed an alignment on this machine: did you experience bolt-bound or base-bound problems, were there any other issues?
4. You should determine your alignment targets. Are there any recommendations from the manufacturer? Do you have to compensate for thermal growth?
5. And if possible, you should review the results from the previous alignment job performed on this machine.

Before you approach the machine you must follow all lock-out and tag-out procedures. And you must close all pump valves, etc. to ensure that the shaft cannot begin rotating because of fluid/gas/air flow.

Figure 16-30 Follow all lock-out and tag-out procedures

Collect "as-found" readings

You should record the as-found readings - take a set of alignment readings to show the initial alignment state. This is not necessary if it is a new installation, but it can be very beneficial to be able to show the improvement to the alignment state as a result of your work.

Create a clean work area

You must create a clean work area. Any dirt, grit, burrs, or other debris that gets between the machine and the baseplate, or under/between shims, etc., can cause you all kinds of problems.

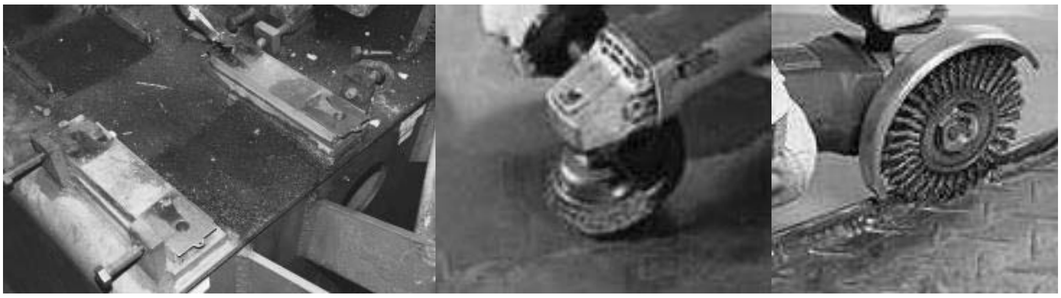


Figure 16-31 Remove dirt, grit, burrs and other debris

Prepare your shims

You should prepare your shims carefully. Remove any shims that are rusted, bent, painted, or dirty (beyond cleaning). Shims can act like small springs under the machine feet which, among other things, can make the alignment process very difficult.

Figure 16-32 Clean shims and remove those that are rusted, bent or painted

Take care of the bolts

You should replace bent, damaged or oversized bolts with new bolts. You should lubricate the bolts, and always use the same torque when tightening the bolts. You should also tighten and loosen the bolts in the same order.

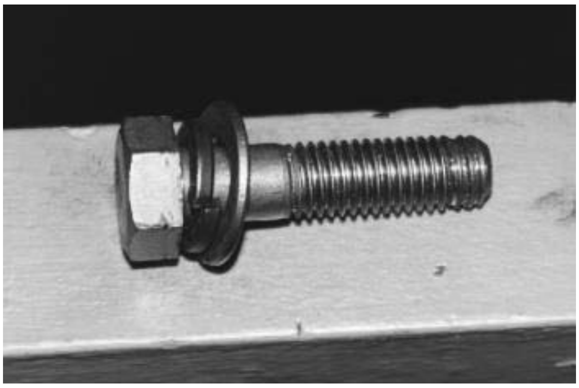


Figure 16-33 Replace bent, damaged or oversized bolts

Prepare the foundations of the machine

You should remove the taper pins and loosen off jacking screws before you begin taking measurements.

You also need to loosen the coupling bolts (in a rigid coupling). If you do not do this, it will be impossible to measure an offset and angle, as the coupling will cause the shaft to bend.

Check the physical condition of the machine

You should check the mechanical "health" of the machine. Rotate the shaft and see if there is any looseness, rubbing or binding. You should check to see if the shaft is bent, and if there is any coupling runout. You should check the coupling for excessive wear and proper fit, and check that the key is the correct length.

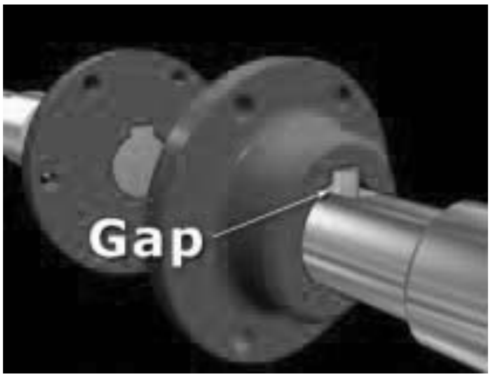


Figure 16-34

You should check for excessive piping strain, conduit strain and other forms of stress placed on the machine. If you loosen the feet and they move more than 0.002" or 0.05 mm then you should make corrections.

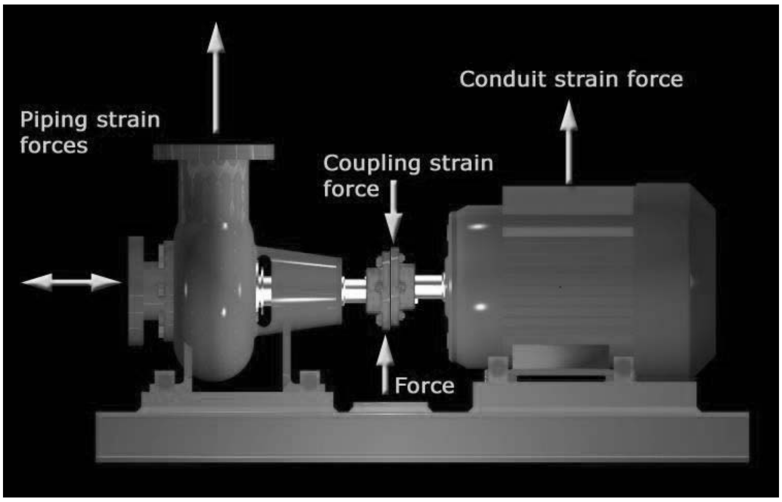


Figure 16-35 Check for excessive stress on the machine

Check and correct soft foot

You should check and correct soft foot. Soft foot is the condition where the feet do not make perfect, flat contact with the baseplate. Some people equate it with a short leg on a chair, rocking back and forth. But there is more to it than that.

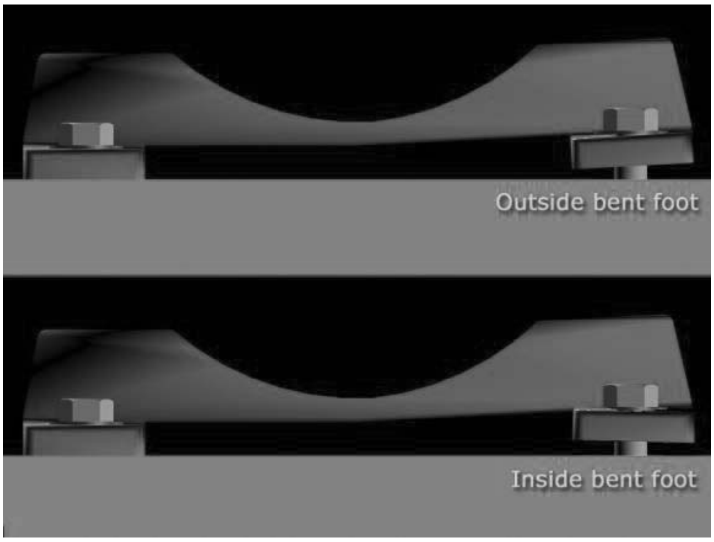


Figure 16-36 Gross soft foot check – obvious gaps under feet

You can start with a "gross soft foot check". Look for any obvious gaps under feet, and shim accordingly. Then you can perform a soft foot check with dial indicators or your laser alignment system.

Figure 16-37 Checking for soft foot using a laser alignment system

This involves loosening each foot, one at a time, and measuring how much the foot lifts. If it lifts more than 0.002" or 0.05 mm, then the soft foot condition must be corrected.

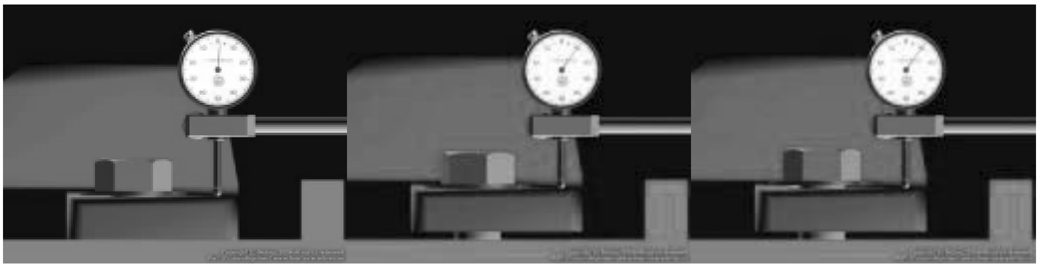


Figure 16-38

Note in Figure 16-38 that as the bolt is loosened the needle indicates that the foot is lifting. But when it reaches 10 the needle stops moving, even though the bolt is still being loosened.

Soft foot can take a number of forms. In summary, if two of the feet diagonally opposite each other have the highest readings then you have rocking soft foot (Figure 16-39). Shims can be placed under those feet. Otherwise the readings may indicate a bent foot condition, a "squishy" foot condition (too many shims or dirty shims under the foot), or the existence of pipe/conduit stress.

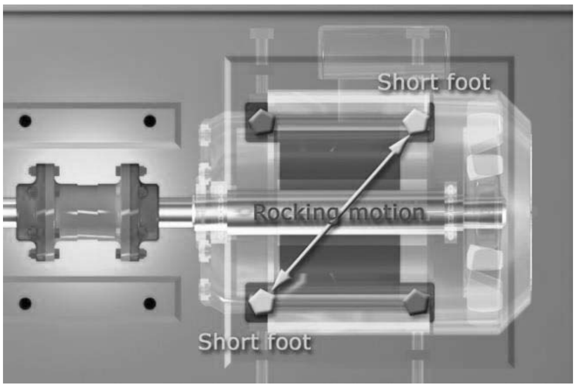


Figure 16-39 Rocking soft foot

Soft foot can make the alignment task very difficult and frustrating, and it can distort the machine frame and bearings - thus reducing the life of the machine. **You must not ignore soft foot.**

Begin the alignment process

Once you have prepared the site, checked the mechanical health of the machine, and corrected the soft foot condition, you are ready to begin the alignment measurements and correction.

Determining the alignment state

We now know what misalignment is, but how do you determine where the rotational centerlines are located so that you can make corrections? If you knew that the shaft of the motor was parallel but higher than the shaft of the pump by 10 mils (0.25 mm), you could lower it by that amount and be done. But how do you measure the offset and angle?

There are three ways to determine the relative positions of the shaft rotational centerlines: by eye, with dial indicators, and with laser systems (there are other methods, but we won't discuss them here).

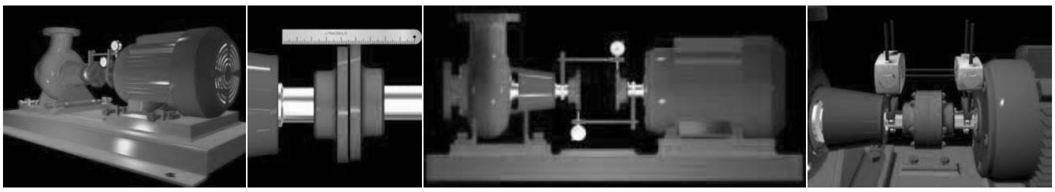


Figure 16-40 Methods of determining shaft rotational centerlines

Using a straightedge or feeler gauge

You can use very rudimentary tools to attempt to determine the relative position of the shafts, but the accuracy is very poor. These methods can be used for an initial "rough alignment", but not for the final alignment.

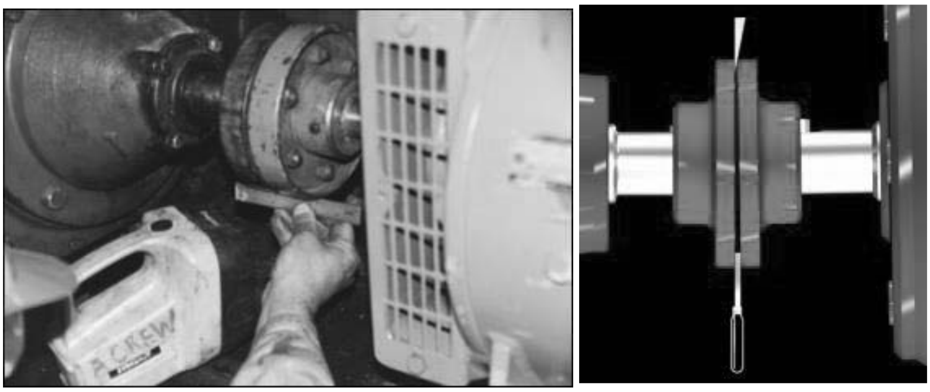


Figure 16-41 “Rough alignment” using rudimentary tools

You can measure the gap in the coupling to determine the angularity. For example, if there is a gap at the bottom of the coupling but not at the top, then you know the motor must slope up and away from the coupling. You can then lower the feet of the motor to close the gap. The shafts will now be parallel.

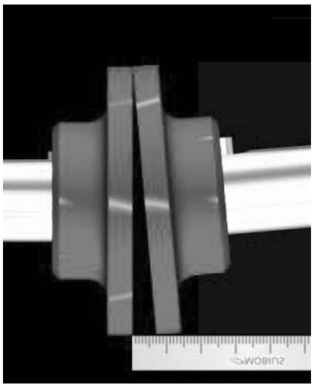


Figure 16-42 Gap at the bottom of the coupling shows the motor is sloping upwards

If you sit the straightedge on the coupling hub you will see a gap between the straightedge and the lower coupling (assuming there is no runout and the couplings are the same diameter). You can measure the gap, or simply move the machine until the gap disappears.

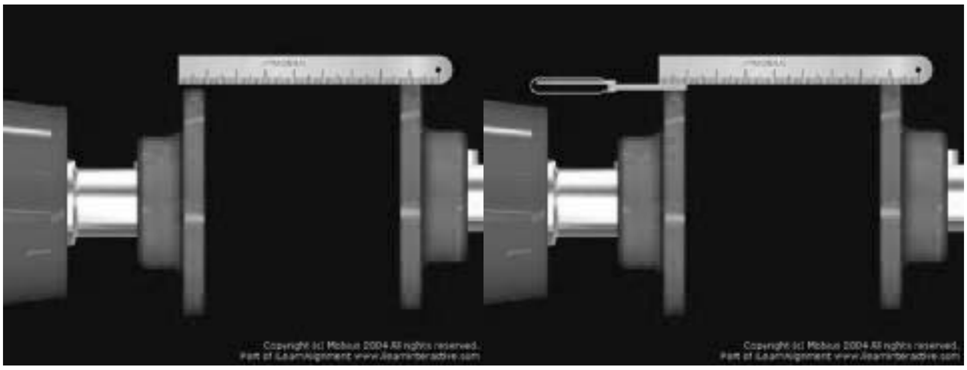


Figure 16-43 Using a straightedge to sight a gap

Again, you will look at the gap on top of the coupling to determine the vertical shimming required, and you will look at the offset at the side of the coupling to determine the feet movements.

It should be noted that we are not actually measuring the position of the shaft "rotational" centerline - we are simply making a rough assessment of the position of the shaft "geometric" centerline. In fact, we are really just aligning the couplings. All of our measurements thus far are taken on the coupling. If the coupling is incorrectly bored, if the shaft is bent, if the coupling is not round, or for a number of other reasons, we will not even be aligning the shaft geometric centerline.

With practice and common sense you can improve the alignment state with a straightedge and feeler gauge, but you will not correct the alignment. It is highly unlikely that the machine will be within tolerance using these methods alone.

Using dial indicators

Dial indicators are commonly used to measure the relative positions of the shaft rotational centerlines. There are a number of configurations that can be used to attach the dials to the shaft, but we will quickly review the two most common methods: the rim and face method and the reverse dial method.

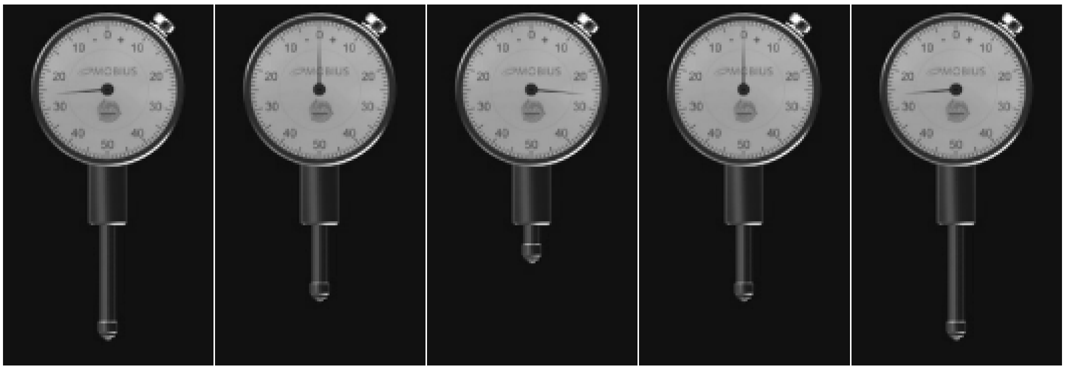


Figure 16-44

Dial indicator limitations

It is important to know that there are a number of limitations with these dial indicator methods that can result in poor alignment results.

Bar sag

One of the biggest issues is "bar sag". When the shaft is rotated from the 12:00 position to the 6:00 position, the bar carrying the rim measurement bends slightly. This adversely affects the rim readings. You <u>must</u> first separately measure the amount the bar will bend, and then compensate the readings.

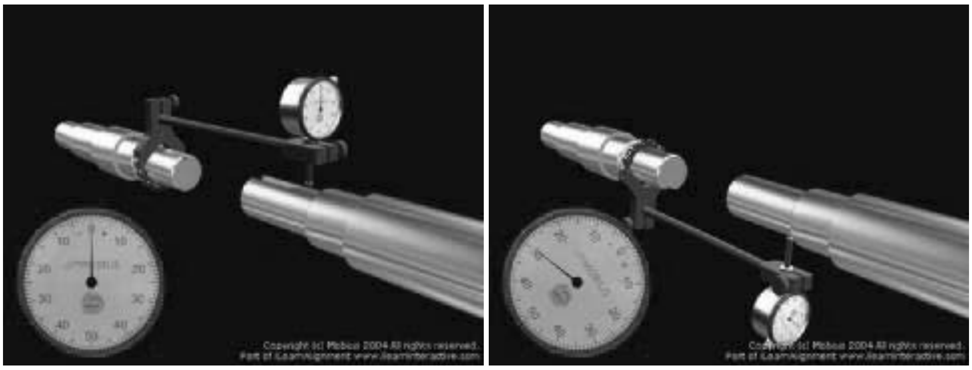


Figure 16-45 Dial indicators showing bar sag at the start of the test and at the end of the test

These two shafts are perfectly aligned yet the dial indicates that there is an offset – this is entirely due to the bar sag.

Reading accuracy

Another issue is reading accuracy. Because you have to take the reading from the dial, you have up to 0.5 mil [0.005 mm] rounding error. It is very difficult to take the reading from the dial, especially in a crowded machine space in poor light when the dial is up-side-down.

Additional problems

There are many other important issues: internal friction/hysteresis, reading errors, play in mechanical linkages, axial shaft play (especially for face readings), and more. And it is also common for people to make mistakes with the calculations and graphical method.

For many years people have been able to do a good job with dial indicators, however laser alignment systems produce more accurate and dependable results (with automated calculations of feet corrections), and the measurements can be taken far more quickly.

The Rim and Face method

The rim and face method requires two measurements to be taken, one on the rim of the coupling (in order to measure the offset), and the other on the face of the coupling (in order to measure the angularity). You can see where this method gets its name!

Commercial systems are available with special brackets that are designed to maximize measurement accuracy. But the basic principle is always the same: one dial measures the face angle, and the other measures offset.

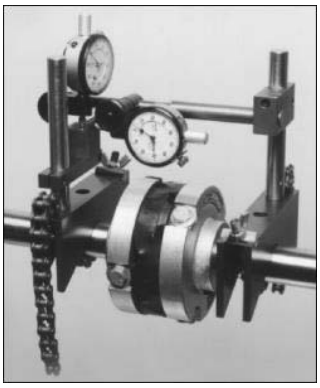


Figure 16-46 Rim and face configuration

Figure 16-47 shows an example of one of the tests in action. Notice that both shafts are rotated. This means that the system is measuring the relative position of the shaft rotational centerlines.

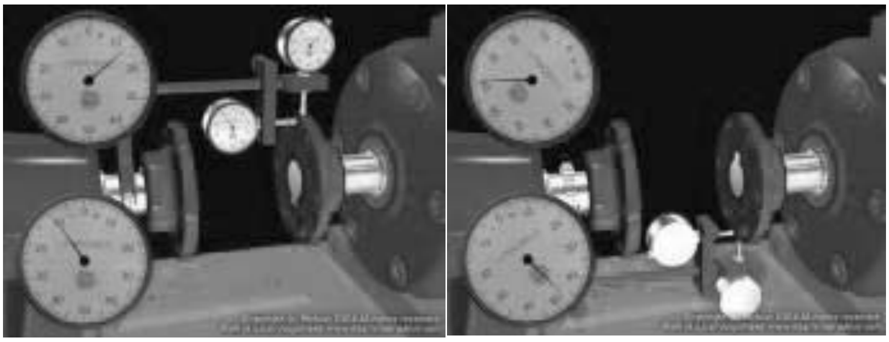


Figure 16-47 Rim and face test

This rim measurement will have revealed whether the motor shaft is higher or lower than the pump shaft (at the coupling). If the motor shaft were lower, the plunger would have been pushed into the dial (as it moved from the 12:00 position to the 6:00 position) which will be recorded on the dial face as a positive reading.

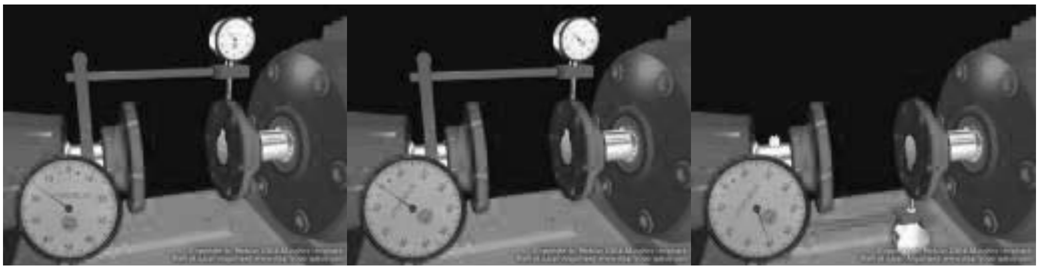


Figure 16-48 Rim measurement

Likewise, the face measurement will tell us if the motor shaft slopes away higher or lower (or parallel) to the pump shaft. If it sloped downward, the plunger would be pushed into the dial as the rig is rotated to the 6:00 position.

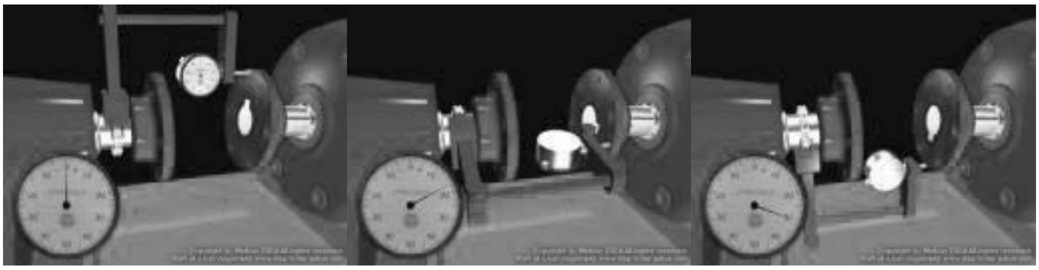


Figure 16-49 Face measurement

Based on the readings we take at the 12:00, 3:00, 6:00 and 9:00 positions, we can determine the offset and angularity, and we can calculate the changes required to the feet height and lateral position in order to make the two shafts collinear.

We can also do this graphically. By drawing the relative positions of the shaft to scale on a piece of graph paper (once in top/plan view, and again in side/elevation view), we can determine how the feet need to be moved.

The Reverse Dial method

The reverse-dial method is the most commonly used dial indicator alignment method used today. In this case we attach two dials to the two coupling rims to take two offset readings.

Again, there are a number of ways that these dials can be set up in order to collect essentially the same information. There are a number of commercial systems that employ this method.

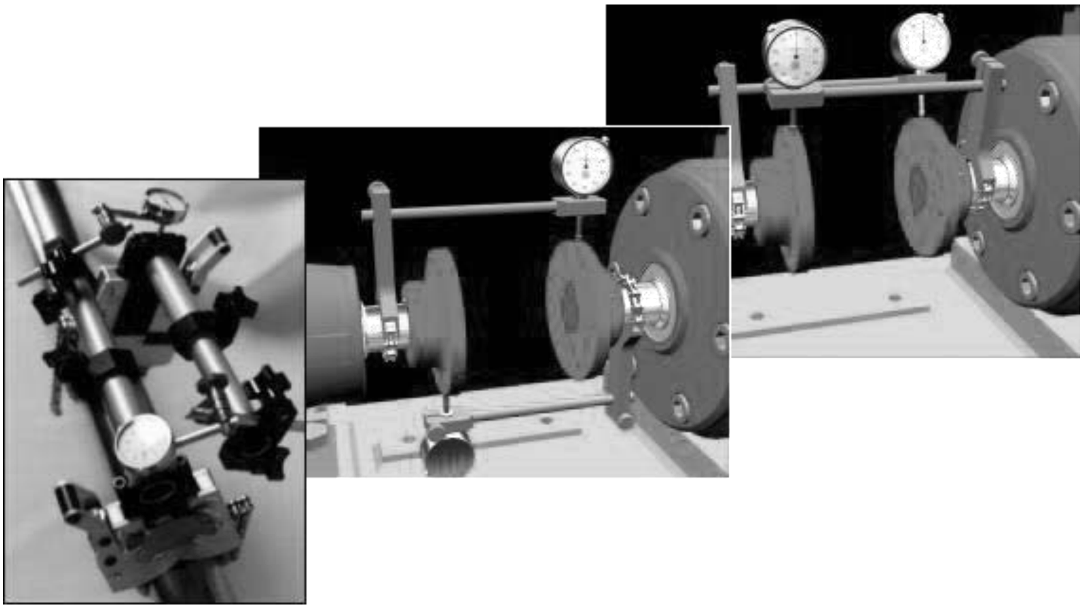


Figure 16-50 Different setups for the reverse dial method

As before, the two shafts are rotated, and the plungers either move into the dial or out of the dial based on the relative position of the shaft rotational centerline at the measurement points.

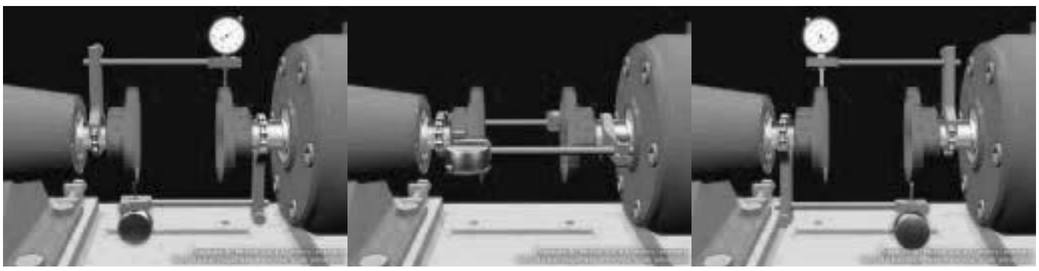


Figure 16-51 Reverse dial test

To determine the relative positions in the vertical direction (for shimming) we start at the 12:00 position and rotate to 6:00 and record the two values.

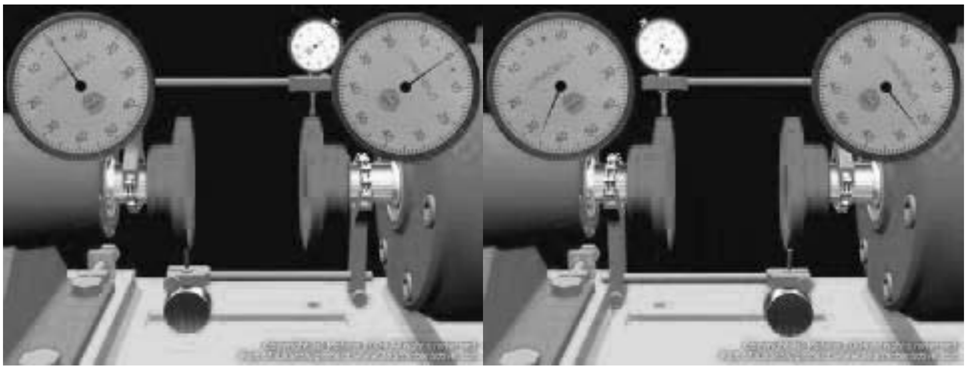


Figure 16-52 Measurements at the 12:00 and 6:00 positions

For horizontal foot movements, we compare the readings at 3:00 and 9:00.

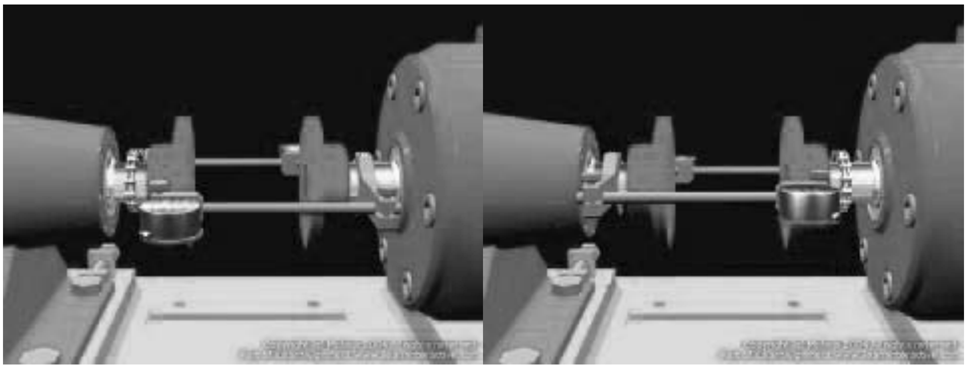


Figure 16-53 Measurements at the 3:00 and 9:00 positions

We can then use these dial readings and perform calculations to determine the offset and angularity at the coupling, and to compute the changes required to the feet height and lateral position in order to make the two shafts collinear.

We can also do this graphically. By drawing the relative positions of the shaft to scale on a piece of graph paper (once in top/plan view, and again in side/elevation view), we can determine how the feet need to be moved.

Laser alignment systems

Laser alignment systems come in many shapes, sizes and colors. They employ a range of technologies, from prisms that reflect the beam, to dual laser transmitters/detectors, and detectors that can assess movement in five axes.

Figure 16-54 Modern laser alignment systems

The two basic components of the laser alignment system are the "emitter" (sometimes called the "transmitter") and "detector" (sometimes called the "receiver").

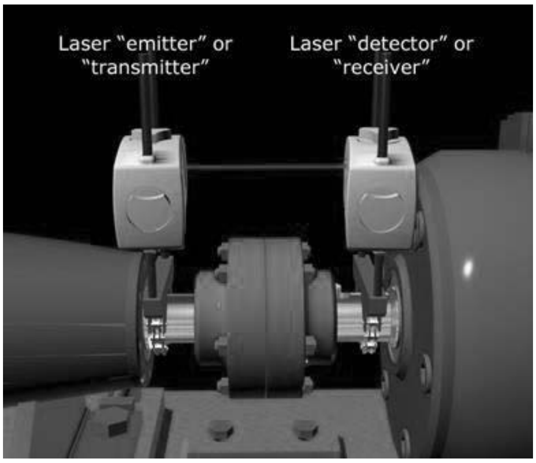


Figure 16-55 Emitter and detector –basic components of a laser alignment system

A number of the systems work in a very similar way to the reverse-dial indicator method: each head has an emitter and detector. The two laser heads are attached to the shaft on either side of the coupling. The laser heads are zeroed, and then the shafts are rotated. The detectors monitor the change in position of the laser - just like watching the plunger being pushed into the dial. In some cases the laser detector can detect movement in two dimensions - "x" and "y".

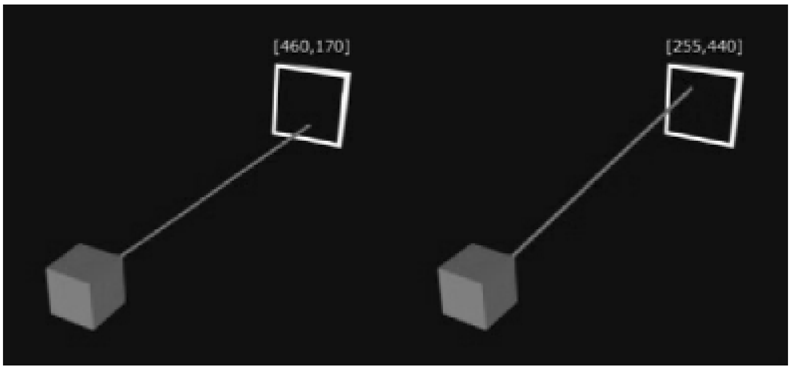


Figure 16-56 Detecting movement in the x and y directions

In other cases, the laser detector is sensitive in a single axis - just like the dial indicator (but FAR more sensitive and accurate).

Regardless of the method used, the shaft is always rotated (if possible) in order to measure the position of each shaft's rotational centerline. Using advanced technology, it is possible to rotate the shaft through only 60 degrees in order to gather enough data.

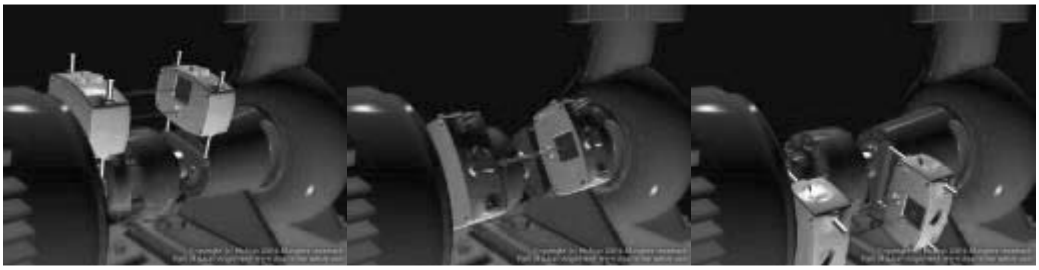


Figure 16-57 Gathering data through a 60 degree shaft rotation

The laser alignment systems are not only extremely accurate, they also come with "computers" that perform all of the calculations. In many cases they show you graphically how the feet should be moved, and can tell you when the alignment is within tolerance.

Laser systems are designed to be used in harsh environments; and they are safe to use - but don't stare into the beam!

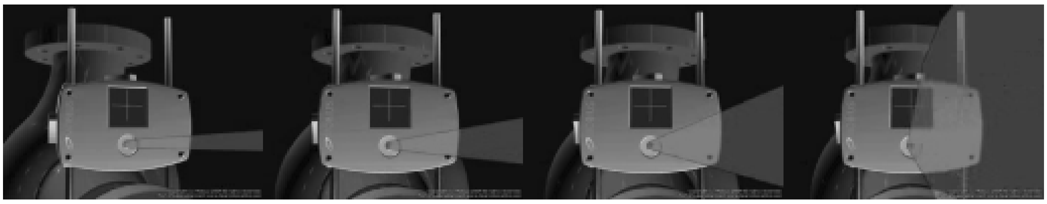


Figure 16-58 Don't stare into the beam!

Measure thermal / dynamic machine movement

Some laser alignment systems come with special brackets that allow you to attach them to the machine structure while the machine is running. The differences between the hot and cold running conditions can help you to measure thermal growth and to refine your alignment targets appropriately.

In an ideal case the alignment tools are placed on the machine when it is cold, a reading is taken and then compared to when the machine is running and hot (not the other way around).

Moving the machine

When the alignment measurements have been performed, you will determine how the feet should be moved laterally and vertically (by adding or removing shims) so that the rotational centerlines are collinear.

Figure 16-59 Adding a shim

As long as there is not a gross misalignment condition, you should always perform the vertical correction before a horizontal correction. A horizontal correction will not affect the vertical position of the machine, but a vertical correction will always affect the horizontal position.

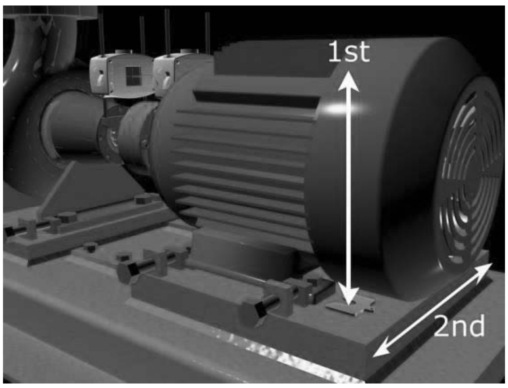


Figure 16-60 Perform the vertical correction first and the horizontal correction second

When a vertical correction is made it is almost impossible not to disturb the horizontal position of the machine. Unless the base plate is severely warped, when you make a horizontal correction the vertical position of the machine should remain constant. If there is gross misalignment (greater than 40 mils or 1 mm), you should make an initial vertical move, and an initial horizontal move, before you make the final vertical and horizontal moves. The initial move is required because unless a base plate is perfectly flat, a large horizontal move will result in a small vertical change.

Moving the machine vertically – shimming

The vertical correction is made by adding or removing shims. You should never leave more than four shims under the foot of a machine. Shims should be clean, straight, and carefully manufactured.

Figure 16-61 Using shims for vertical correction

You may find that you cannot remove enough shims in order to bring the machine into alignment (i.e. it is "base-bound"). In this situation you may need to lift the front feet of the "stationary" component (the pump, for example), as well as the front feet of the "movable" component (the motor) - or you can machine the feet or the baseplate.

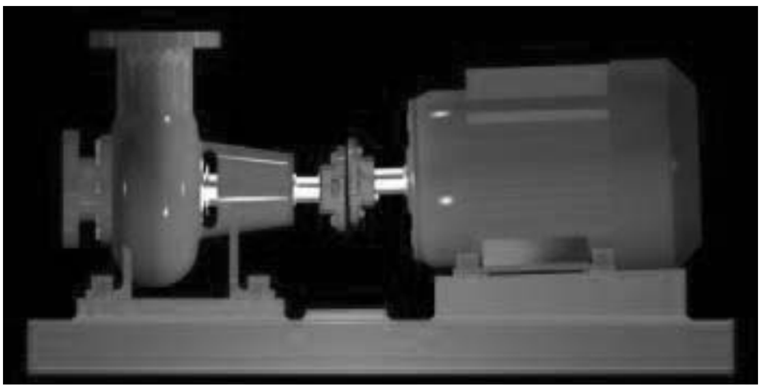


Figure 16-62 “Base-bound” motor

Moving the machine laterally

Although sledge hammers and pieces of 2 x 4 timber can be used to move the machine laterally, it is highly recommended that jacking bolts are used instead.

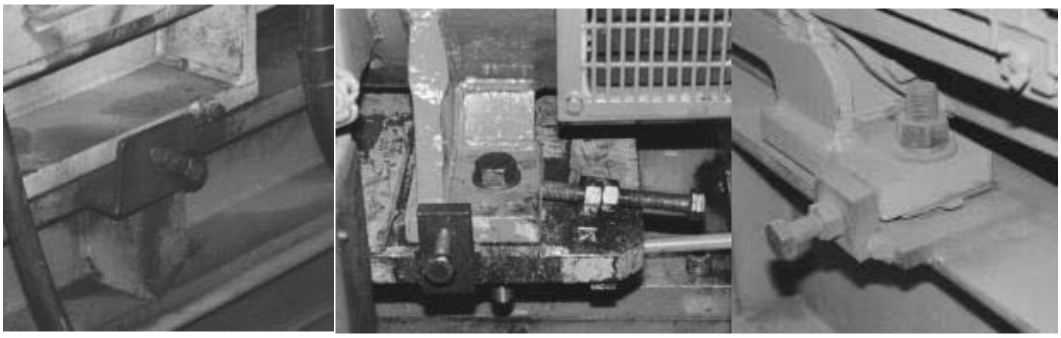


Figure 16-63 Jacking bolts for lateral movement

Laser alignment systems can be used to monitor the movement of the machine so that you can make positional changes until you are in tolerance. Sometimes the movable component cannot be moved far enough to bring it into alignment - it may be "bolt-bound". While you could move the stationary component, you can also consider opening the foot holes, turning-down the bolts, or drilling new holes in the base plate.

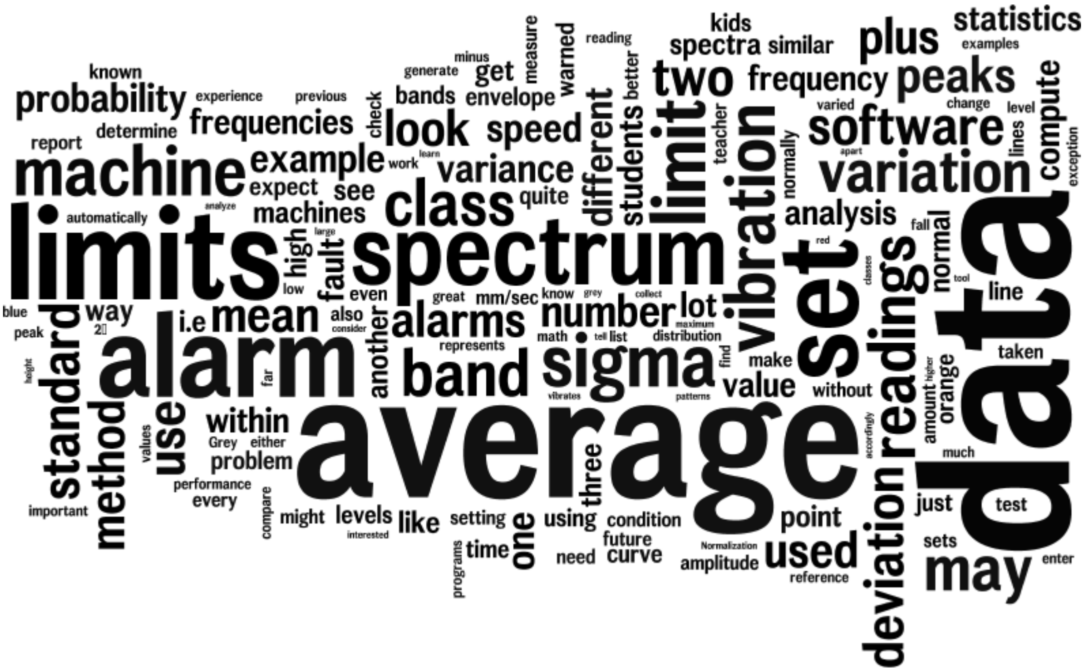
Whenever you tighten the hold-down bolts, you should ensure that they are always tightened to the same torque. It is a good idea to lubricate the bolts before performing the alignment. It is also recommended that you always tighten and loosen the hold-down bolts in the same order.

Figure 16-64 Tighten hold-down bolts to the same torque

After you have moved the machine, and tightened the hold-down bolts, you must always repeat the measurements to check that the machine really is still in tolerance. If you have had to make large changes either vertically or laterally, you should also recheck the soft foot condition.

Conclusion

Precision alignment is critically important to the ongoing reliability, and thus profitability, of your plant. The additional time spent performing the alignment can always be justified. Laser alignment tools make the measurement and movement of the machine easier, however it is essential that you understand the measurements, movements and all of the potential pitfalls in order to ensure that you can successfully align machines under all circumstances.



Chapter 17

Statistical Alarm Generation

Objectives:

- Describe the benefits of statistical alarms
- Give a brief refresher on statistical analysis

In this chapter we will discuss the benefits of using statistics to set alarms. Alarms serve two general purposes. They tell us which machines have problems and they help to reduce the number of data sets that require manual analysis. Despite the huge benefits of setting alarms, unfortunately, many people never get around to implementing this feature of their monitoring systems. Therefore the goal of this chapter is to urge you to use alarms!

Generating alarms with statistics:

In an ideal world, you could set alarm limits for all of the vibration data you collect and you would only get 'exceptions' when there really was a problem, and you would never have a problem without being warned. Sounds great, but for most people it is an unrealistic goal. In this chapter you will learn about a way to set alarm limits statistically; in essence, the machine sets its own alarm limits. This method utilizes previous readings to evaluate how the machine normally vibrates and then establishes alarm limits accordingly.

After collecting a route full of vibration data, the task at hand is to review the data and find any machines that may have problems. If you have tested a large number of machines, a significant amount of time will be required in order to look at each spectrum , check the levels and patterns, and document your findings. If you look at every spectrum then *surely* you won't miss a fault condition, after all, you remember what previous spectra looked like, and you are very careful with your analysis – *right?*

But it takes a lot of work to analyze so much data and it is easy to make mistakes. There is a lot of detail in each spectrum; therefore there is a lot of information (data) to study. People get tired; their concentration wanes. It is human to err. Couldn't we use the computer to do some of this work?

The need for alarms

Alarms can be used to detect faults or to determine if one meets requirements or regulations. Alarms can be set on all data types. One reason to set alarms then is simply to see if a machine is "OK" or not "OK". In other words, alarms act as acceptance criteria or to tell us what levels are or are not acceptable.

A second use for alarms is to help you prioritize your work.

There is something called the 80-20 rule which states that in a typical plant, 80% of the machines do not have problems. Of the 20% of the machines that do have problems, only 20% of those require immediate attention. Now imagine you are testing 1,000 machines per month, with an average of 6 test locations per machine and 4 tests per point (spectra, time waveform, overall, demodulation) – that adds up to 24,000 tests per month. How will anyone have the time to analyze all of this data? Remember, we are monitoring our machines to *save* money, not create an enormous load of extra work!

Using the 80-20 rule, 20% of the 1,000 machines have problems (that is 200) and only 20% of those require attention (that is 40). The second goal of setting alarms then is to let us know immediately which 200 machines have problems and then point us to the 40 machines that require attention in such a way that we do not have to manually analyze each and every bit of data from the other 960 machines.

The problem with alarms

It can be difficult to set alarms in some cases and to be honest and up front about it, it takes effort and work – although much less work then analyzing all 24,000 tests each month! Many people will say they do not have time to set alarms but if you question them further you will discover that the reason they do not have time to set alarms is because they are spending all of their time manually analyzing the data!

Another problem with alarms is that if they are set incorrectly you will get false positives and false negatives – in other words, you will lose trust in your alarms. If you cannot trust your alarms then you will still feel the need to manually analyze all of the data or your program will just turn into a crap shoot. In short, to be useful, alarms must be set up well and you will have to learn to trust them.

Setting alarm limits

All vibration analysis software programs give you a way to set the alarm limits, and the software will check your readings and report on any exceptions. We will briefly describe how this is done in a moment, but the expectation is that the alarm limits are good, and the method used to check new measurements is able to find faults no matter where they are in the data.

Band alarms

A commonly used method is known as "band alarms" (also known as "analysis parameter sets"). Originally six bands were supported by the main software programs, but now many support a larger number. Each band has a start and stop frequency, and a threshold (maximum vibration limit). In most cases you can have more than one limit so that you have an "alert" and an "alarm".

Somehow you have to decide what those limits should be. You have to look at your machine, determine the forcing frequencies, and then set the bands around those frequencies. For example, if a pump had 6 vanes, we may have a band that started at 5.8X and ended at 6.2X.

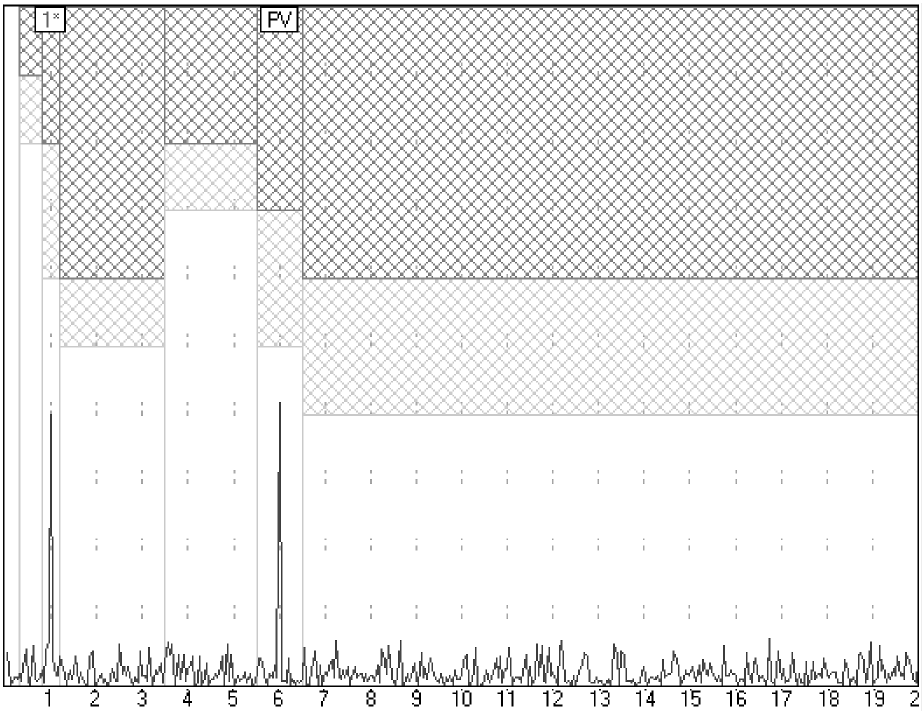


Figure 17-1 - Sample band alarms

But as we know, individual frequencies do not always indicate a fault in isolation. Instead we are interested in harmonics and sidebands, so we may create bands that take in a wider range of frequencies: 1X -> 10X for example.

This can be a tedious task. You have to make that assessment for every point on every machine, and then enter the settings into the software.

But the real challenge is knowing what value to enter for the alarm limit. Is 0.5 mm/sec the limit? Should you use 1.0 mm/sec? Is 0.2 mm/sec too high for this machine? Is 1.5 mm/sec too low for that machine? If you get it wrong you will experience two problems: machines will fail without warning, and you will read exception report after exception report filled with "false alarms" – yes, the limit was exceeded, but the limit should have been higher.

There has to be a better way…

Attempts have been made to define standard limits. The following illustration comes from the General Motors acceptance testing guidelines.

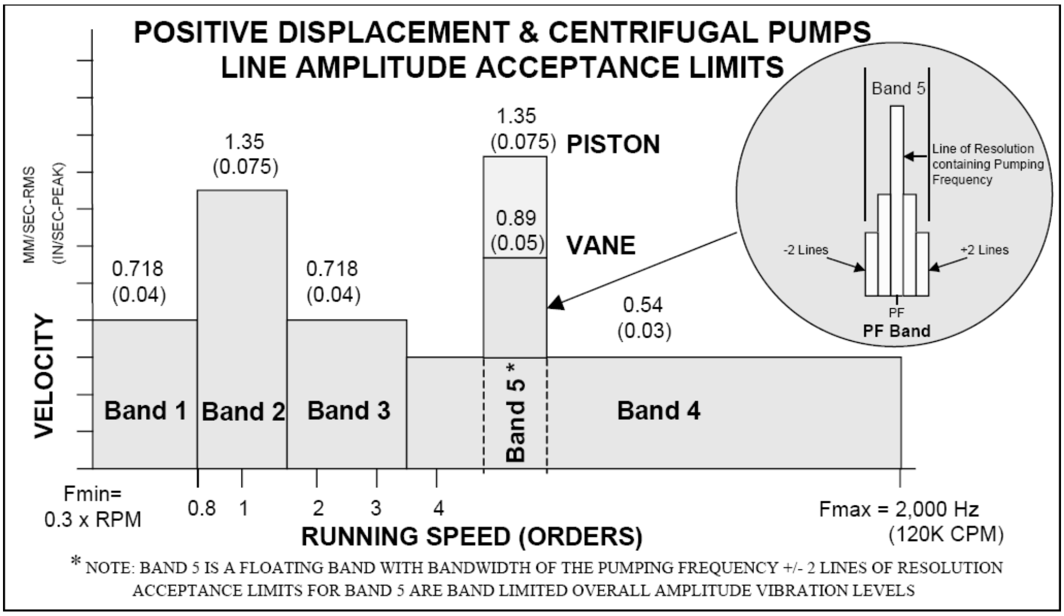


Figure 17-2

In this author's experience, the use of "pre-defined" limits has not been successful. How can you set limits for a generic "pump"? Don't the vibration limits depend upon its size, its load, how it is mounted, where it is located – and more? Even two identical pumps can have different vibration patterns and levels without either of them having a problem.

There is one thing we know about vibration analysis; change in vibration levels is a far better indicator than absolute level.

Of course, even if the limit set on the band was correct, there is another problem, which is perhaps the topic for another paper – the band method itself is flawed. If we make the assumption that there is just one peak in the band, and that is all we are interested in, then the simple band method would work just fine. But if there are sidebands, harmonics, and non-synchronous peaks, peaks may significantly increase in amplitude within a band without triggering the alarm.

One solution is to compute the "power" or "rms" level in the band – a summation of all the vibration within the band, and set a limit on that value. Again, somehow you have to determine what that limit is, but that is a more effective method than setting a single threshold.

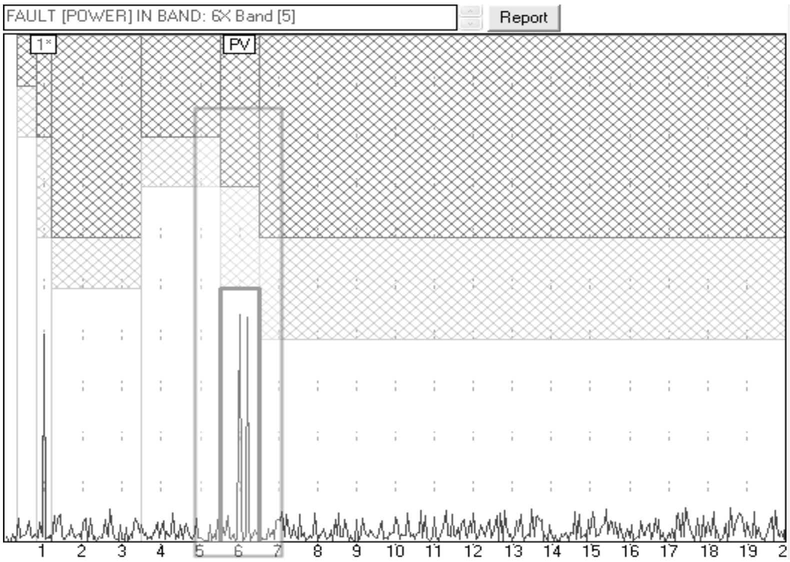


Figure 17-3

It should be said that band alarms do have another benefit: trending. If you trend the peak and rms levels within a band, you can see how the vibration is changing in your spectrum. That is a great analysis tool, and it can help to draw our attention to a fault that is developing – but it is not exactly an automated fault detection method.

Mask or envelope alarms

There is another method used by many of the vibration analysis systems known as either mask alarms or envelope alarms.

As shown in the following example, the envelope is a limit set at all frequencies. There are no gaps.

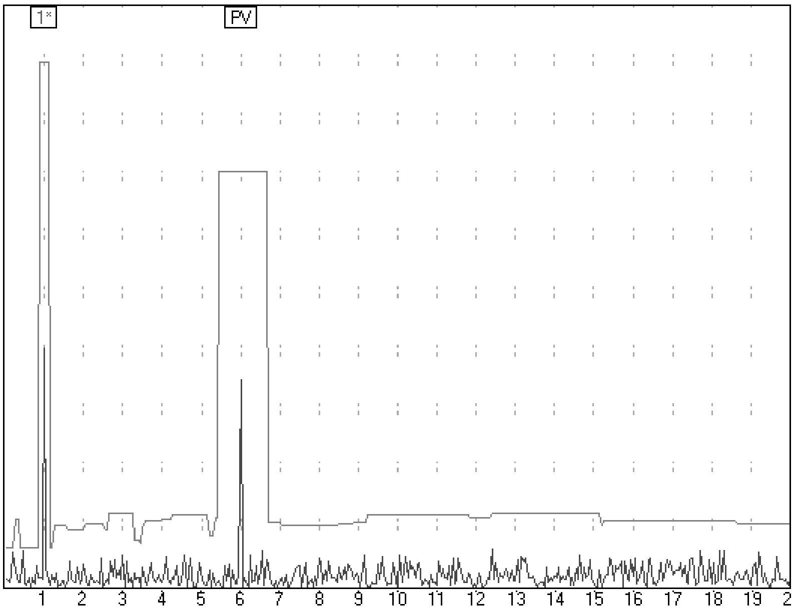


Figure 17-4

The exception report may just list the frequency where a peak exceeds this limit, or it might correlate the frequency to the known forcing frequencies and either list the name itself "PUMP VANE RATE", or even attempt to list the fault condition.

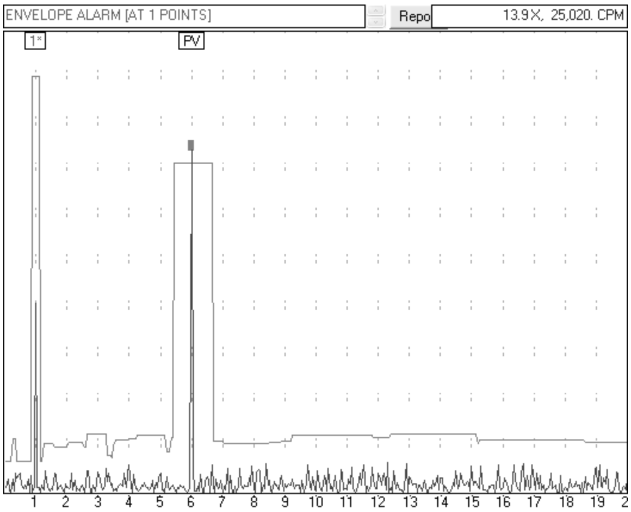


Figure 17-5

When setting the envelope limits, we have a number of decisions to make; what should the amplitude limits be, should we impose a maximum limit, should we impose a minimum limit, and how closely to the peaks should the envelope be created – that affects its ability to deal with speed variation – discussed later.

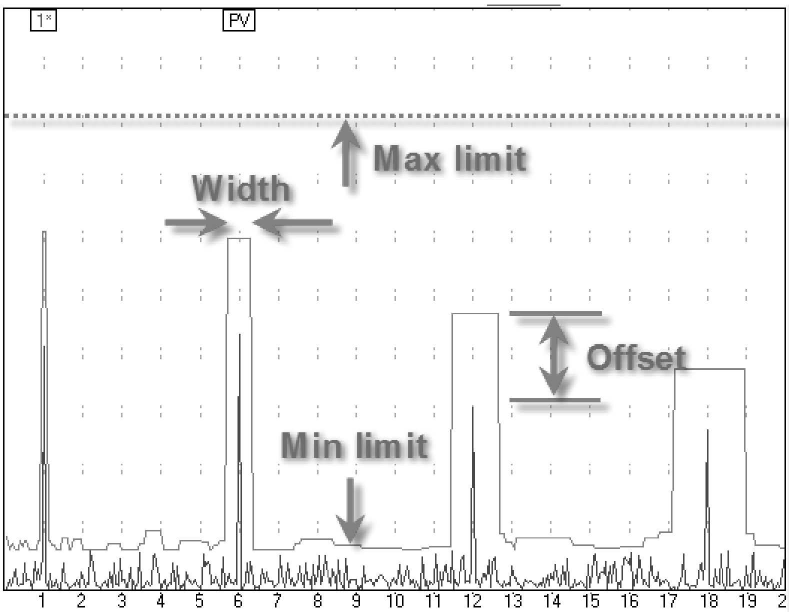


Figure 17-6

Many software programs provide a means to generate these limits automatically, but the key question is; what reference is used? Does the alarm generation software use the "reference" or "baseline" spectrum; does it use the latest reading taken, or does it compute a spectrum based on a statistical calculation that uses all previous readings? We will now explore the possibility of using statistics.

A statistics refresher

Unless you completed an engineering degree, or you studied math at a high level, terms like statistics, variance, and standard deviation may not be familiar to you. So let's get (back) up to speed on what these terms mean, and how they can help us to set alarm limits.

Statistics can be used to tell us how similar a set of values are. For example, if we asked a group of grade 6 students to write down their heights, we would find that there would be an average height with some variation; some kids are taller and some kids are shorter. The average height is determined by adding all of the heights together, and dividing by the number of students. In a situation like that we might expect there would be quite a lot of variation: some kids can be quite tall at that age, while others are still quite short. Let's look at a different example.

If we wanted to compare the performance of three math teachers, we might give all the students in the three classes the same test. We can then add all the scores in class 'A' and divide by the number of students, and do the same in class 'B' and 'C'. If we compare the averages we get a rough measure of the performance of each teacher. The terrible fact is that in most classes these days, 50% of students may be below average ;)

But what if class 'A' has three children with learning disabilities and they perform poorly on the test? They will drag the average score of the class down. Is that a real reflection of the ability of the teacher? And what if class ‘B’ has three brilliant students that always blitz the test? They will drag the average up. But their performance may reflect the abilities of the teacher; he was just lucky to have smart kids in the class. The third class may happen to have a lot of average kids.

So the average may not be a great measure of the abilities of the teacher (unless each class had a very large number of students, and the teachers were in control of the class for a long time).

In order to capture information about the highs and lows in the class, we would like to have a measure of the “variation” – or “variance”. The variance would be high in class ‘A’, and high in class ‘B’, but low in class ‘C’. We will see an example of this shortly.

Truth be told, we could use such a measure in vibration analysis during data collection. When you collect spectral data the analyzer keeps the average. But you don’t know if the analyzer averaged a number of very similar spectra, or whether at certain frequencies there was a lot of variation – which means you have reduced repeatability, and you could have beating.

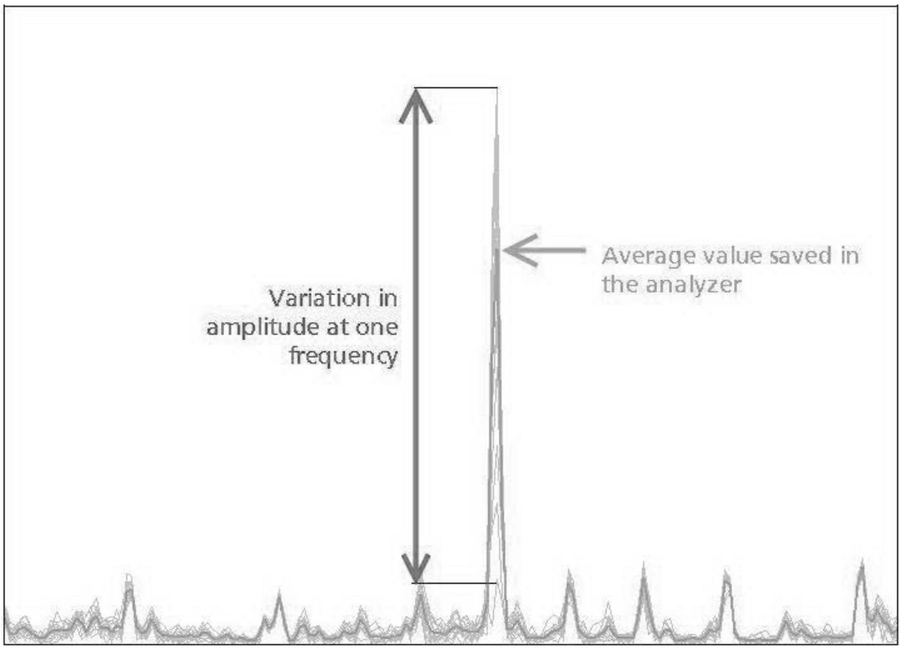


Figure 17-7

So, whether we are studying children’s math scores, vibration readings, or anything else, the average value does not tell the entire story. If we understood how much variance there is in the data then we get a better picture of what is really going on.

Let’s have a look at four examples. In each case we have a set of data that has the same average. It is obvious that the data sets are quite different. The average (or mean) may be the same, but there is a grossly different amount of variation in the data.

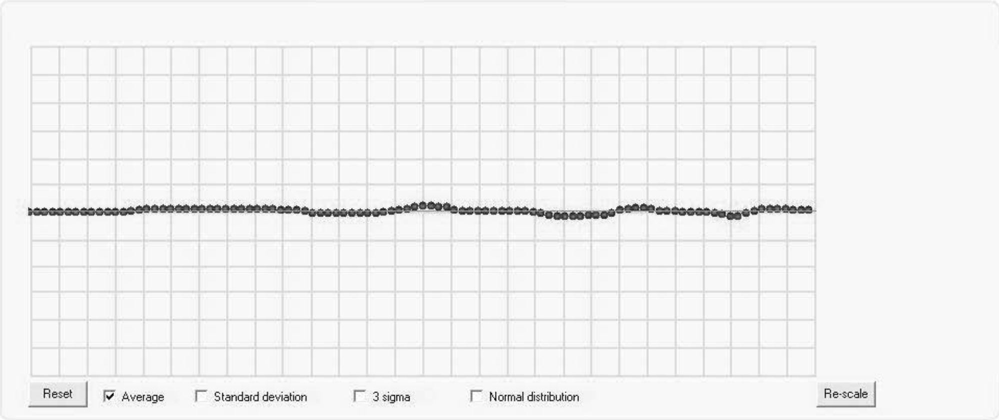


Figure 17-8

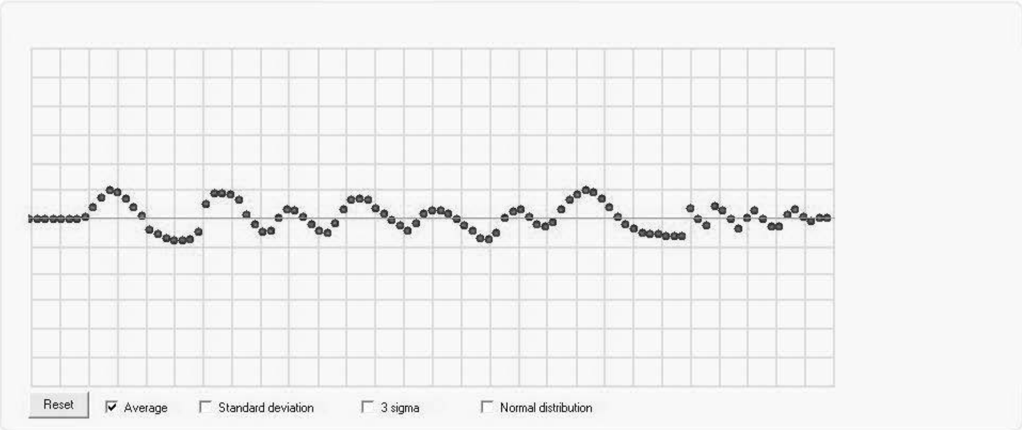


Figure 17-9

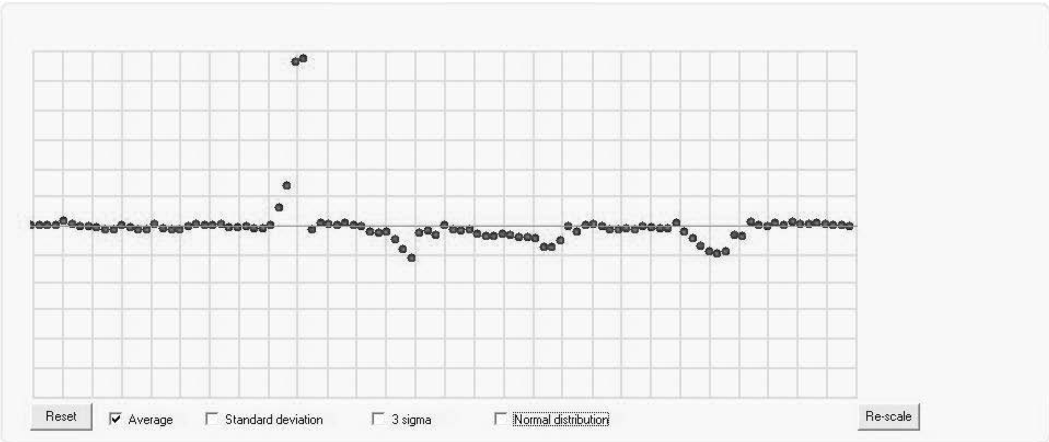


Figure 17-10

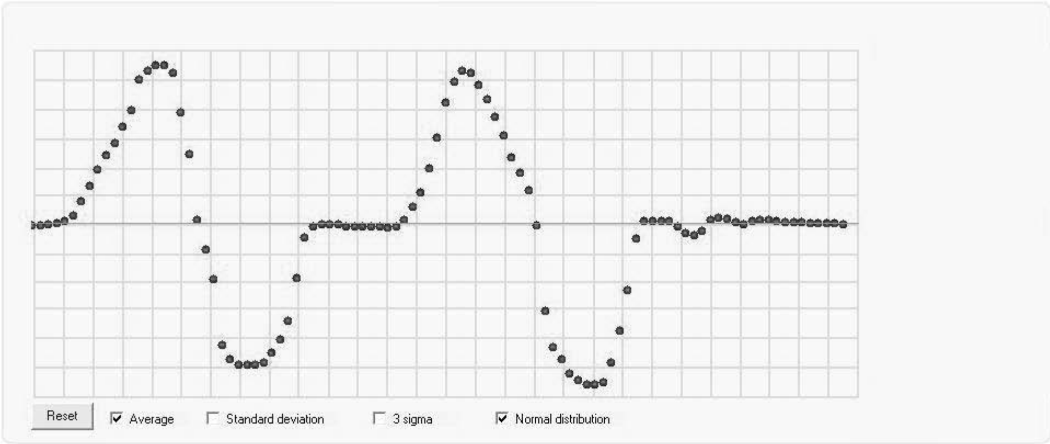


Figure 17-11

Variation

There is a way to quantify the variation in the data – it is called *variance*. Variance is the average of the squared differences between data points and the mean (average). But we do not need to explore this further as we will not use it.

However the concept is important. If the values do not vary much, the variance will be low, and we expect all of the readings to be similar to the average (e.g. Figure 17-8). If the variance is high, then we do not expect readings to be similar to the average (e.g. Figure 17-11).

Standard deviation

The standard deviation measures the spread of the data about the mean value. The standard deviation is usually denoted with the letter σ (lower case sigma). It is defined as the square root of the variance; therefore it is measured in the same units as the data (and the average).

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

(N is the number of samples or reading, x_i represents the ith sample, and $\bar{x}$ is the mean value.)

We don't need to look closely at the formulas used to compute the standard deviation, but there are a few very interesting facts that are important.

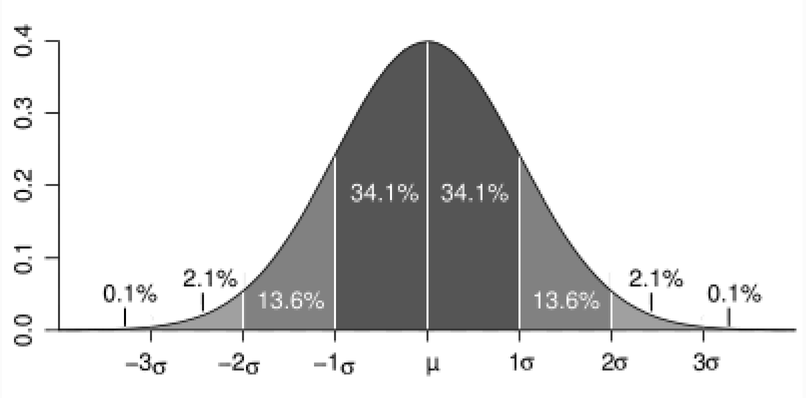


Figure 17-12

The curve you see above is the "probability distribution" curve, assuming that the data is normally distributed. It shows the "probability distribution" of the data. If we compute the mean value (μ) of a set of readings, and compute the standard deviation (σ), then we can say that there is a 68.2% probability that any future readings will fall within the mean plus or minus the standard deviation
(i.e. μ ±1σ). We can also say that there is a 95.4% probability that any future readings will fall within the mean plus or minus the 2 x standard deviation (i.e. μ ±2σ). And we can say that there is a 99.6% probability that any future readings will fall within the mean plus or minus the 3 x standard deviation (i.e. μ ±3σ).

So let's have a look at our four sets of data, but we will overlay the ±1σ, ±2σ and ±3σ lines, and we will show the probability distribution curve.

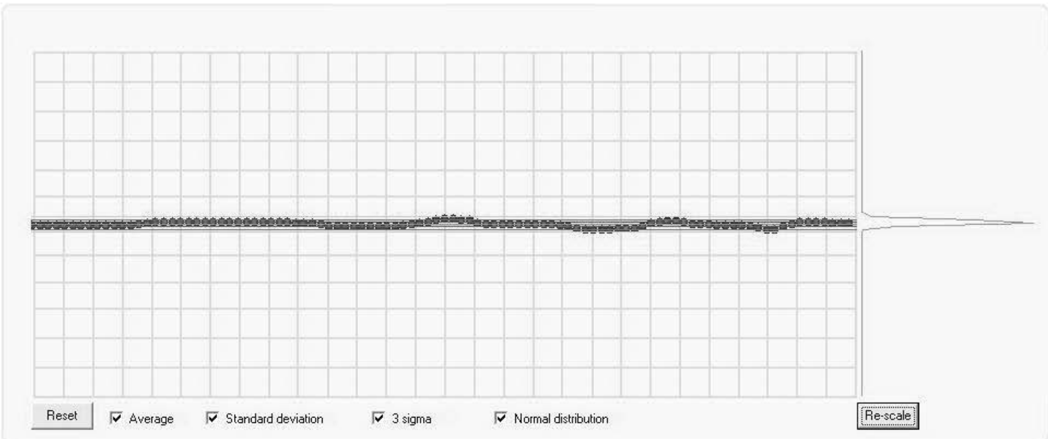


Figure 17-13

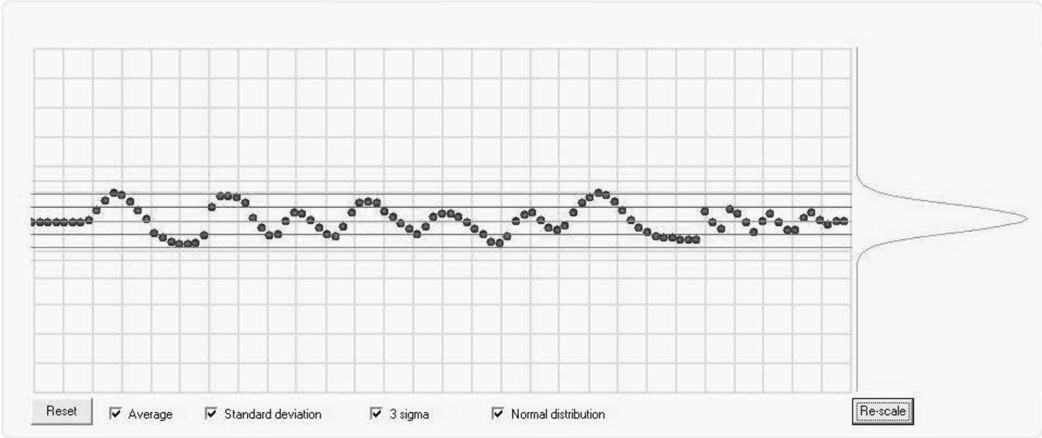


Figure 17-14

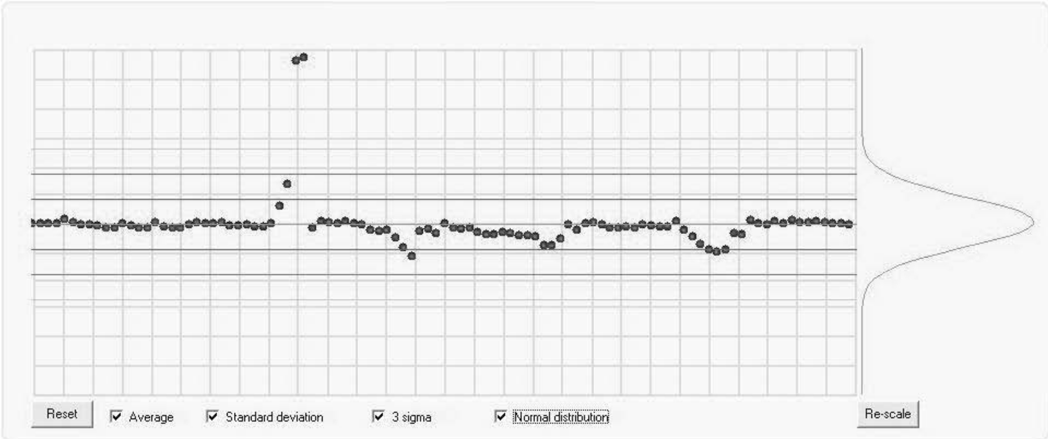


Figure 17-15

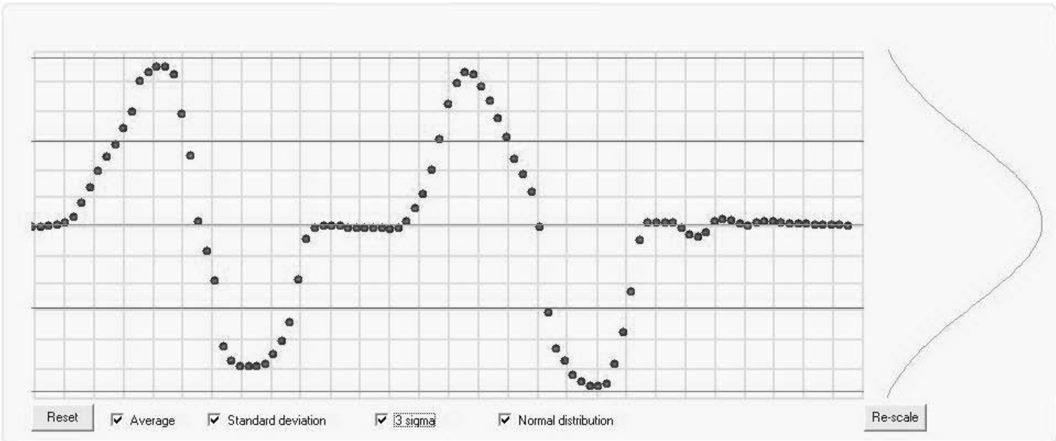


Figure 17-16

Notice how in the first set of data the curve is very tight, and you can barely see the sigma lines – the probability of the next reading being very close to the mean value is high. Whereas with each set of data the curve changes and the sigma lines move further apart.

Using statistics to set alarm limits

OK, so what does all this have to do with setting alarm limits? Well, if you were to look at a set of ten spectra from a point on a machine (taken one month apart) and saw that the amplitudes at each frequency in the spectrum were the same (all within 1% variation), then we would expect the eleventh spectrum to also be similar. If it was 10% higher in amplitude at any frequency then we would consider that to be unusual – and as an analyst we might like to be warned that this has happened.

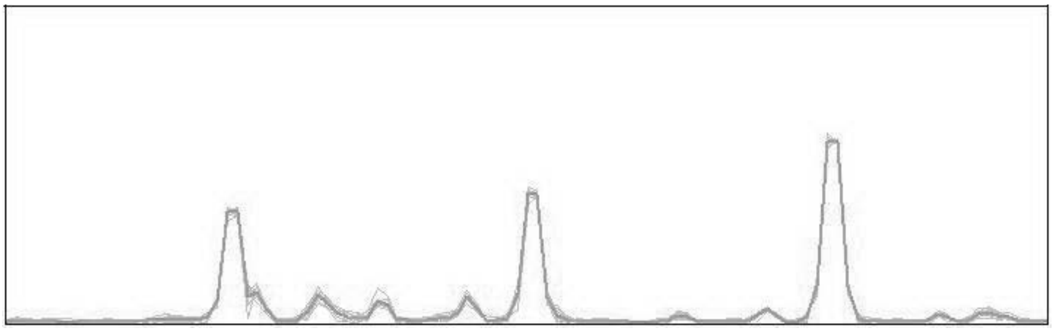


Figure 17-17 Zoom of three peaks in the spectrum – grey data represents the past readings and orange is the average

Let's consider a different example. If we again looked at ten spectra taken from a point on a machine, but this time we found that the amplitude levels at each frequency varied by 10% then we would expect the eleventh spectrum to also vary by up to 10%.

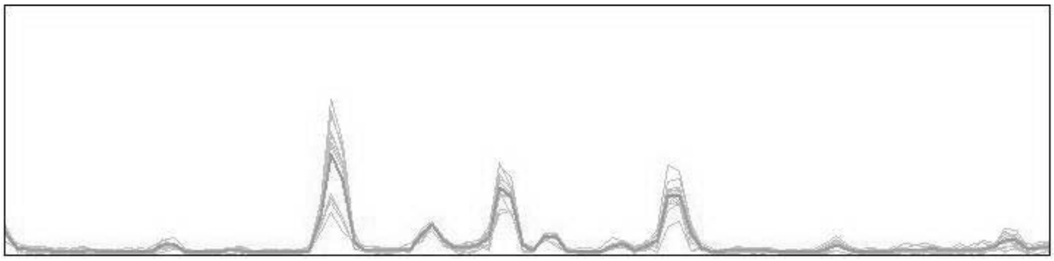


Figure 17-18 Zoom of three peaks in the spectrum – grey data represents the past readings and orange is the average

We can use this to set alarm limits – and we can *potentially* do it automatically in software. The software can analyze the history of measurements, and compare all of the amplitudes, line by line, spectrum by spectrum, point by point. It can compute the average at every frequency, and determine the amount of variation (i.e. the standard deviation). Then it can set the alarm limit based on the average, plus sigma (or two sigma) plus an additional offset if you wish to be warned only when the variation is even greater.

Just to clarify, we only compute the average and standard deviation using data from the same point on the machine.

You need to check the capabilities of your software. The author has personal experience using average plus 2 sigma (i.e. $\mu \pm 2\sigma$) as the reference, and then adding offsets to that calculated spectrum for alarm and alert values. By using average plus 2 sigma, you 'believe' there is a 95% probability that future readings should be below this limit. Let's have a look at these two examples again to see what average plus 2 sigma limits would look like.

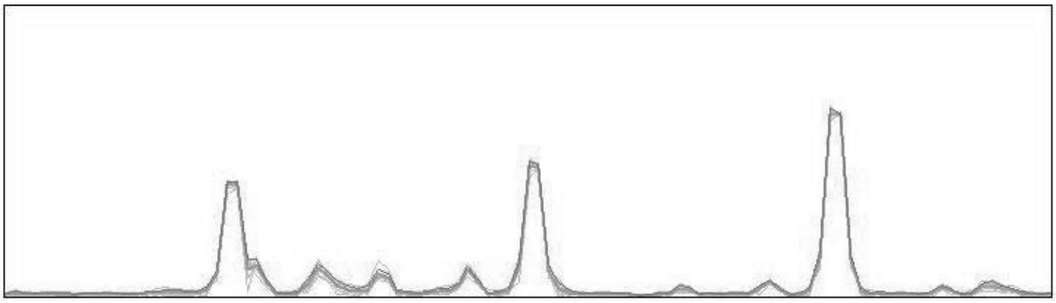


Figure 17-19 Grey is normal data, orange is average, and blue is average + two sigma

There is so little variation in the data that it is difficult to see the difference between the data, the average and the average plus two sigma lines.

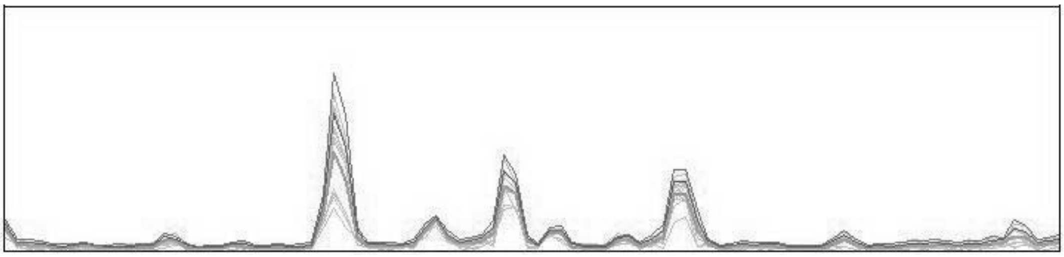


Figure 17-20 Grey is normal data, orange is average, red is average + one sigma, and blue is average + two sigma

In the second set of data the average plus two sigma data is above all of the measured data (in this example). If we used this spectrum to set the alarm limits, then we could feel confident that only 'unusual' data would exceed the data.

Guess what. The data displayed in the last two examples is actually from two sections of the same spectrum – one group of peaks varied very little, while another part of the same spectrum varied quite a lot. So statistics can be used to protect different parts of the spectrum in different ways – all with the same effect – determine what is normal and set alarm limits accordingly.

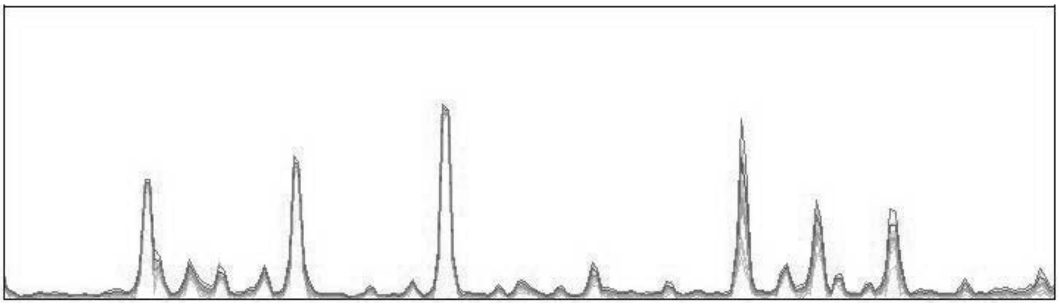


Figure 17-21 Grey is normal data, orange is average, red is average + one sigma, and blue is average + two sigma

Data quality

When we use this method to generate the alarms, it is important to only use "good" data. If spectra were collected when a fault condition existed, or the data was taken incorrectly, or the test conditions were wrong - then the data should not be used. We are trying to learn how the machine would *normally* vibrate so that we can be warned when it no longer vibrates normally. Therefore the data used in the calculation should represent *normal* operation of the machine.

Normalization: Dealing with speed variation

There is another problem when using this method – what if the speed varies from one spectrum to the next? Unfortunately, the speed *will* change, even if just slightly. When dealing with variable speed machines the challenge is far greater because load may change, you excite different resonances – and because the peaks do not line up and all.

Some software packages provide options to 'normalize' spectra. Normalization ensures that all of the 1X peaks line up, and the 2X peaks line up, etc. so we are comparing apples with apples. You may be required to enter the machine speed (or the paper/belt/roll speed), or use a tachometer while collecting the data.

Another way to utilize statistics is to apply it to bands of vibration (i.e. the alarm bands described earlier) instead of to individual peaks in the spectrum.

Conclusion

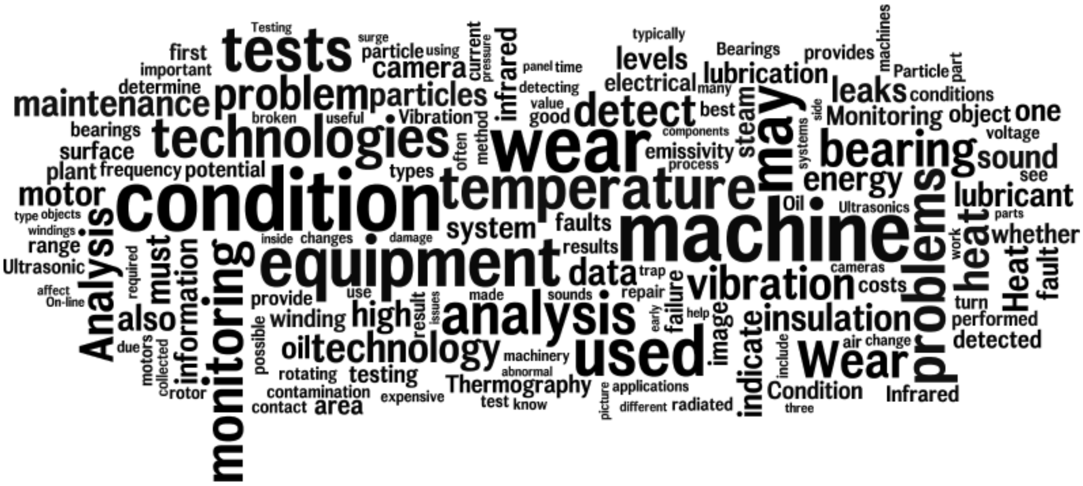
Statistics can be a powerful tool for a number of reasons:

1. The alarm limits can be set automatically (depending upon your software), saving you a tremendous amount of time.

2. The alarm limits are customized to each machine – in effect your machines are setting their own limits.

3. The alarm reports should be far more effective – you should only be warned when a machine has changed from what is normal.

If you choose the data carefully you can generate a set of limits that will save you time, reduce the chance of missing a fault condition, and enable you to focus on the machines that need most attention.



Appendix A

Condition Monitoring

***Objective:** List six Condition Monitoring Technologies and three applications for each.*

Condition Monitoring is the art of monitoring plant equipment to determine its health or condition at a point in time.

Condition monitoring has two elements:

1. **Determine** whether the machine is running in a stable condition or if it is deteriorating. If the machine condition is deteriorating, then determine how fast it is deteriorating, and where it is in the life cycle.

2. **Convey** the conditions to the appropriate people so that decisions can be made regarding any maintenance or process changes that may be required.

Introduction

Often the terms "condition monitoring" and "predictive maintenance" (or "condition based maintenance") are used interchangeably. In reality, they do not mean the same thing. "Condition monitoring is the act of determining the condition of a machine."

Condition Monitoring ≠ Predictive Maintenance

Predictive maintenance involves taking action based on the condition. Sadly, in some facilities, an excellent group of people monitoring the condition of important equipment, but action is not always taken because the relevant staff do not really believe in the predictive maintenance philosophy. These people will only believe that maintenance is required when you can actually hear the bearing screaming or feel the high vibration levels. It can be frustrating to work in this environment, however you must have patience, and a willingness to educate these people on the validity of the technology and the benefits of predictive maintenance.

In the previous chapter the four maintenance practices were compared to the health of human bodies. It was pointed out that doctors use different methods to determine the condition of the human body: temperature, blood pressure, heart pulse rate, even fluid samples. Several technologies are used to arrive at an accurate picture of our condition.

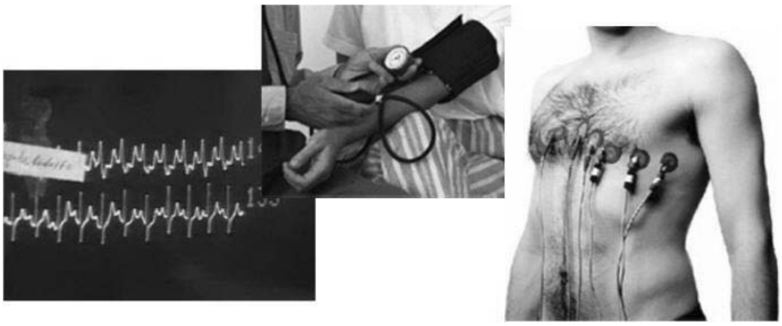


Figure A-1 Various technologies are used to determine the state of our health

For plant equipment, the process is similar. Rotating machines try to tell us their condition through changes in **vibration, temperature, lubricants, sound, and motor curren**t. So, processes are checked…temperature, pressure, flow, speed, motor current, vibration, fluid and lubrication analysis, and more. When the information is gathered, a clear picture can be built of the machine's condition.

It is true… it is possible to know what is happening inside a machine.

In determining the condition there are three points that should be considered.

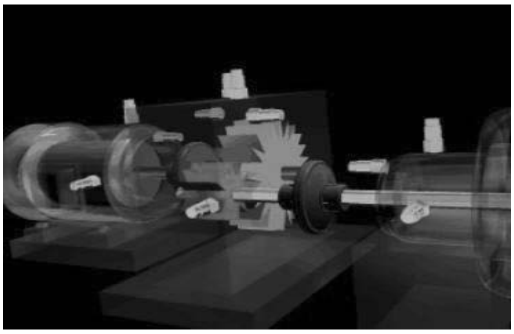


Figure A-2 Various technologies let us see the condition inside a machine

First, the best picture of a machine's health is arrived at when we perform all the tests and correlate the results. One test alone may be misleading or may give only part of the picture. In fact certain tests will not give us any information regarding some fault conditions. For example, wear particle tests on an oil sample won't indicate whether the machine has an imbalance problem. Vibration analysis won't be able to indicate whether there are contaminants in the lubricant. The technologies must be combined to provide **Integrated Condition Monitoring**.

The **second point** is that certain technologies and tests can provide information to indicate situations that can cause problems for a machine, not that they necessarily are causing problems. Examples:

Resonance per se is not a fault. All machines have resonances. The fact that a resonance at running speed can result in excess vibration and may destroy the bearings is a problem.

Contaminants in a lubricant is not necessarily a problem. The fact that the contaminants can damage bearings and other components is the problem.

The **third point** is that the best diagnosis will be achieved when there are previous results to compare to and data is available to use as a reference. Doctors know what a person's blood pressure should be, and what cholesterol levels are safe. Some guidelines do exist for vibration analysis and other technologies (ISO standards, for example) however for the most part we must rely on comparisons with previous data taken from the machine.

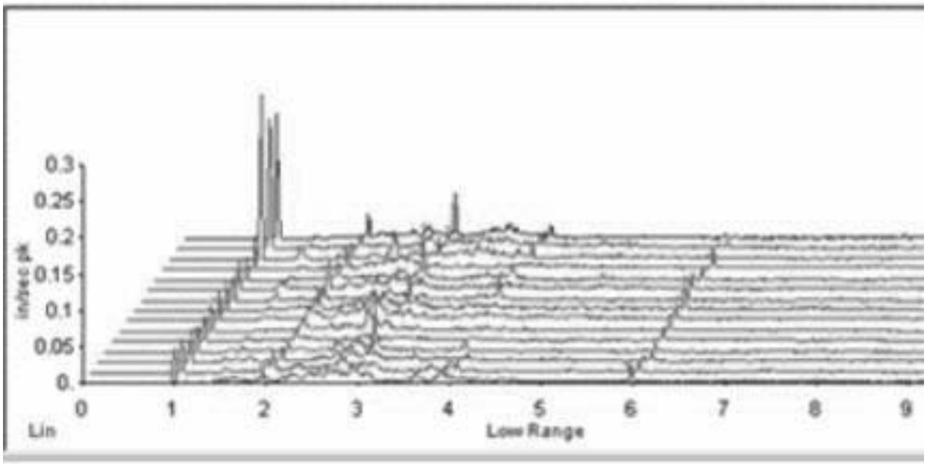


Figure A-3 Compare to previous measurements to see changes in condition

Condition Monitoring Technologies

Why Perform Condition Monitoring?

In one word the answer is profitability. When a machine fails **it can stop production**. Downtime can be incredibly expensive. Loss of production and spoiled product (and wastage) eats into corporate profits.

When a machine fails, **secondary damage** can occur. Instead of simply replacing a bearing, the entire rotor may need to be replaced. The costs can escalate. In addition, the repair time is longer, resulting in longer downtime and higher labor costs.

There may also be safety or environmental issues; machinery failure can result in injury (or worse), and liquids or gasses may escape causing environmental damage.

When a machine fails, spare parts must be on hand to complete the repair. Those parts must be held in stores "just in case" a failure occurs. Holding and managing those spares is very costly.

So, if we can assess the risk of failure the outage can be planned for the most cost effective time. Parts can be ordered, and labor organized. Expensive preventive maintenance does not need to be performed. This goal can be achieved through appropriate monitoring.

If the reliability can be improved, the machine will not fail as often. If a machine can reach its design life it is providing the greatest return on investment, and will result in the lowest possible maintenance costs - and thus greatest company profits.

The Whole Picture – Select the technologies and the test frequency carefully. The best results are achieved when the results from multiple technologies are integrated into one report.

Six technologies are typically used in Condition Monitoring programs. They are:

- Airborne Ultrasonics, also known as Acoustic Emissions
- Infrared Thermography
- Electric Motor Analysis
- Oil Analysis
- Wear Particle Analysis
- Vibration Analysis.

Each technology provides information about the machine. Combining the technologies helps in pinpointing the condition.

Why So Many Technologies?

One physicist who is a machinery troubleshooting expert says that his method is very simple; "I just ask the machine a lot of questions". He uses the various technologies to ask and get answers to specific questions regarding the condition. He then compiles the answers and comes up with the best response for the particular condition.

The physicist's plan sounds simple enough, but it requires that he know the types of information that can be obtained from each technology and the limits of each as well. While it not necessary to have the physicist's level of knowledge of each technology, a brief overview of each technology can aid in choosing what may be the best technology for a given situation.

Some technologies indicate whether a problem exists now and they help in determining the severity. Other technologies indicate that a condition does indeed exist that could become a problem in the future.

There are situations where acceptable levels are not really known. In these cases, it is wise to trend the levels and monitor for any changes.

Vibration Analysis

All rotating machinery like fans pumps, motors, turbines and compressors will vibrate. The level of the vibration and the pattern of the vibration indicate the condition of internal rotating components.

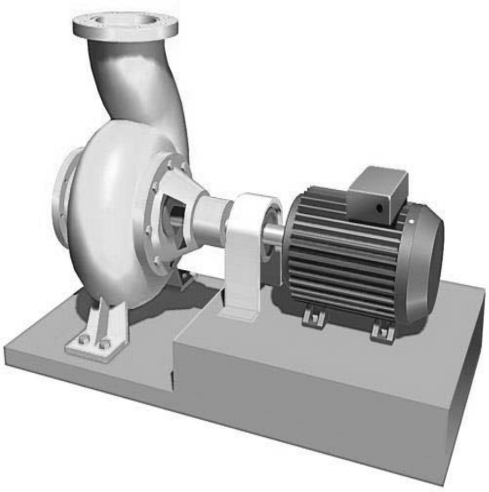


Figure A-4

If we use electronic instruments to measure the vibration, those levels can be monitored and the pattern studied. To a large extent, if the levels increase, and the patterns change we can not only detect that there is a problem, but we can diagnose the type of problem.

A number of different types of problems can be detected with vibration analysis. The vibration pattern can indicate a misalignment condition or an imbalance condition. The pattern can point to a rolling element bearing problem or a journal bearing problem.

Fault conditions detectable with vibration analysis include:

- Bearing problems – both journal and rolling element bearings
- Imbalance
- Misalignment
- Looseness
- Soft foot
- Electrical faults
- Eccentric rotors
- Belt and coupling problems
- Gear mesh
- Broken rotor bars

Figure A-5

Vibration analysis utilizes a special sensor mounted to a bearing housing that is sensitive to movement. A “snapshot” of the vibration is captured in a portable data collector and transferred to a computer for analysis. The snapshot data is generally collected on a monthly basis except for critical machinery which may have permanent sensors mounted for continuous monitoring.

The “snapshot” data is studied to determine whether a problem exists, and the severity of the problem.

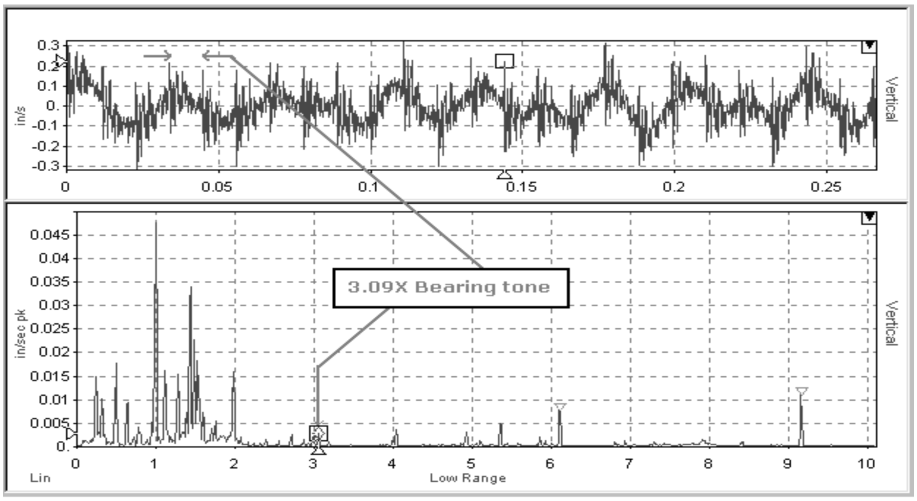


Figure A-6

Data is collected from more than one location, and in more than one direction. A machine vibrates up and down (vertically), side to side (horizontally), and end to end (axially), and different types of faults reveal themselves in different ways.

Sensors can be permanently mounted so that data can be collected at junction boxes in safe environments.

For a machine that is critical to the process, and machines located in remote or hazardous environment (such that routine measurements cannot be taken), sensors will be mounted permanently on the machine, and a monitoring system will monitor the vibration levels to give an early warning of a fault condition.

Online monitoring may be set up for the most critical machines. They will typically have a protection system installed so that if vibration levels reach predefined alarm levels, the machine will shut down. This type of system is often used for turbine/generators.

Four phases of Vibration Analysis

Detecting whether a problem exists
Diagnosing the severity
Performing a Root Cause Analysis to determine why it happened
Verifying the problem is resolved after repairs

Detection Phase: The first task of the vibration analyst is to *detect* whether the machine has a problem. Most vibration analysis software programs can assist in this step. Unless an expert system is used, the software compares the vibration levels to a set of preset alarm limits. It then becomes the user's job to go through the "exception report" to find out which machines possibly have a problem. This begins the Analysis Phase.

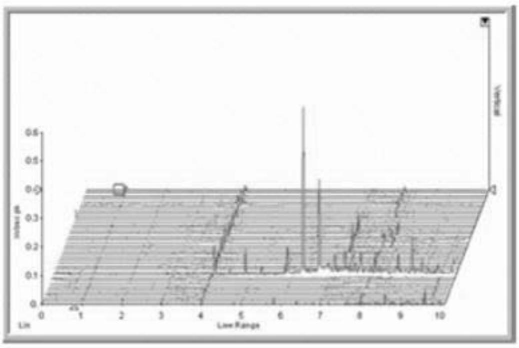


Figure A-7 Comparing historical data can provide valuable information to the analyst.

Analysis Phase: It is possible to determine what kind of problem exists and how severe it is by studying the spectral and waveform data. Then a report is generated along with recommendations. The maintenance group uses the report to schedule any recommended actions.

Root Cause Analysis Phase: The same set of tools are used to perform critical root cause analysis investigations. From these tests a determination can be made as to why the machine developed the problem in the first place. Additional recommendations are then issued to make changes that will prevent the problem from occurring again.

Verification Phase: Once the machine has been repaired, additional measurements are made to make sure the problem really is solved, that the repair was sound, and that the machine is fit for continued operation.

Vibration analysis technology continues to grow in popularity because it can be used for most plant equipment and advances are continually being made through many companies that increase its capabilities while reducing the initial investment. The Return on Investment is one of the best of all the condition monitoring technologies.

Acoustic Emission (Airborne Ultrasond)

Rotating equipment and other plant assets emit high frequency sounds that provide clues to potential problems. Ultrasonics testing is a useful technology for a variety of applications.

Finding air leaks	Excellent
Finding bearing problems	Good
Finding steam leaks in steam traps	Good
Detecting lubrication problems	Good
Detecting electrical faults (arcs, coronas)	Good
Finding flow related problems in pipes and valves	Very limited success

Table A-1

How it works

The human ear can only detect sound in the sonic frequency range of 20 Hz to 20,000 Hz. Sounds above this range are referred to as “Ultrasonic”, meaning above human hearing capability.

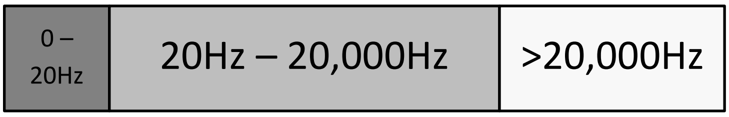


Figure A-8 Sound is categorized into 3 regions; Sub-sonic range, Sonic range, and Ultrasonic range.

The ultrasonic sensor is used to measure the signal and heterodyne (demodulate) it to a frequency range within the human hearing range.

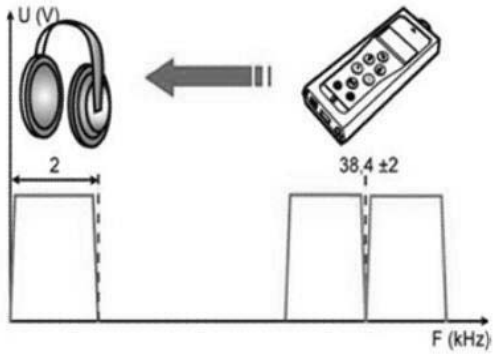


Figure A-9 Ultrasonics high frequency sounds are converted to an audible range

A few considerations in using the ultrasonics technology:

- The sounds are directional and the sensor in the "gun" is very directional. Therefore the sources of the sounds can be pinpointed (but may easily be missed).
- The volume depends on the distance from the source.
- Sounds travel through air, liquid, and solid objects, but not through a vacuum.
- There are two modes of detecting ultrasonic sounds – airborne, and direct contact. The typical airborne sensor or "gun" has an open end which is good for scanning areas in close proximity. For distances, a concave dish similar in style to a satellite dish is used. The dish reflects the sound to the sensor in the center.

Figure A-10 A "dish" captures sound from a distance

A typical "direct contact" system makes contact with the surface of the equipment to provide better transmission of the high frequency sound.

Ultrasound data can be listened to via headphones, which is very useful when searching for leaks and detecting faults in noisy environments.

Figure A-11 A Direct Contact probe uses a "stinger" to capture the sounds

Note: The sound can also be measured and displayed in db units in waveform and spectra displays.

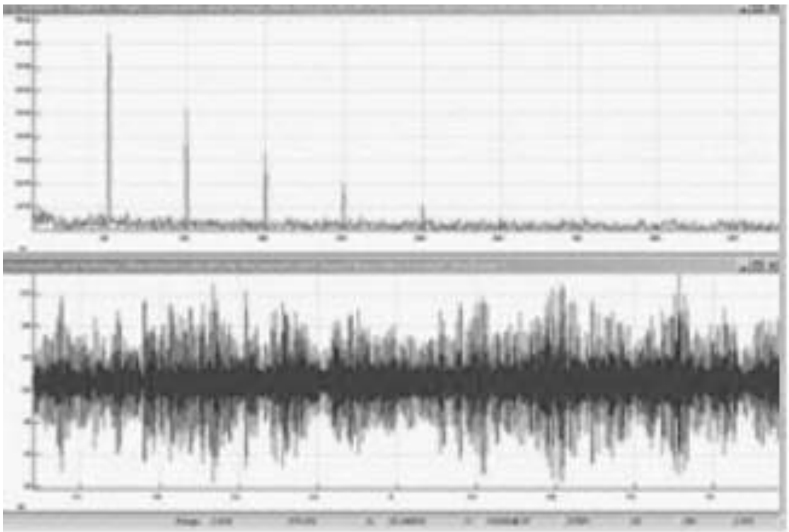


Figure A-12

Air Leaks

Air leaks are the most expensive utility leaks in manufacturing. Turbulence from leaks creates white noise with a strong ultrasonic component. Ultrasonic instruments pinpoint pressurized gas and vacuum leaks regardless of ambient background noise. Finding and correcting leaks can save hundreds of thousands of dollars annually.

A compressed air maintenance program includes a complete inspection of the airlines three to four times a year. Leaks are tagged when detected so that action can be taken. Regular inspections ensure that new leaks are found and confirm that tagged leaks from previous surveys were repaired.

Know the system – familiarize everyone with the supply side, the demand side, the number of compressors, operating pressures and any additions to the system since it was installed. Be sure the equipment user is properly certified and trained by a reputable trainer. Find, tag, repair, and re-check the leak area with the ultrasonic detector. Remember that the person making the repair may not be the same person who tagged it. It is also possible that a new leak was created while the old one was being repaired. Document everything.

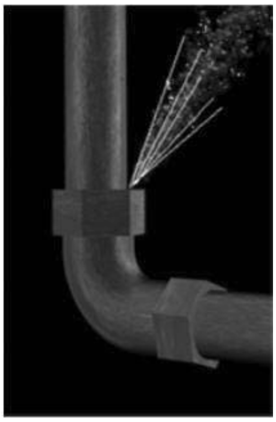


Figure A-13 Air Leaks emit high frequency

Boiler, Heat Exchanger, and Condenser Leaks

Scan for external pressure or vacuum leaks in Boilers and Heat Exchangers and condensers All pipe connections, flanges, seals, and access doors should be inspected as part of regular PM's. Listen for the same sound that is associated with compressed gas and air leaks. Tube leaks in condensers and heat exchangers can be checked using either the pressure method, the vacuum method, or the bisonic transmitter method. Choose the method that suits your application best to save valuable inspection time.

Sweep the microphone across the area to detect the leak. Sweep side to side and up and down.

Detecting Faulty Steam Traps

Steam traps open occasionally. This opening and closing produces a sound that is very distinct from the normal flow. A direct contact probe is ideally suited for listening for correct operation. There should be a steady sound interrupted by the distinct opening and closing of the valve. If the sound is a continuous fluttering sound, it is not operating correctly. A consistent rushing sound indicates the trap is stuck in the open position.

If the steam trap is stuck open, it wastes hundreds or thousands of dollars. It generates excess steam in the system which creates back pressure and causes failure in other steam traps.

Figure A-14 Checking a steam trap with a Direct Contact probe

When the steam trap is stuck shut, it produces a water hammer sound. In cold climates the line may freeze causing the piping or trap to rupture.

There are four common types of steam traps, and all work to remove impurities from the steam system.

The four types are known as:

- Inverted Bucket
- Float and Thermostatic
- Thermostatic
- Thermodynamic or Disk

Traps work on one of three operative modes.

- Change in Density
- Change in Temperature
- Change in Velocity

Detecting Electrical Problems

Ultrasonic monitoring can be used to detect arcing, nuisance corona, tracking, and line bushing conditions such as may be found in:

- Motor control centers
- Breaker Panels
- Power Lines
- Connections
- Insulation breakdown

Figure A-15 Ultrasonics is a good choice for monitoring distribution lines

Ultrasonics is one of the top choices for electrical inspections of transmission and distribution lines, Substation inspections, Switch gear, Transformers, Corona, Arcing and Tracking, Radio and television interference faults. Sounds associated with frying, popping, buzzing, and humming are characteristic of these faults.

Bearing Faults and Lubrication

Rolling element bearings produce ultrasonic friction as the internal rolling elements turn against the raceway. Likewise, the friction absorbing properties of grease means that a well-lubricated bearing produces less friction than a bearing that lacks lubrication.

True digital RMS readings provide accurate and reliable advanced early warning of impending failure in rotating equipment. Acoustic Vibration Monitoring trends high frequency bearing energy to determine proper lubrication intervals and indicates when the bearing is entering its first stages of wear.

Over lubricating bearings can damages seals, build internal pressures on the bearing, and cause premature failure. Over greasing an electric motor can push lubricant into the windings causing shorts and more severe damage.

Figure A-16 Direct Contact probe is useful for detecting bearing problems

Under lubricating the bearing negatively impacts the lifespan of rotating machinery also. A high percentage of bearing failure is due to incorrect or insufficient lubrication which generates frequencies above 30 kHz. Bearing impacts can be heard and dB levels trended.

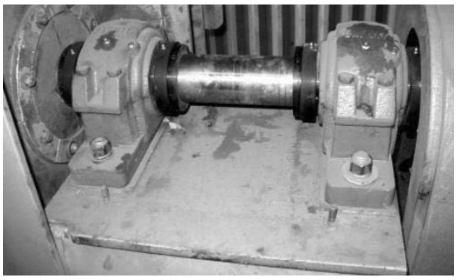


Figure A-17 Under lubrication generates Ultrasonic energy

Ultrasonic technology is used to let the bearing indicate the amount of grease lubricant to add. The sound can be monitored through headphones as the grease is pumped into the bearing. The bearing noise reduces substantially as the grease reaches the bearing. The waveform from the bearing during lubrication clearly reveals the improvement due to the lubrication. Training is required to ensure that this method is used correctly.

Figure A-18 Capture of Ultrasonic energy during lubrication

Mechanical Fault Detection

Ultrasonics can be used to detect early bearing wear in rotating equipment such as Gearboxes, pumps, motors, and compressors. The values can be trended for changes which can be an early warning indicator.

Ultrasonic monitoring should be an integral part of the condition monitoring program

Benefits:

- Ultrasonic equipment is not expensive and is not difficult to operate.
- It is often used to detect a fault while other technologies are used to follow-up and determine the severity.
- It is very direction which enables the pinpointing of particular faults, especially leaks.
- It is useful in high noise environments.
- Can be used in peak production hours
- Integrates with other predictive maintenance (PdM) technologies

Infrared Thermography

Infrared Thermography is the study of *radiated energy* using a thermal infrared imaging system.

Thermography is a popular technology applied to rotating and non-moving equipment in the plant. It involves the study of temperature as increased wear, steam leaks, and electrical arcing (to name but a few conditions) result in a change in temperature.

Excessive heat is an indicator of problems or potential problems in plant equipment including moving and stationary parts and equipment such as electrical panels, boilers, transformers, and electrical power transmission conductors, insulators and switchgear. Infrared Thermography is an ideal, non-intrusive technology for detecting these problems.

The technology uses sensors that are sensitive to the radiated electromagnetic energy associated with heat. The device translates the detected level of radiated energy into a temperature based on information entered by the user. Two types of devices are commonly used in our industry: spot radiometers and infrared cameras.

Spot radiometers sum up the energy in a small area and display a temperature reading. Radiometers often use a laser beam to help you to target where the measurement is to be taken. It should be understood that this reading is not based purely on the temperature at that point; the further you are from the target, the larger the area used to determine the temperature. You should visualize a cone radiating from the device – the greater the distance to the target, the greater the measurement area.

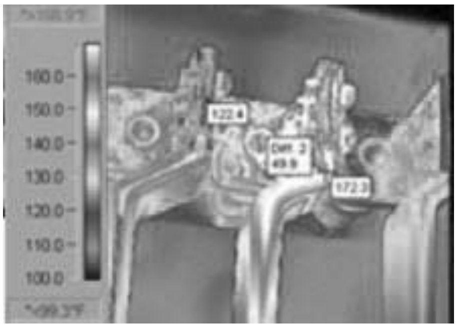


Figure A-19 Thermography image of bus bar connections. White area is the hottest.

It is important to understand that the actual temperature indicated may be incorrect as surface type, air flow, and other factors affect the accuracy.

An infrared camera can create a thermographic image. The instrument used to "see" radiated electromagnetic energy is generally referred to as a camera. These cameras use special sensors to detect the heat which is displayed in a visual image similar to a photograph. Many thermographic cameras also have a standard photographic camera built in so that a photographic image can be compared to the thermographic image. Most thermography systems include a software program for transferring the images to a computer for analysis and printing reports.

Note: the infrared camera does not "see" temperature. Temperature is calculated from inputs by the user in the camera or software.

Figure A-20 Infrared Thermography instruments include Spot Radiometers, still cameras, and movie cameras.

The instruments range from a "spot radiometer" used to detect temperature in an area, to still cameras and movie cameras which can record the changing temperatures. The price ranges from inexpensive to expensive with the more expensive models typically having a higher resolution image, an ability to zoom in to the area of interest, more field options, and additional software options.

The typical infrared camera has the ability to adjust the sensitivity so that the color scale shows the hottest area as "white hot." This "white hot" area could be 50 degrees f to several hundred degrees depending on the application. **The primary usefulness in most applications is the relative temperature rather than the absolute temperature.**

Figure A-21 Thermographic and photographic image of overheated bearing

It is possible to adjust the image to make any part of it glow "white hot." A scale indicates the relationship between color and temperature. Although the scale indicates temperature, comparing temperature at selected points provides the best indication of severity. Temperatures are calculated based on inputs the user provides to the camera.

Infrared Thermography is typically used in the following applications:

- Mechanical
- Machines, pipes, bearings, belts
- Electrical
- Overhead lines, transformers motors, control panels
- Steam Systems
- Piping, steam traps
- Refractory plant

A few things to know…

The infrared camera looks at **radiated electromagnetic** energy from the first 1/1000" (one mil) of the surface.

The infrared camera does NOT "see" temperature. Temperatures are calculated from inputs the user provides to the camera or computer.

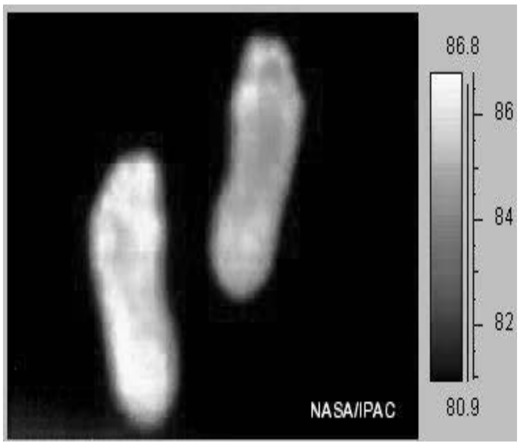


Figure A-22 IR image of footprint after standing a few seconds

Heat Transfer

Heat is transferred in three ways. These can affect the accuracy of the temperature reading of the object.

Radiation or Emission – is the type of heat transfer where the heat is transferred from the surface of an object as an infrared energy.

Conduction - A type of heat transfer mainly through a solid object

Convection - A type of heat transfer where the heat is transferred by the heated part of gas or liquid moving upward.

See Figure A-23. These issues must be understood to correctly measure temperature.

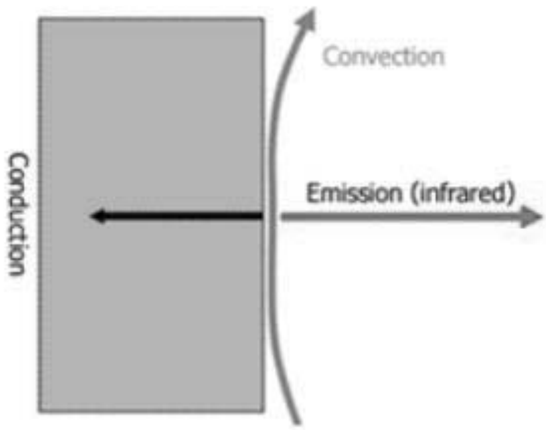


Figure A-23 Heat is transferred via conduction, convection, and radiation (emission)

Emissivity - How does it work?

The infrared camera detects 3 forms of radiation:

- Emitted Heat
- Reflected Heat
- Transmitted Heat

These three must be understood and considered in order to obtain accurate temperatures. All objects emit, reflect, and transmit heat differently. Therefore it is important to know the emissivity rating of an object.

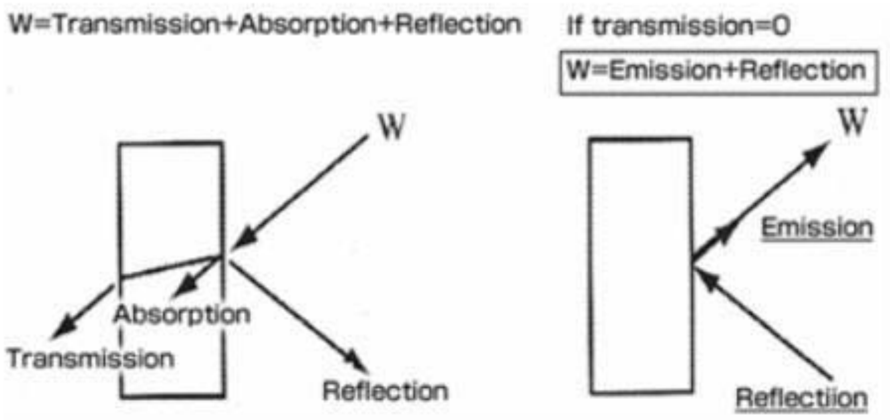


Figure A-24 Relationship of Emitted Heat, Reflected Heat, and Transmitted Heat

Emitted Heat – Emissivity

Emitted heat is the heat radiated from an object as infrared energy.

- The 'black body' has an emissivity value of 1.0. The black body has an anodized black coating.
- All other bodies have an emissivity value of less than 1.0.
- All objects emit, radiate, and transmit heat differently; therefore it is important to know an object's emissivity.
- Examples of high emissivity objects are water, snow, ice, vegetation, glass, paper, soil and minerals. Oxidized copper = .0.68, polished copper = 0.02
- Examples of low emissivity objects (.2 and below) are polished metals (copper, aluminum, steel, silver, chrome)
- Tables of emissivity values for various objects are available for inputting emissivity values into the camera or computer.

> *Note: Emissivity is the term used to describe the ratio of the energy emitted by an object compared to the energy emitted by a 'black body' at the same temperature.*

The infrared camera looks at radiated electromagnetic energy from the first 1/1000" of the surface. The source of heat may come from inside the object (tank, machine, cabinet), but the detector only sees the energy emitted from the surface. Therefore it is necessary to open cabinets to detect electrical faults – unless special windows are used or the cabinet door gets hot.

The content and color of paint can affect the emissivity.

Figure A-25 The paper is all the same temperature. The color affects the Emissivity.

Reflected Heat

Heat can be from a source other than the object being measured and may drastically affect the accuracy. Other sources could be intense light, sunlight, heat from a nearby object or machine. The reflected heat could be at such an angle that the camera is looking directly at the reflected heat rather than the heat actually emitted from the object. For example, an infrared image of a shiny tank on a sunny day may reflect sunlight directly into the lens and thereby affect the accuracy of the reading.

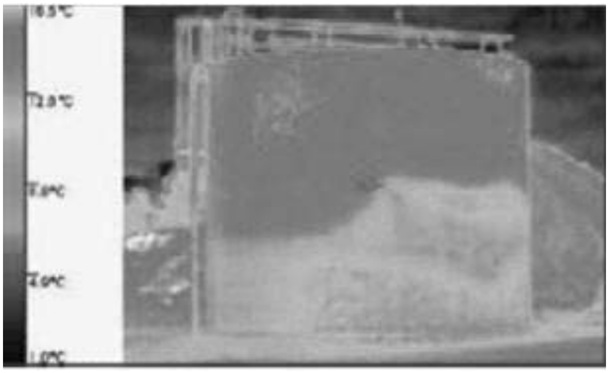


Figure A-26 Storage tank in sunlight

Transmitted Heat

Heat from another source can be absorbed into an object and actually transmitted out again. This affects the actual true temperature and the radiated electromagnetic energy

Generating Reliable Measurements

To generate reliable measurements:

- The emissivity must be greater than 0.6
- The background must be known or controlled
- There must be controlled or no wind

Unreliable measurements are generated when:

- Emissivity is less than 0.6
- Radiation is emitted poorly
- Reflects background radiation
- Extreme, unknown or changing background

Warning: If there is wind, even relative temperatures will be inaccurate.

Electrical Applications

Infrared Thermography is especially useful in many electrical applications. In substations and distribution lines watch for influences of sun and wind.

Loose connections and overloading can be detected.

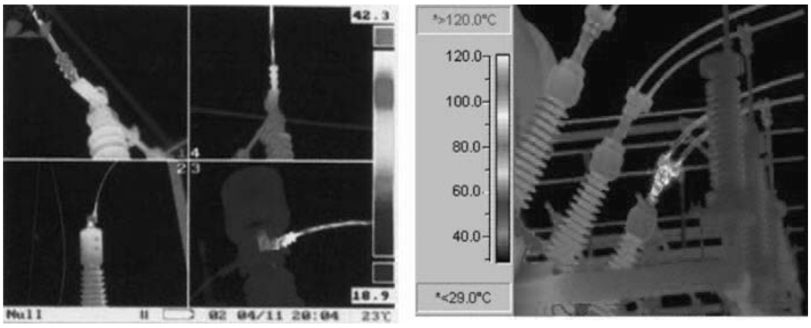


Figure A-27 Loose connections and overloading can be detected. Overcome the effects of sun by scanning at night.

Bus Bar Connections may be loose causing excessive heat, or the loads may be unbalanced. These conditions can be detected when the panel is open or there is a window in the panel cover. See Figure A-28.

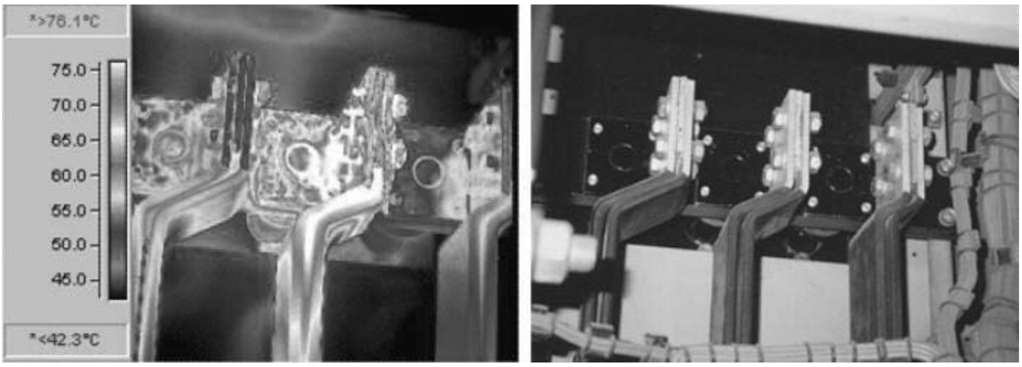


Figure A-28 Loose or overloaded connection.

Fuses and fuse blocks can have similar problems which can be detected. Candy striping due to broken strands is very apparent using a thermographic image.

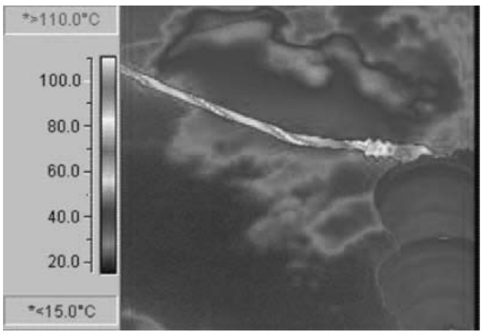


Figure A-29 Candy striping due to broken strand. Light colored strand is carrying the load.

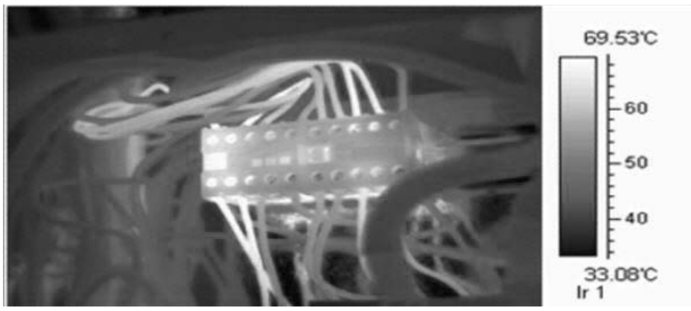


Figure A-30 Hot terminal block

Mechanical Applications

Many mechanical conditions can be detected using Thermographic imaging including:

- winding problems in motors See Figure A-32
- cooling issues,
- belt problems
- overheated bearings See Figure A-32
- abnormalities in pumps, pipes, and compressors

Note: IR is not a good early-warning indicator of bearing wear.

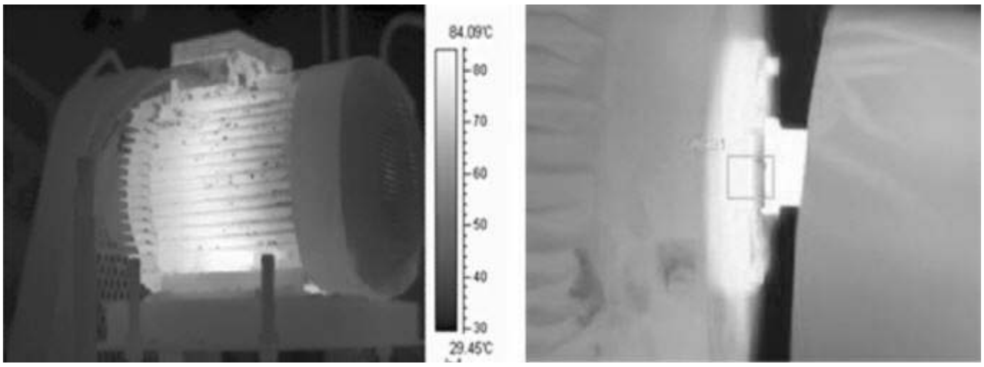


Figure A-31 Thermographic and Photographic image of failed steam trap. Note the delta temperature of 20 degrees.

Steam Systems

Leaks and blockages in steam systems are easily seen with Thermographic imaging. It is also a good way to check Steam Traps to verify whether they are functioning properly. One key is to check the relative temperature across the trap from the inlet to the outlet. Figure A-32 shows a delta temperature of 20 degrees.

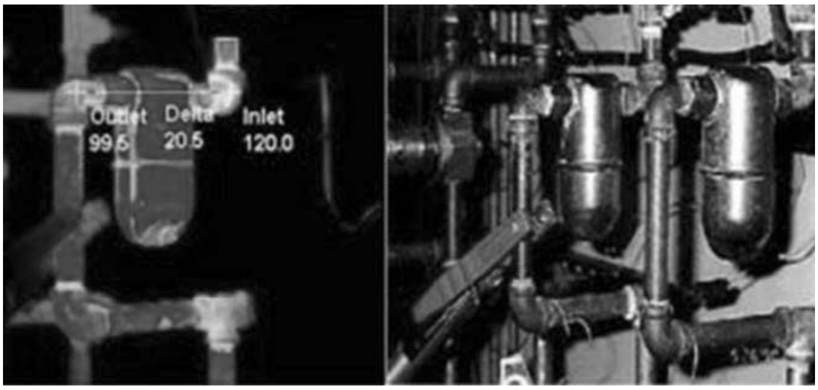


Figure A-32 Overheated winding in motor; Overheated bearing

Refractory Plant

Boilers, kilns, and other refractory equipment often have insulation problems that are detectable with Thermographic imaging. It can also provide and indicator of blockages or buildup that may prevent the system from functioning at an optimum. Hot areas in the lime kiln in indicate areas of heat loss which affect the process efficiency.

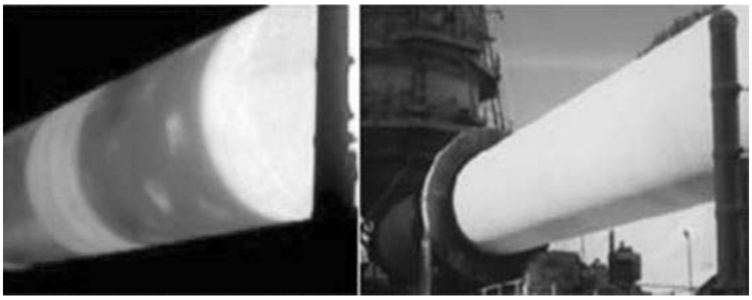


Figure A-33 Lime kiln with hot areas indicating poor insulation.

Infrared Thermography is a technology that has many applications in and around a plant. It has useful applications in processes, in support equipment and systems such as electrical power, piping, tank levels, leaks, and more. The price of IR equipment has a wide range so that it is affordable to most plants.

Electric Motor Testing

Electric Motors are the main equipment component in most plants. It is imperative to know their condition in order to ensure uninterrupted processes and to schedule downtimes rather than have surprise downtimes.

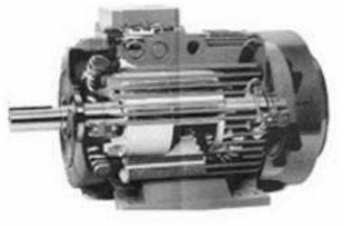


Figure A-34

Mechanical problems of motors can be detected with Ultrasonics, Infrared Thermography, and Vibration monitoring. However, there are special tests that can detect the electrical condition. The tests fall into two categories:

1. Static / off-line tests
2. Dynamic on-line tests

Static / Off-line Tests

Static or off-line testing is usually performed once a year or during outages with the motor shut down. Off-line testing is also used as a quality assurance tool when first receiving reconditioned or rewound motors from the motor shop before they are stored or returned to service. Testing these incoming motors provides proof the motor shop is doing its job properly and becomes the new base-line for future trending. Off-line equipment can also be used as a troubleshooting tool. Any time a problem has occurred the motor involved should be tested for insulation integrity. Overload situations, contamination issues and voltage problems can compromise the insulation.

Off-line testing includes:

- winding resistance
- meg-ohm
- polarization index
- high potential
- surge testing.

The tests should be performed in that sequence with modern, state-of-the-art test equipment. Equipment is manufactured today that can adequately reproduce “real world” experiences without causing damage to the motor’s insulation system. It is important to test motors at voltage levels and conditions they will see in their normal, day-to-day operation.

Motor Circuit Analysis - MCA Traditional Test Methods. - Most of the traditional test methods require a significant voltage application in order to work. The purpose is to stress the insulation system by forcing a reaction of the insulation dipoles or to force a potential across a resistive or capacitive fault.

Meg-ohm or Insulation to ground testing - Polarization Index; Resistance Testing; and, Surge comparison tests. Insulation to Ground Testing (Meg-Ohm meters)

PdMA – MCE - Static

Figure A-35

A DC potential is placed across the motor winding conductors and ground. The applied potential is set and a value of current (leakage) crosses the insulation boundary. This value is converted to resistance, usually in meg-ohms. It is, in effect, a method of measuring leakage across the insulating boundary, but only between the surfaces of the conductors and ground. As the insulation dipoles are only excited with DC, some time is required for them to polarize. Standards normally indicate a winding charging time of about 1 minute and, as insulation resistance is directly affected by temperature and moisture, normalization for temperature.

Polarization Index - The polarization index (PI) test is a measurement of leakage at one minute then at ten minutes. The results are shown as a ratio of the ten minute to one minute reading. It is assumed that a fault will polarize slowly (high ratio) or rapidly (low ratio) due to contamination and changes to the circuit capacitance.

High Pot or High Potential Resistance Testing - Resistance tests use a low voltage DC output and a bridge. The primary purpose is to detect high resistant joints, loose connections, broken connections (or conductors) and direct shorts.

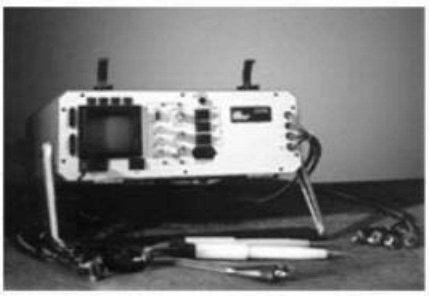


Figure A-36 Baker Instrument used for Static tests

Surge Comparison Testing - This is an older method of evaluating windings for shorts. A series of steep-fronted higher voltage pulses are sent from the instrument to the stator. The higher voltages force the dipoles in one direction leaving the ability to detect a reactive fault as creating enough potential to cross the barrier (Paschens law) either being shut down after partial discharge occurs or an arc is drawn. Both methods of detecting cause a change to the properties of the insulation at the point of defect either accelerating the fault or completing the fault. In order to force slight defects, a greater potential must be applied, stressing the complete insulation system. Due to the steep fronted pulses, the applied voltage is normally impressed only on the first 2-3 turns in the first coil of each phase.

The situation is quite different for detecting the breakdown of the turn insulation in a winding (parallel or phase) having many coils. The breakdown of the turn insulation in a single coil in a winding of many coils produces a very small relative change in the characteristics (L, C, R) of total load impedance seen by the surge generator. Hence the change in the VFW [voltage wave form] shape produced by the breakdown of the turn insulation somewhere in a winding of many coils is relatively very small. Therefore the surge tests may not reliably verify the presence of one shorted turn in a single phase winding or three phase winding in a machine.

“The surge tests on windings in a machine may possibly lead to wrong conclusions. Perfectly intact windings may appear to have a turn short. More importantly, a turn short induced by the surge test by breaking down the weakened turn insulation may not be detected. In such a case, the stator winding would likely fail after the machine is put back into service. “In view of the above facts, caution is advised in surge testing of the turn insulation in complete windings. These tests carry very significant risks, which should be carefully considered. Such caution is more important for diagnostic tests on machines in service as such tests are carried out quite infrequently in contrast to frequent tests on new, or refurbished, or repaired machines in a manufacturer’s plant.”[8] As shown, traditional testing has specific flaws in the ability to detect, and the ability to detect defects in a non-destructive manner.

Dynamic On-line Tests

On-line tests enable testing at the motor and at the panel while the motor is in service. The tests view the current and voltage spectra depending on the test. The data is treated like vibration data.

Figure A-37 Dynamic On-line Tests can be done at the electrical panel.

Online Tests that can be performed include:

- Winding shorts between conductors or coils
- Winding contamination
- Insulation to ground faults
- Air gap faults, including eccentric rotors
- Rotor faults including casting voids and broken rotor bars.
- Vibration which detects broken rotor bars, air gap eccentricity, eccentric rotor
- Current Analysis with a current clamp – broken rotor bars
- Flux Coil – uneven flux field.

Current spectra – Current spectra can be collected on each power leg at the panel using a current clamp. See Figure A-38. One general comparison that can be made is the amperage draw from each leg. These should be approximately the same for proper load balancing. An exceptionally high or low load indicates problems.

Figure A-38 Current Clamp

The current spectrum can indicate the probability of broken rotor bars and other defects such as uneven air gap or a bowed rotor.

The flux coil is used on the motor rather than at the panel. It collects data of the flux field generated by the motor. The data can be viewed as a spectrum. It can indicate potential problems in the windings.

Figure A-39 Flux Coil used at the motor

The various on-line tests tend to be less destructive than off-line tests and have the added benefit of keeping the motor in service and avoiding starts and stops. The on-line tests can typically be performed by certified in-house personnel while many of the off-line tests are normally performed by outside personnel such as a motor shop.

Oil Analysis

Oil is the life-blood of rotating equipment. Rotating machinery needs correct lubrication. But it is surprising how often the incorrect lubricant is used, or the lubricant is contaminated. The result is increased wear and equipment failure. There is also an economic issue - the lubricant is expensive, both to purchase and dispose of.

Figure A-40

Too often perfectly good lubricant is changed out, at great expense. So testing is performed on the oil and grease. The tests indicate:

- whether the lubricant is still able to perform its job (is the additive pack OK, etc.)
- whether there are any contaminants such as water or dirt
- whether there are any metals or other elements, which may give an early warning of wear

Oil Tests and what they measure

Samples are collected routinely for analysis. They may be sent to an outside lab or an in-house lab. Various tests on the oil include:

Test	Measures…
Oil Bath 40c and 100c	Viscosity
R. D. E. Spectroscopy	Elemental Concentrations
FT – IR (Infrared)	Degradation, contamination, additive depletion
Total Acid	Acid Levels
Total Base	Base Levels
Water Crackle Karl Fisher	 Concentrations to 200ppm Concentrations to 10ppm
Particle Count	NAS & ISO Cleanliness

Table A-2

Strengths of Oil Analysis

- Detects normal wear particles up to 6-10 microns.
- Determines lubricant additive depletion
- Detects fluid contamination

Weaknesses of Oil Analysis

Does not detect the onset of abnormal wear – wear particles in excess of 10 microns
Does not detect the sources of wear (bearings, gears, seals, rings, etc)
Does not provide information regarding machine condition.

Note that these tests do not provide all the information that may be needed for detecting abnormal wear inside the machine. The particle count only recognizes particles up to 6 microns in size and ignores larger ones. Neither does it indicate the sources of wear, i.e. bearings, gears, seals, rings, etc.

Oil analysis provides good information about the oil condition. It provides little information regarding the machine condition. It is easy to assume that a positive oil report means the machine is in good condition. That is not necessarily true. It is important to recognize the limitations of oil analysis and couple it with Wear Particle Analysis for an accurate picture of the true condition of the machine.

Wear Particle Analysis

Ferrographic wear particle analysis is a machine condition analysis technology that is applied to lubricated equipment. It provides an accurate insight into the condition of a machine's lubricated components by examining particles suspended in the lubricant.

By trending the size, concentration, shape, and composition of particles contained in systematically collected oil samples, abnormal wear-related conditions can be identified at an early stage.

Figure A-41

Wear particle analysis complements vibration analysis by providing, in some cases, earlier fault detection and is less susceptible to the limitations imposed by slowly rotating or reciprocating machinery.

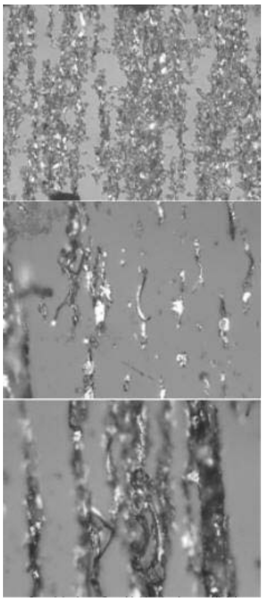


Figure A-42 The different wear particle types indicate specific problems

Although it is possible to purchase laboratory equipment and perform tests in-house, most industries rely on external commercial laboratories for their testing. The oil samples must still be collected on-site in a controlled manner, but they are then sent off-site to a laboratory. Test results are typically available electronically for integration into the condition monitoring program.

Wear Particles are typically divided into six types. They each have particular characteristics and causes.

Oil Analysis vs. Wear Particle Analysis

Figure A-43 shows the relationship between standard oil analysis capabilities using Spectroscopy and Wear Particle Analysis. Spectroscopy only recognizes particles up to 6 microns and ignores the larger abnormal wear particles. For this reason it is a good practice to have Wear Particle analysis performed on oil samples.

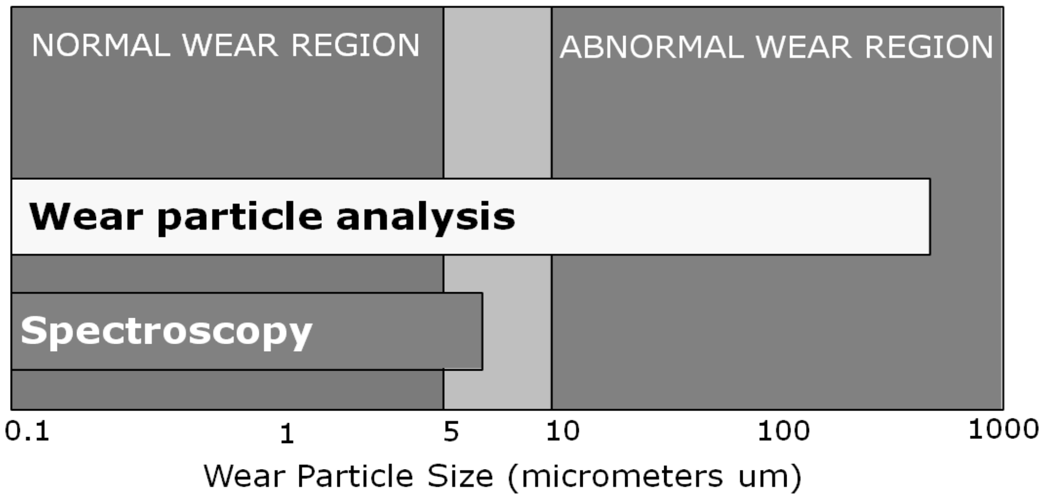


Figure A-43 Normal oil analysis does not see abnormal wear particles.

Abrasive Wear

Abrasive Wear is the result of hard particles coming in contact with internal components. Such particles include dirt and a variety of wear metals. Introducing a filtration process can reduce abrasive wear. It is also important to ensure vents, breathers, and seals are working properly.

Adhesive Wear

Adhesive Wear is generated when two metal surfaces come in contact allowing particles to break away from the components. Insufficient lubrication or lubricant contamination normally causes this. Ensuring the proper viscosity grade lubricant is used can reduce adhesive wear. Reducing contamination in the oil will also help eliminate adhesive wear. Cavitation occurs when entrained air or gas bubbles collapse. When the collapse occurs against the surface of internal components, cracks and pits can be formed. Controlling foaming characteristics of oil with an anti-foam additive can help reduce cavitation.

Corrosive Wear

Corrosive Wear is caused by a chemical reaction that actually removes material from a component surface. Corrosion can be a direct result of acidic oxidation. A random electrical current can also cause corrosion. Electrical current corrosion results in welding and pitting of the wear surface. The presence of water or combustion products can promote corrosive wear.

Cutting Wear

Cutting Wear can be caused when an abrasive particle has imbedded itself in a soft surface. Equipment imbalance or misalignment can contribute to cutting wear. Proper filtration and equipment maintenance is imperative to reducing cutting wear.

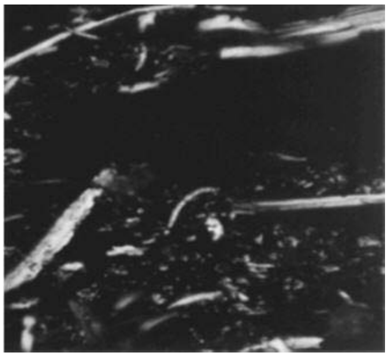


Figure A-44 Cutting wear

Cutting wear appears as long, curly strips of material with aspect ratios ranging from 5:1 to 50:1 (length to width.) It is never considered to be normal.

Fatigue Wear

Fatigue Wear results when cracks develop in the component surface allowing the generation and removal of particles. Leading causes of fatigue wear include insufficient lubrication, lubricant contamination, and component fatigue.

Sliding Wear

Sliding Wear is caused by equipment stress. Subjecting equipment to excessive speeds or loads can result in sliding wear. The excess heat in an overload situation weakens the lubricant and can result in metal-to-metal contact. When a moving part comes in contact with a stationary part sliding wear becomes an issue.

The following table is a useful tool in finding possible sources of wear particles.

Wear Metal	**Possible Origin**
Aluminum	Bearings, Blocks, Blowers, Bushings, Clutches, Pistons, Pumps, Rotors, Washers
Chromium	Bearings, Pumps, Rings, Rods
Copper	Bearings, Bushings, Clutches, Pistons, Pumps, Washers
Iron	Bearings, Blocks, Crankshafts, Cylinders, Discs, Gears, Pistons, Pumps, Shafts
Lead	Bearings
Nickel	Bearings, Shafts, Valves
Silver	Bearings, Bushings, Solder
Tin	Bearings, Bushings, Pistons

Table A-3

Wear particle analysis is a powerful tool for non-intrusive examination of the oil-wetted parts of a machine. It can detect particles from 1 micron to 350 microns. The analysis considers the particle shape, composition, size distribution, and concentration. The results aid in determining operating wear modes within the machine, resulting in specific maintenance recommendations. Wear Particle Analysis detects abnormal wear. The standard oil analysis detects normal wear particles up to 6 microns.

Figure A-45

Selecting the Best Technology

Before selecting the technology it is best to step back and perform a review of all the plant equipment. Issues that must be considered are:

- reliability requirement
- the importance to the process
- whether there is redundant equipment
- physical accessibility and location
- hazards.

All of these issues are financial issues. Everything must be justifiable financially. If the time and effort required to monitor a machine cannot be justified, don't monitor the machine.

Figure A-46

Consider the **history of the machine**. Consider its reliability and failure modes along with ways to detect the failure modes. What is the impact of failure? Will there be losses due to downtime and secondary damage? What are the associated costs of parts and labor? What are the production downtime costs and capital costs involved? Is there a spare unit?

All these risks must be weighed against the likelihood of catching the fault, the costs involved with monitoring the equipment, and the costs of the technologies required – including the training needed to be successful.

A risk analysis should be performed. That analysis will help decide whether it will be monitored, and also how often and with which technologies. It will also help later on when a fault is detected: at what point should the repair be made?

The risk may change during the year. At times of high demand, or adverse weather conditions (summer for cooling plant, winter for power generation, for example), the monitoring frequency and repair plan may need to change.

Then consider the technologies required to detect future problems. Can you be successful with vibration? Should other technologies be included: infrared, wear particle analysis, etc.

Technology

Application	Vib	Lube	Wear	MCA	IR	US	Vis
Generator	☺	☺	☺	☹	☺	☺	☺
Turbine	☺	☺	☺	☹	☺	☺	☺
Pump	☺	☺	☺	☺	☺	☺	☺
Elec. motor	☺	☺	☺	☺	☺	☺	☺
Diesel eng.	☺	☺	☺	☹	☺	☺	☺
Fan	☺	☺	☺	☺	☺	☺	☺
Gearbox	☺	☺	☺	☹	☺	☺	☺
Cranes	☺	☺	☺	☺	☺	☺	☺
Elec. Circ.	☹	☹	☹	☺	☺	☺	☺
Transformer	☹	☺	☹	☺	☺	☺	☺

Table A-4

The following table shows which technologies are good for specific fault types.

	Vib	Lube	Wear	MCA	IR	US	Vis
Wear	☺	☹	☺	☹	☹	☺	😐
Heating	☺	☺	☺	☹	☺	☹	😐
Impact	☺	☹	☺	☹	☹	☺	😐
Corrosion	☹	☺	☺	☺	☺	☺	😐
Fatigue	☺	☺	☺	☹	☹	☹	😐

Table A-5 From Keith Young, paper in Maintenance Technology, June 1995

The Future of Condition Monitoring

Condition Monitoring has proven to be well worth the investment to implement. The future involves:

- tighter integration of the technologies and systems
- continued integration between condition monitoring and process monitoring
- increased use of automated diagnostic systems
- an emphasis of information over data
- greater use of the internet

Technology Integration - It is not possible to get the full picture from one technology alone. Analysts and technicians need to become multi-skilled rather than only dealing with a single technology.

MIMOSA (Machinery Information Management Open Systems Alliance) and other groups and vendors are continually working on standard for sharing data and information, however, more work must be done is this area. Groups like MIMOSA need greater support from vendors and users of the technology.

Merger between Process Management and Condition Monitoring - Large companies from the Process Monitoring and Control industries noticed the value of condition monitoring and have embraced the concepts and technologies. They are realizing the value of condition monitoring to the control of maintenance dollars and overall bottom line improvement to the organization. The line between condition monitoring and process monitoring will become blurred.

On-line Monitoring - With the popularity of the technologies and the value clearly seen for industry, advancements are rapidly being made in the area of On-line Monitoring. This application provides the capability for immediate detection and recognition of potential problems even in remote locations. On-line systems will become more sophisticated, less expensive, and smarter. Rather than transmit megabytes of data, diagnostic information will be transferred.

The Internet – Most plants today involve people and services that are located far from the equipment site. Many systems have automated email generation and the portable web browsers enable individuals to view data and make recommendations remotely. We can be sure that applications for condition monitoring will quickly follow the introduction of new advancements in communication.

Review

The Goal is to reduce plant operating costs (or increase profitability) by lowering maintenance costs and minimizing costly downtime. This is done by increasing equipment reliability, and regaining control of the maintenance schedule.

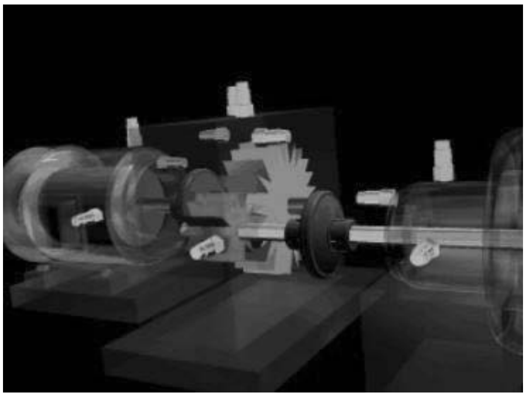


Figure A-47 Condition monitoring technologies enable us to look inside a machine and see its condition.

This goal will be achieved by making the required changes to ensure that equipment is more reliable. The first thing to do is determine why the equipment failed in the first place – high vibration levels inducing wear due to imbalance, misalignment and resonances; contaminated lubricants and improper lubrication; poor repair work; low quality parts; design and manufacturing problems and more. Condition monitoring technologies really do enable us to look inside a machine and see its condition.

The machines should be monitored frequently enough to detect problems in their infancy so that the most cost effective and timely repairs can be made. These goals can be achieved and are being achieved in varying degrees by industries around the world. It requires management commitment, proper training, and a high level of dedication.

Appendix B

Running a Successful Condition Monitoring Program

This article first introduces vibration monitoring concepts so that the neophyte will have a sense of the challenges involved. Then I discuss the need for training, steps required to get the program up and running, and what is takes to run a successful program. In preparation for this article I surveyed a number of my customers (users of the iLearnInteractive computer based training system). I asked them for their comments on training, and their experiences in starting and running a program. I share their comments in this article.

While the focus of this article is vibration monitoring, the issues raised and recommendations apply equally to all condition monitoring technologies.

Introduction

At first glance, it may seem easy. Just connect the sensor to the machine, press the record button, and then look at a graph on the computer screen. Those graphs tell you what's wrong with the machine – don't they…

Well, as anyone knows who has been involved in vibration analysis for any amount of time, it is just not that easy. Knowing how to take the measurement correctly, knowing where to place the sensor, knowing which machines to test, knowing what to do with the measurements, and knowing how to interpret the data is all quite challenging. Sometimes it can be very challenging – even for someone who has been doing it for a long time. That does not mean that you should back away and drop the vibration monitoring program; it just means that you need to be prepared.

And even if you are experienced in vibration analysis, does that ensure the success of the maintenance program? Ask yourself these questions: are recommendations being acted upon; are recommendations provided in a timely fashion (so that corrective action can be taken at a 'convenient' time); dos everyone believe in the recommendations; are other condition monitoring technologies being utilized; are improvements made to avoid failures in the future?

Quick review

First we need to *quickly* review how vibration monitoring works. In short, the sensor converts the vibration of the machine (which is due to the rotational, dynamic and frictional forces within the machine) into an electrical signal which can be measured by the electronic instrument – let's call it the "data collector".

Now, before we go too much further, we already have a couple of challenges. This sensor; what is it and how does it work? Do you need to know? Can't you always use the same sensor? And where do you place it on the machine? Does it matter? How is it attached to the machine? Does that matter? Well, of course, everything matters. All machines are not the same, so there are different kinds of sensors available, and where (and how) they are placed on the machine is very important.

Figure B-1

OK, we have a data collector that can record the electrical signal. Are all measurements the same? Well, of course you know that they are not all the same.

The simplest measurement is called the RMS or "overall level" measurement. It is a single number that reports the overall level of vibration. The idea is that if the level increases we must have something wrong with the machine. The trouble is that this reading is not sensitive to all fault conditions (the bearings may be failing and you may not know), and it does not give you any specific indication as to what is wrong with the machine – which can be helpful when planning the maintenance activity.

Figure B-2

There is another measurement called the Shock Pulse reading (and Spike Energy, etc.) which focuses on the state of the bearing (and lubrication). The RMS reading is more likely to increase if there is an out-of-balance, misalignment or looseness problem, so, together with the RMS reading, you have the most common fault conditions covered. Well, sort of. You should get a warning of impending doom, but you need more information.

Figure B-3

Instead, people use a time waveform and spectrum measurement. This is not the place to explain what these measurements are, but suffice to say that the waveform represents the signal that comes from the sensor and the spectrum (which is derived from the waveform) represents the amplitude level of all of the individual frequencies emanating from the machine.

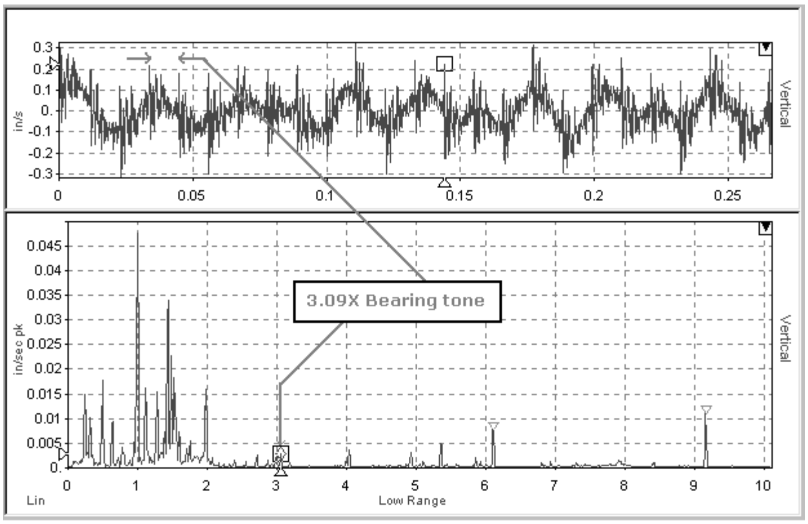


Figure B-4

The job of the vibration analyst is to look at the spectrum (and waveform) and try to determine what the patterns mean. They have to relate the pattern to what they know about the machine (its running speed, the gear ratios, the type of bearing, etc.). They have to look for changes in the pattern over time. Trust me, it is not easy.

The bigger picture

But let's step back and look at the bigger picture.

Once the system has been purchased the analyst has quite a challenge ahead. In fact, the vibration analyst is surrounded by different types of challenges.

First they have to look at the machines in the plant. Each machine is a potential trouble maker. Thankfully, through changes in vibration/sound, electrical properties, temperature, and lubricant characteristics, the machines are trying to warn of impending problems. If you read the signs correctly and deal with them appropriately, you can be a hero. Miss them, or misinterpret them, and you can look like a chump…

Figure B-5 Beware, they're out there somewhere

Second there are typically people in the plant who don't understand what this vibration monitoring equipment can do – so they fear it. Often there is an "us and them" environment. The analyst makes a recommendation to make a repair, and others insist that the machine is fine. And then there is the stand-off; one hoping the machine is fine, the other hoping the bearing is shot… This is not a great work environment.

And third there is the manager. The manager means well, but maybe he or she does not quite understand what the vibration program can achieve. Perhaps he or she expects that downtime will become a thing of the past. That puts unbelievable pressure on the vibration monitoring team/person.

So, what is the answer? I believe it is quite simple: you must understand the failure modes, you must have realistic expectations, you must create and follow a plan, and you must have on-going training.

Setting expectations

Everyone needs to have identical (and realistic) expectations.

If every person who maintains, lubricates, repairs, and operates the rotating machinery understood what the vibration monitoring technology can achieve, and chooses to help/cooperate rather than do nothing (or actually work against the program), then the situation would be much improved. Imagine if these people actually told the vibration team what they knew about the machine, and gave them a "heads up" when they noticed a change in vibration (audible) or operating state.

And if the maintenance management, operators, production and planning folk understood what can be achieved with vibration monitoring (the capabilities and limitations), then realistic expectations would be set. When recommendations were made, they would have greater confidence in the information, and would be able to put it to greatest use.

Instead, what tends to happen is that everyone outside the immediate monitoring group has little understanding of the capabilities of vibration analysis, and not only do they question recommendations, but when a machine does fail, blame is quickly focused on the vibration group.

The benefits of training

So how do we make it work? The first step is training – but not just for the vibration team – for everyone.

You only have to look at a few spectra to realize that vibration analysis is not easy. There is a lot to learn. You need to know how to operate the data collector and software, but you also have to know how to interpret the spectrum and waveform patterns. Extracting the information out of the data is very tricky and interpreting the information is even more difficult.

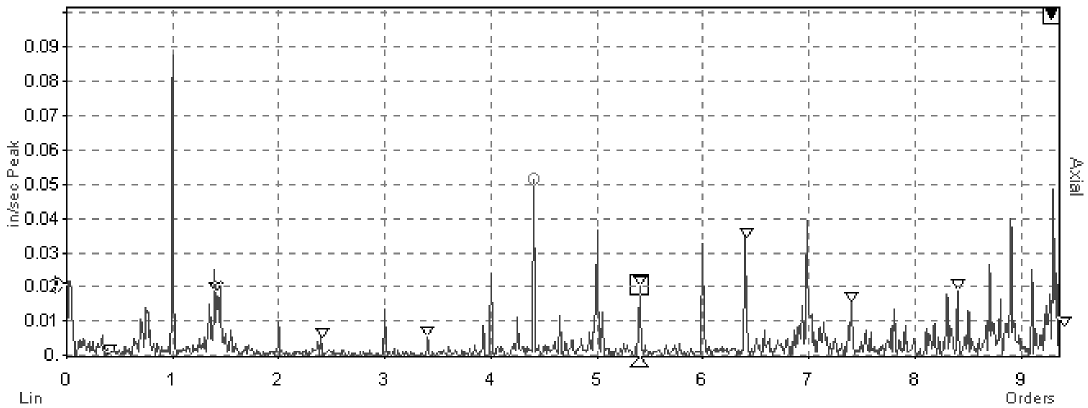


Figure B-6 A typical vibration spectrum – not easy to understand, is it?

Classroom training is very helpful. If the vibration technicians/analysts go to class off-site, they are able to ask questions of the instructor and the other attendees. On-site training is excellent because the instructor can potentially focus the discussion on the type of equipment at that plant, and the experience level of the team.

But "traditional" classroom training does have some limitations. Apart from the high cost of attendance (course fees, travel and accommodation), it is very difficult for most people to stay focused for such a long time. The instructor has his or her schedule of topics to cover, and the student is required to keep up. Most people are too embarrassed to ask questions. And if you dare to reflect in your mind on a topic just discussed, you will be missing what the instructor is presenting now.

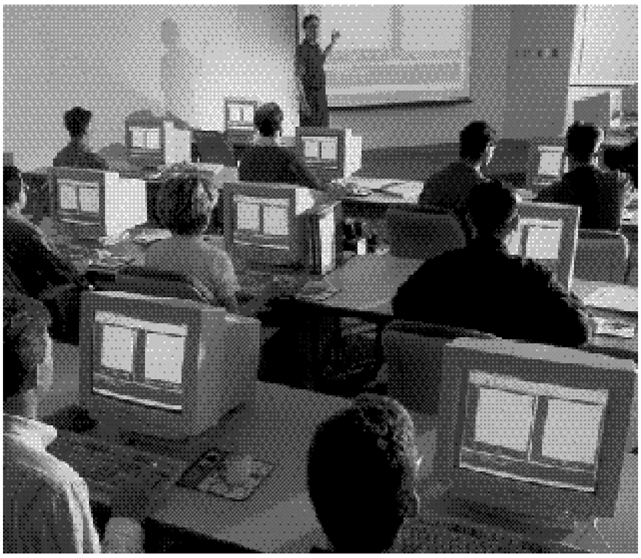


Figure B-7 Classroom training is necessary, but not the only answer

The fact is that by the end of the second day, most people have forgotten what they learned (or were told) on the first day. By the end of the third day, the first day is a blur. And after returning to their office (with a binder of course notes that will never be opened) and being confronted with a week of work to catch up on, the new knowledge quickly gets filed under "I wish I could remember that stuff".

The other fact is that most of what is "learned" never really gets embedded in your memory until you have to put it into practice. If you are lucky, the classroom course will provide some opportunities to analyze data and participate in interactive activities, but in general the important questions will arise *after* the course has ended when you are faced with a new situation.

And that is why computer based training (either via a Web site or CD-ROM based) is so valuable and effective. If you can take the training at your own pace, you are in control of what you learn. If a topic is already understood, you can fly through it. But if it is more challenging, you can go over it, time and time again. A good computer based system will also act as a reference system – when a new situation arises (a pattern in the data you don't recognize, for example), you can quickly turn to the reference and training material for assistance.

Note: Mobius Institute now offers classroom training – you receive a vibration training CD *before* you come to class so that you can prepare. During the class you use the simulators and take challenges so that you are more active and involved (and thus more interested). And you keep the CD which can be used as a source of refresher training and reference for years to come.

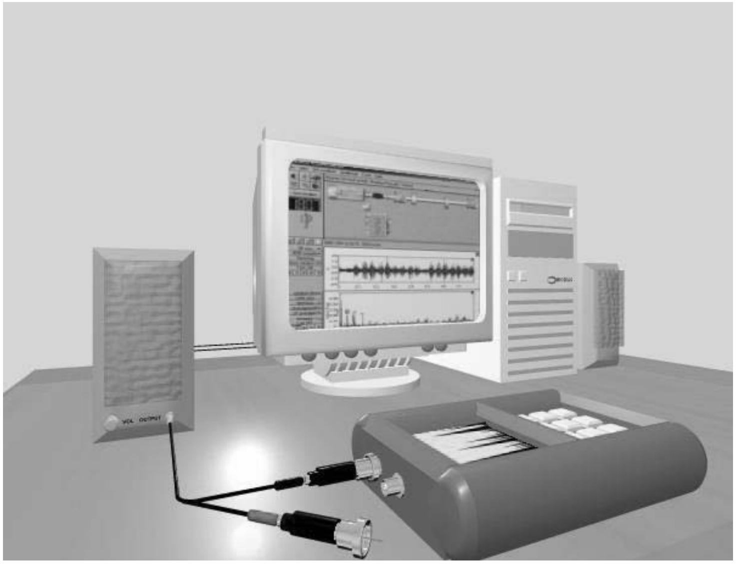


Figure B-8 Computer based training is interactive, self-paced, and available 24/7

But wait, so far we have only discussed training for the vibration monitoring staff. What about the rest of the maintenance staff?

I firmly believe that practically everyone needs training. No, the guy who lubricates the machine does not have to know how to analyze spectra, and the maintenance manager does not have to know how to diagnose bearing wear, however:

1. Everyone should know why we perform vibration monitoring. They need to know that if the condition of the machine changes, the vibration level and/or pattern *will* also change.
2. Everyone should also have an idea of the limitations of the technology. Everyone should know that vibration can be applied to certain types of machines, and can detect certain types of faults. They need to know that some machines can present real challenges (for example machines under varying speeds and loads), and that some fault conditions can develop too quickly to be caught by vibration analysis.
3. And they also need to understand the concept of predictive maintenance versus breakdown and scheduled maintenance. A good idea of reliability centered maintenance would help too.

The goal is:

1. When they see the vibration people collecting data, they know it benefits them (and their company).
2. When they witness a change in operating and maintenance state, they may like to tell the vibration team. They should volunteer relevant information.
3. When the vibration guys make a recommendation, whether it turns out to be right or wrong, everyone should know that they did so with the best intentions; using technology that gives a good, but not perfect, insight into the machine's condition.
4. When the budget becomes tight, management should still allow staff to receive training, and continue the program, even though there may not have been a spectacular save (or failure) recently.

While the vibration people need in-depth technical (and practical) training, and an excellent reference, the remainder of the maintenance staff, both the field workers and managers, need to have a training program suited to their needs. It may only require an hour a week for a few weeks. The benefits are significant.

Survey results

I have been developing a training product for the last six years, and I have now seen it implemented in a number of ways – from one extreme to the other. Recently I conducted a survey of users, in part so that I could write this article.

1. In some cases, people had implemented plant-wide training. Technicians, trades-people, and others, right on "up" to key managers were asked (or required) to view selected modules of the training. In some cases they viewed all of the information at a single sitting, in other cases they assigned staff to take the training in one hour sessions each week. Examples include General Motors, General Electric, Arizona Power, and International Paper.
2. In some cases the training was presented only to the condition monitoring group; so that no matter what technology was being used, they all understood vibration analysis.
3. In some cases, key staff would utilize the system to give their own internal classes to staff. In that way they could focus the training on issues relevant to their plant, but still utilize all of the material.
4. And sadly, in one case (and maybe there are others), the comment was that "I don't let anyone else view the training – I want to be the only on-site expert".

The benefits of planning

Now, let's assume that everyone has the required training, what else can you do to ensure success? The key is to create a program that is designed to succeed. Based on my own experience, and the feedback from my survey, the following sections offer a number of important tips.

Start small

It is essential that the program start small. If you try to test too many machines you will spread yourself thin. It will take time to get the system set up correctly (the database, etc.), and it will take time to streamline the measurement routines and analysis procedures. Choose a small group of machines (more on this shortly) and try to do a good job with those.

I cannot stress this enough. Take small steps. Start with simple machines, just so that you can get some experience using the equipment and the software. Analyze data that is relatively easy to understand. It is important for your own confidence that you do not immediately begin testing a large number of machines, or the most complex machines – even if they are critical.

Once you have some experience, work up to the larger machines. That is when you can establish the program and start a regular program. But do not rush into it. Take your time, and do the "serious" work *only* when you are ready.

Select the machines wisely

If the plant stops when a machine stops, monitor that machine. (Also look for machines that pose safety risks, and expensive machines.) However, if that machine poses technical challenges, then it may be wise to avoid the machine – to begin with. For example, if it operates under varying speed and load, or it is very difficult to access the monitoring points, or it has a very complex gearbox, then it may be beyond you to successfully collect or analyze the data.

Figure B-9

Especially to begin with, it is better not to monitor the machine than to try and fail. You need to build your own confidence, and the confidence of the rest of the maintenance and production staff.

However, as I will reinforce in a moment, *it is essential that your manager knows which machines are being monitored*.

Select an appropriate test schedule

I would typically recommend that you start with a thirty-day cycle. If a machine appears to be "healthy", and the condition is not changing, and it is not critical, then you could push it out to a sixty-day cycle. However, if the machine does appear to have a fault condition, then depending on the severity of the condition, and the criticality of the machine, you may even monitor it weekly or daily.

Don't work in isolation

Vibration monitoring is great, but it is not perfect. It *cannot* help you to understand the condition of *all* plant equipment. Vibration monitoring will not necessary give you the earliest warning of a fault condition.

Infrared analysis, oil analysis, wear particle analysis, acoustic emission, motor signature/circuit analysis, and other technologies all play a vital role. If someone is performing these tests within your plant (either in-house or a consulting service) make sure that you learn about the technology, and befriend the people involved. Share the data to get the best results.

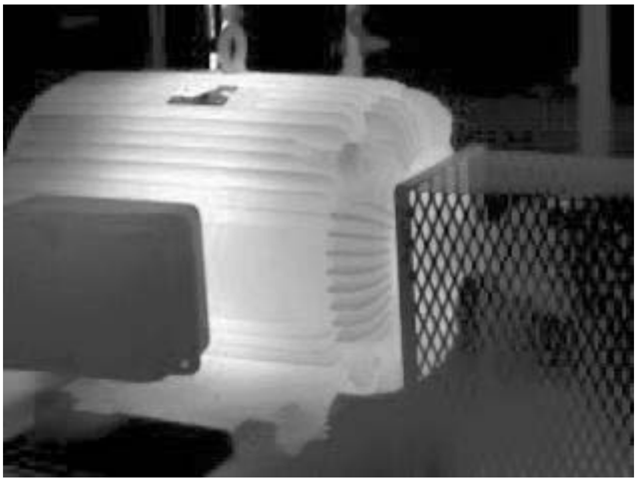


Figure B-10 A thermographic image of a motor

Understand the failure mechanism

Before you add a machine to the vibration program, you should do your best to learn as much as you can about the machine. If you have access to mechanical information (bearing number, gear ratios, vane counts, etc.), that will assist your vibration analysis work. But it is more important to understand the maintenance history and failure mechanisms. For example, knowing that a machine has a history of bearing failures will help to focus your attention, and you can choose the measurement types (shock pulse, demod spectra, PeakVue, etc.) accordingly.

Figure B-11 A sure sign that something is wrong with the machine

Understand the reporting process

When you detect a fault condition, be sure to report it to the right people. Make sure that your report is clear and understandable. It does not make sense to present a series of spectra and vibration levels to people who do not understand them. Find out what they need to know, and report it to them in the desired format.

Understand the production process

Sometimes you may detect a fault and report it to the required people, but they don't take the desired action. Don't get too stressed about it. At the end of the day, they have to balance the risk of failure (and the subsequent downtime and equipment damage) against any loss of production due to the unscheduled maintenance outage.

I have one case history where a bearing on a fan showed increasing signs of wear. Reports informed managers, but the decision was taken to keep running the fan. The fault condition worsened month by month, but still nothing was done. There came a time when the analyst reported that failure was imminent. Still no action was taken. And then the machine failed. And there was secondary damage to the machine, and there was downtime. In part it was a calculated risk – in part it was a lack of understanding and confidence in the vibration monitoring program (and condition-based maintenance).

Be realistic about what you can detect

When you understand more about vibration analysis and the rotating elements of your machine, you can begin to understand what types of faults can be detected. However, you must understand that some fault conditions can appear quickly, and failure can occur rapidly. So, if a machine does fail, be sure to investigate the failure mechanism, and then review the technologies utilized.

Figure B-12 Inner race bearing fault – why did this happen?

Utilize automated diagnostic reports

Most software packages are able to automatically "screen" the collected spectra, comparing new readings against mask or band alarms, and in some cases providing a diagnostic report. Use it!

Most systems require a considerable amount of work to set up the alarms, and/or describe the machine to the software. So many people have purchased these systems (or they come as a standard part of the vibration analysis software), but they are not being used. Instead, each new measurement is viewed manually, which takes a considerable amount of time.

Invest the time to set them up, and then allow them to save you time (and focus your analysis effort) every day.

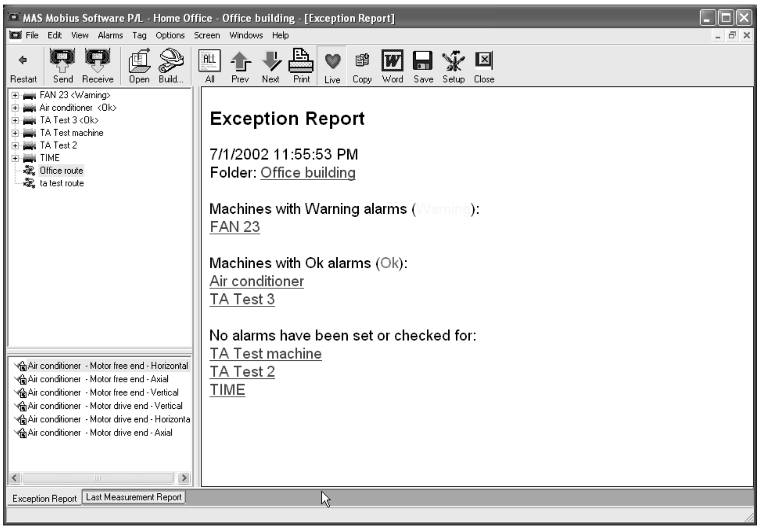


Figure B-13 Sample exception report

Find a mentor

If you can find someone with experience to oversee your program, either from a sister plant or an outside consultant, you may avoid making costly mistakes. Don't be too proud to ask for help. The profitability of your plant (and the continuation of your employment) may depend on it.

Figure B-14 Don't be afraid to ask for help

Don't be too "gung-ho"

If you think you will change the world in a day, and try to force people to accept your recommendations, and try to take on too much, you are likely to fail. You need to be cooperative. You need to be seen as an ally. You need to help people achieve *their* goals. Once people perceive you as someone who can help, rather than as a threat, then you will have success.

Figure B-15 Teamwork was its rewards

Keep management informed

No matter which machines you choose to monitor, no matter which technologies are used, no matter how often they are monitored – make sure your manager has a clear understanding of what can be achieved. It is the wrong time to explain your plan *after* there is an unexpected failure.

You need a "champion"

Most of the successful condition monitoring groups have a champion. He or she is the person who drives the program forward, inspires the people collecting the data and performing analysis, but also involves people from other sections of the plant – providing training and involving them in meetings.

And most importantly (for the success of the group), the champion inspires management to maintain and grow the program.

Mature program

The following is a list of tips for an established condition monitoring program – how to keep the program alive and continue to be successful and effective.

Record your successes

In order to keep the program going, it is **essential** to keep track of the money saved (or costs avoided) by the group. Quite often there can be exciting saves in the early days of a program, as there is a good chance that you will come across machines that have existing fault conditions. Record the repair cost, and estimate the downtime avoided.

It is important to understand the cost of downtime. For example, let's say that a repair to a pump cost $2,000, but it is deemed that it would have costs $10,000 if it were allowed to fail; the saving is not "just" $8,000. You must take into account the cost of downtime avoided. Depending upon the application, and manufacturing process, etc., it may have cost $20,000 if production stopped while the repair was performed. So the actual amount saved is $28,000.

You must be vocal and clear. Take pictures, keep notes, build case histories, and make sure everyone knows about the effectiveness of the program. Don't "show off" or annoy people, but if management does not understand the group's value, budgets will be squeezed (there goes the training budget) and squeezed (there goes the software upgrade) until the group no longer exists (there goes your job).

But you have to be sensible. If you continually report that you are avoiding thousands of dollars every time you detect bearing wear, nobody will take you seriously.

From a survey respondent:

"*Vibration readings indicated an outer raceway defect on a 1200 Hp motor. Those can only be shut down twice a year. We were able to trend and replace during the normal semi-annual outage without interruption to production. Failure during production would have set serious limitations on production and possibly plant wide shutdown.*"

Investigate other technologies

As I have already stated, vibration is not the only condition monitoring technology, so explore other methods (infrared, oil analysis, etc.).

From a survey respondent:

"During a thermography survey, I found a 34,500 Volt disconnect with a 250°F hot spot on the arm. Repairs showed major metal fatigue and cracking. The supervisor stated to management on my behalf that if I would not have found that problem it would have gone undetected. It definitely would have failed and caused a total loss of electrical power distribution to major portions of plant. Loss to production would have exceeded US$100,000."

Expand the program

Try to expand the program in small steps. Slowly increase the number of machines monitored. That may require additional staff to collect data, and as the collection periods are increased on certain machines you will naturally have more time available.

Consider installing permanent sensors on machines that are difficult to access, and on-line monitoring systems on machinery that is either difficult to monitor, difficult to predict faults (bearings seem to fail quickly), very critical to production, or that pose a safety risk.

You should learn the language of upper management. You cannot simply report the *benefits* of a new technology. You should determine the ROI, NPV, IRR and other financial measures when you propose additions to your program.

Figure B-16

But please, **don't build an empire**. Don't add people and technologies just so that you have a larger group. A technology might be "cool", but it may not be appropriate. Remember, the goal is to save the organization money in reduced downtime and reduced repair/maintenance costs. Bigger is not always better!

Perform root-cause analysis

When you determine that a fault condition exists you must find out what caused the problem in the first place. Why did the bearing begin to wear prematurely? Was the lubricant contaminated? Was the machine misaligned?

With that knowledge you can review your monitoring practices AND you can review the operation/maintenance of the machine so that the machine is more reliable in the future.

Build case histories

When you make a recommendation, don't just wash your hands of the situation and move on. When the repair is performed, verify what was actually wrong with the machine. Take photos, talk to the repairers, and save the bearings or other failed part. You can learn a great deal, and it is useful information you can present to management.

Verify the repair

When a machine is returned to service, collect more data to make sure that the fault was actually repaired (and that it is balanced and aligned). Your diagnosis may have been incorrect, so the machine may not be fit for service - and you still need to determine what is really ailing the machine.

Survey results

The survey was sent to a number of my iLearnInteractive users from small and large companies, from a variety of industries, and from different countries. Given that they are iLearnInteractive users, they are probably value training, but their comments and results were interesting nonetheless. Here are a few comments:

Management support

Almost all respondents said that their biggest concern was that their management did not understand what they did. They felt that with greater understanding they would get better support, and greater cooperation when either funds were needed for training, upgrades, etc., or more importantly, when a recommendation was made to correct a fault condition.

Figure B-17

Training

The majority of people said that their management appeared to value training. A number reported that when “times got tough” the training budget was slashed.

All respondents believed that training was essential to their success. All respondents believed that on-going training, and quick access to a good reference was essential. All respondents believed that computer based training was the most effective form of training, because they could learn at their own pace, at their own timetable, and could get “refresher” training at any time.

Most respondents also believed that it was essential to also attend seminars, conferences or classroom training so that they could get face-to-face with trainers and their peers.

Figure B-18

Return on investment

Almost all respondents said that they were sure that the investment in condition monitoring technology was an excellent investment, with an ROI achieved in less than one year. In many cases, however, they did not have the financial data to justify that statement.

Sometimes it was difficult to gather this information. One person queried "*how can you tell what the cost saving is by not delaying a space shuttle launch by a few days*". While we don't all work for NASA, it is not always easy to determine the exact financial benefit of avoiding unexpected failures and improving reliability.

Financial information

Almost all respondents believed that recording financial data was important, so that they could justify their program, however less than half of the respondents appeared to actually record the figures. One person said "*I don't do money!*".

Certification

Certification options have changed a great deal in the past couple of years. The ISO has released the standard 18436, and part 2 deals specifically with vibration analysis. To the best of the author's knowledge, only Mobius Institute and Vibration Institute follow this standard in the USA.

The ASNT has also defined a "recommended practice" for vibration analysis which includes a specification of the topics that must be covered and the format of the exam.

Certification is predominantly used to measure a person's understanding of the concepts and skills that must be used as a vibration analyst. It is not yet possible to provide a practical examination of a vibration analyst's competency.

However it is believed that certification to an internationally recognized standard will become a requirement for vibration analysts in the future. The advice provided by the analyst, and the potential for significant financial ramifications if an analyst misses a fault condition (or misdiagnoses a fault condition), will result in employees and clients demanding a proven level of proficiency.

Figure B-19

The costs required to run a program

Often when a company is looking to purchase a system, a great amount of time is spent analyzing the purchase costs. Bargaining and horse-trading continues until the lowest price is achieved. But little thought is given to the ongoing costs – the life-cycle costs.

All respondents stated that salaries were by far the largest cost. In the first year (when equipment is purchased), the purchase costs represents a large percentage (if you do not take depreciation into account), but still not as much as the staff salaries. In subsequent years, the maintenance costs (vendor support contracts, upgrades, etc.) can still be quite expensive, but still the dominant cost is staff (depending upon the size of the program, the number of employees, etc.).

So, what is the point? It is essential, in my opinion, to ensure that the equipment you purchase will do the job for you. If you understand your plant, and the failure modes of your equipment, you will have a better chance to buy appropriate equipment. Make sure you buy equipment that enables you to achieve the desired goals. That does not mean that you have to order the optional "bells and whistles", but you should spend money wisely. (And remember, it is easier to buy wisely in the first place, than to go back later and ask for more money.)

The second point is that you should analyze the on-going costs of the program – training, upgrades, and maintenance contracts. Consider the future.

Failed programs

Three main reasons were given for failed programs:

1. Lack of management support. Management either assigned staff to too many tasks, not realizing what it took to run a vibration program, or they did not understand the financial benefits (whose fault is that?) and simply cancelled it because it was considered "too expensive".
2. Lack of understanding of the real benefits. In a mature program there should be fewer failures. Based on root cause failure analysis, maintenance procedures (and purchase procedures and design processes) are changed so that machines are more reliable. Greater reliability will result in fewer failures, which can result in upper management asking "what have you done for me lately" – just before they cut the program. You can suffer as a result of your own success.
3. Lack of training. Staff felt so frustrated at not knowing how to use equipment, or more-so not understanding the technology, that they could not be successful. They grew frustrated and "gave up".

Final comment

I have given a number of reasons why programs can fail, and what you can do to grow the program and make a success of it. Success is not measured by the size of your budget or the number of people in your team – it is measured by how much you are saving your organization.

There is nothing worse than when an organization runs out and buys a system without really giving it much thought. They don't have a plan, they don't have training, they don't have a champion – they don't have a clue. But when it inevitably fails the comment is "*vibration analysis – yeah, we tried that and it does not work!*".

Vibration monitoring does work, it can return great benefits, and with some effort and planning (and training), you can make it work.