An observation on non-linear behaviour in condition monitoring

Apirak Jiewchaloemmit[1], Janewith Luangcharoenkij[2], P.W. Bland[3*]
[1-3] Department of Mechanical Engineering, Faculty of Engineering, Asian University,
89 Moo 12 Highway 331, Banglamung, Chonburi, 20260.
Tel: 038-754 450, Fax: 038-754 460, Email: pwbland@asianust.ac.th

Abstract

Condition monitoring (CM) techniques sometimes assign each single observed change in measured data to a single predicted fault, with the healthy condition as the reference state. Warning systems may assume that when two different faults develop at the same time, the measured data would still allow identification of the two faults, following the laws of linear superposition. The main objective of the work was to demonstrate the basic principle of using vibration methods for the CM of rotating machinery, using a simple mechanical system comprising of a rotating shaft and two sets of bearings. Tests included several different defects, with vibration measurements taken for each defect in isolation, then again when defects were combined. This paper briefly reports the experimental work and highlights three classes of non-linearity. The strongest case of non-linearity is described as "non-linear negative superposition", where the effect of one defect was reduced by adding a second defect.

1. Introduction

CM (also known as health monitoring and other derivatives) can be an integral part of modern maintenance strategies, such as predictive or preventative maintenance [1,2,3]. The common theme is to maximise the relevant information (type, source or location, quality, quantity or sample frequency) about the status of the hardware of interest. This data is used to make the best decisions regarding repair, replacement, downtime and scheduling, to reduce maintenance costs and increase availability. These gains should outweigh the cost of the systems needed to record and analyse the data required. The greatest financial benefit comes when critical or costly components or machinery is protected. Advances in trend analysis, types and resolution of sensors, PC based data acquisition and analysis have reduced the time from damage initiation to detection. Furthermore, with PC hardware costs reducing and capacity increasing, improved connectivity and mobility, both the engineering and business arguments for using CM are getting stronger.

The basic aim is to benchmark the healthy condition, identify how the condition may deteriorate and hence list possible or at least most likely modes of defect initiation, growth and then failure. Instrumentation will be selected to detect the presence of a given mode above a certain threshold, and hence give a warning. Early warnings can allow for planning, rather than late warnings forcing panic, or even catastrophic failure and hence high repair and downtime costs.

A common application is that of rotating machinery. Vibration methods are usually used for relatively low frequency effects, with the majority linked to the main shaft rotation speed giving sub-synchronous, low and high harmonics. Shaft effects include whirl, unbalance, misalignment, and eccentricity, resulting in frequencies from just under half to most often twice the shaft speed, sometimes up to four times the shaft speed. Damaged bearings can introduce higher frequency signals based on the impact rates of the defect, as a function of shaft speed, geometry, number of balls or rollers, and contact angle if carrying an axial load. The functions vary if the defect is on the inner race, the ball/roller or the outer race. Radial resonances of the bearings can become relatively high, in the order of up to 100kHz. Gears can also contribute to higher frequencies, proportional to the number of teeth. This also applies to any bladed rotating components.

Vibration methods will typically measure time-amplitude and frequency spectrum data from each sensor at different locations, with some degree of trending analysis. There are advanced methods (beyond the scope of this paper) of measurement, analysis and algorithms. These include displacement, velocity or

* Corresponding author

acceleration data, cepstrum, wavelets and fuzzy concepts, applied to standard operating conditions as well as run-up/run-down scans and impact tests [3-6]. When compared to the expected healthy measurements using instantaneous or trend lines techniques, each single feature in the plot is assigned to known or predicted damage initiation or failure modes. Warnings are given when trip levels are met, pointing towards a specific defect. Each defect should be clearly identifiable, separated from any other defect, giving non-overlapping feature space.

As usual, quality and not just quantity of information allows for better decisions. The path from mechanical effect to instrumentation signal to drawing conclusions about the mode, current extent and rate of growth of damage is crucial. Any false assumptions in this chain of reasoning will reduce the quality of any decisions made. If the analysis is based purely on linear effects, and there is a strong presence of a non-linear effect, then the process path could be fatally flawed.

2. Initial project scope

A 2nd year undergraduate project as part of a course entitled "Engineering Research Methods" was set up to:

1. Conduct background research into the discipline of CM.
2. Identify a simple concept that could be tested in the lab, with only a short time and small budget available for the design, build and testing process.

These constraints and the available lab equipment resulted in a study using vibration methods to analyse a simple rotating shaft with simulated defects.

3. Experimental work

3.1 The test rig

The test rig consisted of an AC motor driving a 3/8 inch diameter steel shaft through a relatively soft coupling, with the shaft supported by two radial bearings as shown in Fig. 1. The motor used mains supply, rated to 60W and 1300 rpm (21.6 Hz). The coupling allowed a small degree of radial movement between the motor and main shaft for the case of shaft misalignment, whilst providing sufficient torque and resilience against small oscillations in rotational speed due to ball bearing roughness. The main shaft was housed in two radial bearings, SKF 609-2Z, bolted to a steel platform. The bearings could be easily replaced and their vertical position incrementally changed by adding thin washers. An accelerometer was positioned centrally on either set of bearings at locations 1 or 2 (L1 or L2), in the vertical plane. Signal conditioning, amplification and processing were performed using National Instruments PC based dynamic signal acquisition hardware and Labview software, giving time-amplitude and frequency spectrum plots. Factory tested sensitivity and calibration factors for the accelerometer were used, but there was no independent verification of these through a separate test.

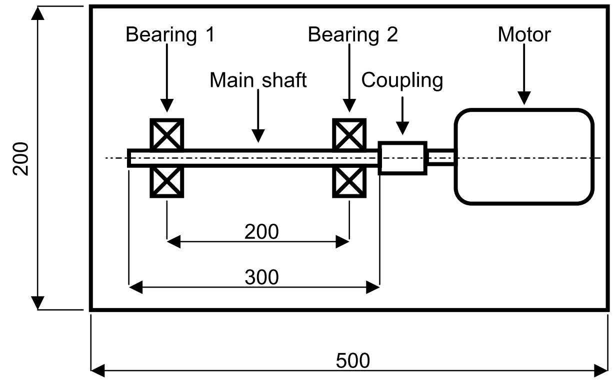

Fig.1: Schematic of the test rig, top view with sectioned bearings and overall dimensions (mm).

3.2 Defining the defects

The following define the single defects, denoted by D1, D2, D3 and D4. Some of these types of defect can be applied at either L1 or L2.

3.2.1 Simulating a defective bearing "D1"

The bearing casings were slightly over stressed by applying a lateral loading to a shaft as shown in Fig.2.

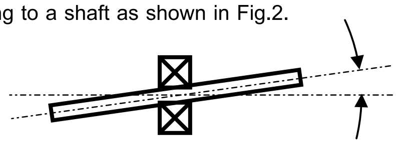

Fig.1: Lateral loading of the shaft causing overstressing of the bearing casing.

The bearings were also heated to add uneven distortion and minor localised changes in material properties. These effects were not fully quantified, however the end result was an increase in rotational friction with an oscillating amplitude as a function of shaft position. A static test measured the torque required to overcome friction, by applying known weights at 50mm from the bearing axis. This classified the bearings into three groups: "undamaged", "slightly damaged" and "severely damaged" needing masses of less than 1g, 3g and 35g respectively to overcome friction. All tests used an undamaged bearing, unless denoted by D1, where a slight damaged bearing was used, unless otherwise stated. Assuming no axial loading, the bearing geometry would give an expected defect frequency of 5/3, 4/3 or 20/3 times the shaft speed for damage to the ball bearing, outer raceway and inner raceway respectively.

3.2.2 Simulating shaft misalignment “D2”

1mm thick washers were used to raise a bearing in the vertical direction, giving a small angle of misalignment. The restraint of the bearings induced some bending of the shaft and would therefore include the effect of an unbalanced load as well as loading on the bearings. Coupling of defects such as this is typical in CM. However, this coupling was expected to be low due to the small angle, compared to the forces present that relate to the misalignment.

3.2.3 Simulating an unbalanced load “D3”

A small mass of 14g was attached to the end of the main shaft, near bearing 1, with its centre of mass approximately 10mm from the axis of rotation. This could represent an unbalanced component such as a rotor. At 1300rpm, the radial force generated is 2.6N. Even for the scale of the test rig, this was a small load, compared to the rated axial static load of the bearing of 1.66kN. However, this defect might be more representative of an early sign of an unbalanced load.

3.2.4 Simulating loose bearings “D4”

Securing nuts were removed leaving only the bolts in place to resist lateral movement, but limited resistance in the vertical plane. The coupling to other effects could again be present here, and hence not a true single defect. The possible motion of the centreline of the shaft and bearings could be complex, and may introduce noise to any measured data, rather than any clear cyclic signal.

3.3 Test matrix

A series of tests compared single and multiple simulated defects to the reference vibration measurement of the assumed perfect condition. Not every combination of defects was tested, but a selection was chosen to try to catch some interesting behaviour. For each test condition, the motor speed was set to its maximum of 1300rpm (21.6Hz), and vibration measurements were taken from L1. Sample times were set to 0.5s, relating to approximately 11 shaft cycles. A dynamic range of 0-100 Hz was used to limit observations to harmonic multiples of the shaft speed, avoiding higher frequency effects. Whilst this may miss a lot of information from higher frequencies, it would still capture any interesting lower frequency examples of non-linearity and simplify the analysis. Each test condition was repeated giving five sets of data for statistical confidence, giving a total of 60 tests. The complete list of tests with incremental numbering was therefore:

3.3.1 Reference test

1. Healthy condition. No induced defects, using new bearings, carefully aligned and balanced shaft, giving a baseline vibration measurement.

3.3.2 Single defect tests

2. D1 (at L1).
3. D2 (1mm at L1)
4. D3
5. D4 (at L1)

3.3.3 Multiple defect tests

Multiple defects used combinations of D1-4. The number of combinations was limited to reduce the total time for the experimental work.

6. D1 (severely damaged bearing at L1 and a slightly damaged bearing at L2).
7. D1 (at L1) and D3.
8. D1 (at L1) and D4 (at L1 and L2).
9. D1 (at L1) and D2 (2mm at L1).
10. D2 (2mm at L1) and D3.
11. D2 (2mm at L1), D3 and D4 (at L1 and L2).
12. D1 (severely damaged bearing at L1 and a slightly damaged bearing at L2), D2 (2mm at L1), and D4 (at L1 and L2).

3.4 Key results and discussion

In all tests, the fundamental and harmonic frequencies of the shaft speed could be identified, giving measured peaks at 23, 46, 70 and 90 ±2Hz. When identifying the presence of defects, the separation of the data as presented in the frequency spectrum was not convincing, needing improved resolution and data points in the frequency axis. Hence all other results rely on comparing peak or RMS average acceleration amplitude during the sample period of 0.5s. Although this limited the conclusions drawn from the experimental work, it still allowed for the following key results. Table 1 shows a summary of the RMS and peak values for tests 1 to 12. For test No.s 2-12 and each of the five columns, two values are quoted. The top value is the peak amplitude, the bottom value a scaling factor between the given peak value and the equivalent in test No.1, hence normalising the results. These numbers are shown in italics. The peak value is split into four columns. As the measurements gave acceleration, the data included both positive and negative values, labelled as “+” and “-”, each with the absolute “ABS” maximum magnitude, hence “ABS +” and “ABS - ”.

Table 1: RMS and peak amplitude acceleration (g)

Test No.	RMS	Peak ABS +	Peak 1% +	Peak 1% -	Peak ABS -
1	0.1015	0.3279	0.2394	0.2238	0.3397
2	0.1826	0.8183	0.4450	0.4394	0.7947
	1.80	*2.50*	*1.89*	*1.96*	*2.34*
3	0.1511	0.5052	0.3490	0.3271	0.4993
	1.49	*1.54*	*1.46*	*1.46*	*1.47*
4	0.0997	0.3220	0.2364	0.2175	0.3486
	0.98	*0.98*	*0.99*	*0.97*	*1.03*
5	0.1726	0.4963	0.3951	0.3538	0.4254
	1.70	*1.51*	*1.65*	*1.58*	*1.25*
6	0.3304	1.6101	0.7813	0.7991	1.5362
	3.25	*4.91*	*3.26*	*3.57*	*4.52*
7	0.1124	0.3870	0.2499	0.2482	0.3752
	1.11	*1.18*	*1.04*	*1.11*	*1.11*
8	0.2738	0.8597	0.6255	0.7501	1.3560
	2.70	*2.62*	*2.61*	*3.35*	*3.99*
9	0.2031	1.3442	0.4653	0.5014	1.0576
	2.00	*4.01*	*1.94*	*2.24*	*3.11*
10	0.0966	0.3486	0.2238	0.2202	0.3309
	0.95	*1.06*	*0.94*	*0.98*	*0.97*
11	0.1427	0.5436	0.3347	0.3401	0.5879
	1.41	*1.66*	*1.40*	*1.52*	*1.73*
12	0.3299	1.2438	0.8351	0.7724	1.2438
	3.25	*3.79*	*3.49*	*3.45*	*3.66*

Fig.s 3 and 4 show test data for test No.s 1 and 9 respectively, in the form of histograms. For test No.1, the ABS + and ABS - values are close to the main body of the histogram, in terms of the position along the x-axis. For test No.9, the ABS + and ABS - values are much further away from the main body of the histogram. After 1.3442g, the next highest positive magnitude for peak acceleration for test No.9 was 0.8295g, with zero number of events between these two values indicating that the ABS + value was an anomaly. For this reason, 1 % cumulative values are given for both positive and negative peak values, labelled as “1% +” and “1% -” respectively. This will discard the first 1% of maximum values, making sure that the 1% + and 1% - quoted peak values are near the main body of the histogram curve. Anomalies such as occurred for test No.9 are highlighted in grey, giving only two cases out of a possible twenty-four.

Statistical analysis across the five data sets for test No.1 gives a measurement uncertainty of ±5%. Apart from the two anomalies, nearly all peak values give symmetry about zero amplitude within this uncertainty range. The exception being test No.8 (compare peak 1% + to peak 1% -), although this may be explained by the 1% cut-off not being high enough, still giving an anomaly for the peak% - value of 0.7501g.

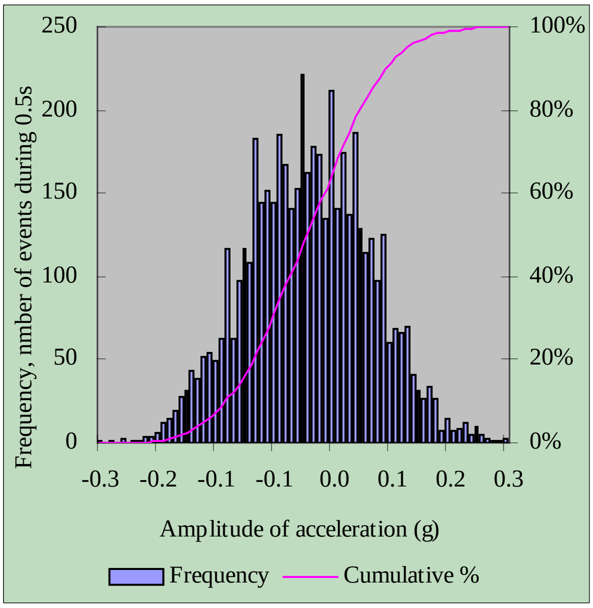

Fig.3: Histogram for test no.1.

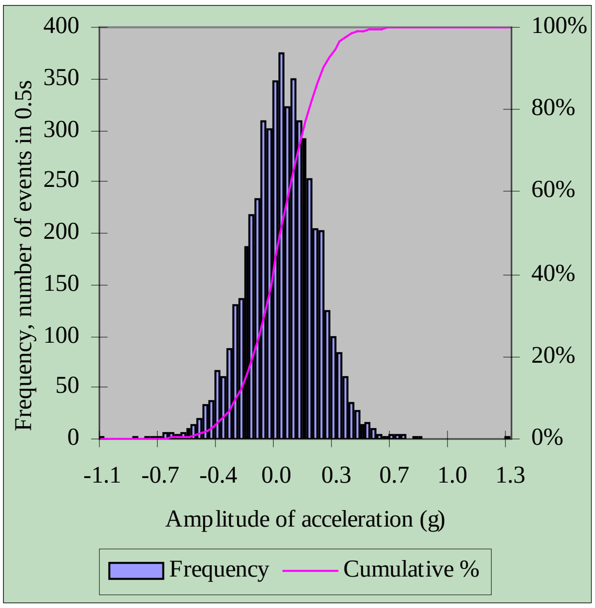

Fig.4: Histogram for test No. 9.

It should be noted that the histogram data gives the summation of amplitude contributions from the full frequency range at a point in time, and counts the number of times that this total amplitude occurs during the sample period of 0.5s. The term frequency in the context of the histogram should not be confused with the actual frequency related to the vibration behaviour. This means that the histogram does not give any vibration frequency information, so is only useful for checking for statistical anomalies.

3.4.1 Non-linear scaling factors (weak non-linearity)

For most tests, the RMS and peak amplitude values seem to scale linearly compared to the reference test. For example, in test No.7 the scaling factors for all 5 amplitudes are within 1.11 ±0.07. This would indicate a linear relationship between the methods of representing amplitude, and is further confirmed by looking at the histograms, which show very similar profiles. Exceptions being the maximum peak anomalies highlighted in grey in Table 1 and test No.s 2 and 6. Fig 5 shows the histogram for test No.7, included for further comparison with Fig.s 3 and 4.

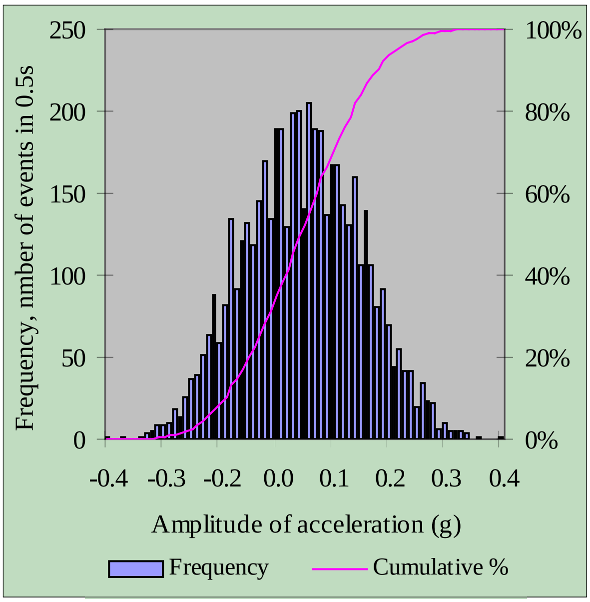

Fig.5: Histogram for test No. 7.

For test No.6, the scaling factor for RMS is 3.25 but for the peak ABS values, it is between 4.5 and 4.9. The histogram for test No. 6 does not suggest that these are anomalies. Test No.6 uses severely damaged bearings, and has a greater degree of this type of non-linearity compared to test No.2, which only uses slightly damaged bearings. It may be possible to use this behaviour to identify the difference between, for example, damaged bearings and misalignment, without using the frequency spectrum. This can be demonstrated by trying to select a thickness of washer such that the RMS value in test No.3 is 0.1826g, the same as for test No.2. This would not give the same ABS peak values and hence aid identification of the defect.

This type of non-linearity is semi-artificial in that parameters have been identified where there may be no expected link. In this case, the question is why would the scaling factors for any given defect be the same, for RMS, peak ABS +, peak 1% +, peak ABS -, and peak 1% -.

3.4.2 Non-linear positive superposition (standard non-linearity)

Test No.9 combines the single defects from test No.s 2 and 3. Test No.9 shows higher levels of vibration amplitude than either test No.s 2 or 3 on their own, but lower amplitude than adding the effects of tests No. 2 and 3. For example, considering the 1% + peak values:

0.4653g (No.9) > 0.4450g (No.2), and

0.4653g (No.9) > 0.3490g (No.3), but

0.4653g (No.9) < 0.7940g (0.4450g + 0.3490g)

This is a standard definition of non-linearity, with the simple laws of linear superposition are not followed. It may occur simply because the governing equations are non-linear in the strict mathematical sense, perhaps including squared or higher order terms.

3.4.3 Non-linear negative superposition (strong non-linearity)

Test No.4 seems to indicate that there was no detectable rise in vibration amplitude for the single unbalanced load defect. All measured values are well within the measurement uncertainty of ±5%. This suggests that the force generated by the out of balance mass was not high enough to cause any significant increase in the measured data. However, comparing test No.s 2 and 7 to the reference test suggests that the unbalanced load does produce an effect. In test no.7, with a slightly bad bearing at L1 and an unbalanced load, the RMS and peak amplitude values increased by 10% compared to the reference test. In test no.2, with only a slightly bad bearing at L1, the RMS and peak amplitude values increased by approximately 100% compared to the reference test. The presence of an out of balance load appears to reduce the effect of the slightly bad bearing. A similar result can be seen by comparing test No.s 3 and 10 to the reference test, showing that the unbalanced load appears to reduce the effect of misalignment. Note that in test No.10, the misalignment is generated by using two 1mm washers, compared to one 1 mm washer for test No. 3. This is the strongest example on non-linear behaviour from all the test data presented in this paper. It can be thought of as negative superposition of combined defects. This may occur because the small unbalanced force is sufficient to alter the contact mechanics between the bearings and the raceways.

Note that for test No.s 11 and 12, standard and strong non-linear behaviour may occur, but possibly dominated by the effects of both sets of securing nuts being loose, and in the case of test No.12, also by the badly damaged bearing at L1. This could be due to the unbalanced load not being able to transmit the force through to the bearing outer raceway, as it is no longer restrained by the securing nuts.

This type of non-linearity has its source buried within the detail of the mechanics, perhaps at the local level. An example would be the behaviour of an object pressed against a surface, such as a ball bearing against a raceway. If the force applied to the ball bearing is labelled as positive when forcing the ball bearing against the raceway, then the relationship between the positive force and the reaction from the raceway may be linear. The relationship between the force when negative (pushing the ball away from the surface) and the reaction force may be a step function, at the point when the ball is about to loose contact with the surface. Although this is a very simple model, it gives a clue as to the possible source of strong non-linear behaviour, essentially in the form of discontinuities. It should be noted that this is a mere hypothesis, with the detailed mechanics of the observed non-linear phenomena needing further investigation.

4. Conclusion

This paper has demonstrated that even the most simple form of rotating machinery can include non-linear effects in CM measurement and analysis. Three groups on non-linearity were proposed, with the strongest being the surprising result of one defect actually reducing the effect of a second defect by 90%. This is classed as non-linear negative superposition. If the warning system, from defect to measured data to algorithm and decision threshold, is based on linear relationships, then acceptable damage levels could be exceeded but no warnings given.

Another consideration is the feature space. Even if all relationships were linear, if the effect of two different defects was detected by the same type of measurement, then it would not be possible to identify the two separate defects. It is therefore important not to have overlapping feature space, perhaps requiring a range of detection methods. This will also help to identify non-linear effects.

Terminology used to describe the defects often includes more than one mechanical effect, which will automatically build in some degree of coupling, complexity and potential for non-linear behaviour. A sensible classification system is also needed to simplify any detection algorithm.

Recommendations for further work centre on being able to fully explain the observed non-linear behaviour. This includes developing the basic test rig and instrumentation to allow higher precision and sampling, especially in the frequency spectrum, reduce the vibration levels for the reference test and independent calibration. A wider range of methods of measurement are needed to verify observations, to link to specific mechanisms and to allow for feature space separation. The test matrix can also be expanded.

References

[1] Wild P., “Industrial Sensors and Applications for Condition Monitoring”, ISBN 0-85298-902-4, 1994.

[2] Bloch, H.P., “Practical Machinery Management for Process Plants”, ISBN 0-88415-661-3, 1998.

[3] Harris, C.M., “Harris’ Shock and Vibration Handbook”, ISBN 0-07-137081-1, 2002.

[4] Mobley, R. K., “Vibration fundamentals”, ISBN 0-7506-7150-5, 1999.

[5] Peng, Z.K. and Chu, F.L., “Application of the Wavelet Transform in Machine Condition Monitoring and Fault Diagnostics: A Review with Bibliography”, Mechanical Systems and Signal Processing, Vol.18, 2004, pp199-221.

[6] Wang, W.Q., Golnaraghi, M.F. and Ismail, F., “Prognosis of Machine Health Condition Using Neuro-Fuzzy Systems”, Mechanical Systems and Signal Processing, Vol.18, 2004, pp813-831.