A STOPPING CONDITION TO EMPIRICAL MODE DECOMPOSITION BASED ON THE GOERTZEL ALGORITHM FOR DETECTING FREQUENCY-BASED FAULTS

Daniel CORDONEANU[1], Constantin NIȚU [2]

Fault diagnosis has received a lot of attention from the research community recently. Empirical mode decomposition is one of the signal processing methods used in fault diagnosis that is self-adaptive and can be used in nonstationary and nonlinear signals without having to know anything about the signal before decomposing it. This method decomposes a signal into a given maximum number of signals. One of these signals can give information on whether a certain fault is present, rendering the rest of the decomposition useless. This paper proposes a new end condition for the decomposition process for monitoring frequency-based faults.

Keywords: Goertzel algorithm, Empirical mode decomposition, Fault diagnosis

1. Introduction

Fault diagnosis domain is getting a lot of attention nowadays as the computing power available in embedded systems is increasing. Fault diagnosis is mainly composed of fault detection and fault identification. Detecting an early fault can be very useful and it can save a lot of money, preventing system collapse, downtime, and time-consuming maintenance, thus being economically efficient for a business.

Mechanical systems can be easily monitored using vibration signals acquired from accelerometers mounted on different components. This data needs to be interpreted and certain characteristics must be extracted so that faults can be detected and isolated.

In [1] the author presents signal processing techniques in different domains like time domain, frequency domain or time-frequency domain. In frequency domain the Fourier transform is a method that uses an orthogonal basis to decompose a signal into a linear combination of that basis. However, the time information is lost in this transform. Also, the Fourier transform does not behave so well in nonstationary or nonlinear signals.

[1] PhD student, Department of Mechatronics and Precision Mechanics, University POLITEHNICA of Bucharest, Romania, e-mail: daniel.cordoneanu@stud.mec.upb.ro
[2] Prof., Department of Mechatronics and Precision Mechanics, University POLITEHNICA of Bucharest, Romania

To accommodate the loss of information from time/frequency only methods, there are other methods available that offer the needed data in the time-frequency domain. Some of these methods are Short Time Fourier Transform (it's the Fourier transform applied on the signal split into chunks using a windowing method), wavelet transform (a suitable basis needs to be chosen a priori in order to extract the most meaningful characteristics of the signal), Wigner-Ville distribution or the Hilbert-Huang Transform (HHT), presented in [2].

HHT is composed of two parts: the empirical mode decomposition (EMD) and the applying of the Hilbert transform to each of the obtained signals called intrinsic mode functions (IMFs). The EMD decomposes the signal until certain end conditions are fulfilled:

- The maximum number of IMFs has been reached
- The residue obtained from the current signal has at most one extremum

In [3], there are given specific insights on how the EMD has certain pitfalls and why derived algorithms should always be used to accommodate certain limitations. The iterative filtering (which is a method derived from the EMD)[4] is also compared to the EMD under different scenarios. However, all these methods are derived from the EMD. The main problem of the EMD is represented by the boundary issues that may appear in IMFs and there are many approaches on how to solve these problems presented in literature [5], [6], [7]. Given that the vibrations coming from a rotational machine are well defined in the way that they have an oscillatory mathematical representation, the EMD-derived methods are suitable for decomposing this kind of signals.

In this paper an additional condition is proposed for stopping the overall decomposing process of the EMD using the Goertzel algorithm for extracting the discrete Fourier transform coefficients only for a specific monitored frequency. This condition can be used in any of the EMD derived methods.

2. Empirical mode decomposition

For completeness, a short description of the EMD will be given as well as how the IMFs are obtained.

Huang et al. give a definition for the IMF in [2]: An intrinsic mode function can be described as a function that satisfies two conditions:

1. The number of extrema and the number of zero crossings must be equal or differ at most by one
2. At any point, the mean value of the envelope defined by the local minima and maxima is zero

From the signal point of view, an IMF is an oscillation, the decomposed signal representing a superposition of these modes of oscillation. The IMFs are obtained through a process called sifting. The sifting process stops when both the above-mentioned conditions are fulfilled, when the standard deviation size of two consecutive sifting results is out of a certain interval or when a predefined number of siftings have been made.

The sifting steps are the following:

1. The local extrema are found, and the signal's upper and lower envelopes are created by connecting the extrema points by means of a cubic spline
2. The mean signal between the two envelopes is derived
3. Let $x(t)$ be the initial signal and m_1 the envelope mean, then the first signal obtained by sifting is h_1. For repeatable steps let r_0 be $x(t)$:
$$h_1 = r_0 - m_1 \quad (1)$$
4. After the first signal h_1 was obtained, it is checked if the stopping conditions had been met. If not, h_1 is considered the new signal and the sifting process continues up to the number of maximum siftings (k) or until the new signal is an IMF:
$$h_{1k} = h_{1(k-1)} - m_{1k} \quad (2)$$
5. At the end of the process, the first IMF $c_1 = h_{1k}$ is found
6. The IMF is extracted from the signal r_0:
$$r_1 = r_0 - c_1 \quad (3)$$
7. The steps from 3 to 6 are repeated, increasing the indices of the signals until either the IMF signal c_n or the residue r_n becomes too small to be significant or the residue because a monotonic function from which there is no IMF to be extracted

For a fault diagnosis system that monitors a specific mechanism, it is useful that once a fault has been found, an alarm is logged, displayed and/or the entire system is stopped. The EMD algorithm can take some time and if the fault is detectable in the first IMF, then the inference system must wait until all the IMFs are extracted to assess the presence of the fault. In step 7, a new condition will be added related to the monitored fault frequencies.

2. The Goertzel algorithm

Standard Goertzel algorithm

Sysel & Rajmic show in [8] that the original algorithm described by Goertzel in [9] that computes the Discrete Fourier Transform (DFT) term of the signal $x[n]$ with length N is equivalent to a discrete linear convolution between the analyzed signal and a signal $h_k[n]$. Let $h_k[n]$ be:

$$h_k[l] = e^{j2\pi k\frac{l}{N}} u[l] \quad (4)$$

Where *u[l]* is the unit function. If y_k is the result of the convolution, then:

$$y_k[m] = \sum_{n=0}^{N-1} x[n] e^{j2\pi k\frac{m-n}{N}} u[m-n] \quad (5)$$

The authors further show in their paper [8] that equation (5) means that the N-th sample of the convolution is the desired DFT term. This means that the required value can be computed as the output sample in time N of an Infinite Impulse Response (IIR) linear system with the impulse response $h_k[n]$. By using the impulse response of a linear system, any response can be computed afterwards. Using the z-transform and a mathematical trick, the transfer function of the IIR system is brought to a second order IIR system. Using differences, the equation of the system becomes:

$$s[n] = x[n] + 2\cos(\frac{2\pi k}{N}) s[n-1] - s[n-2] \quad (6)$$

With output of the system being:

$$y_k[n] = s[n] - e^{-j\frac{2\pi k}{N}} s[n-1] \quad (7)$$

As the creator of the original algorithm mentioned, only *N* multiplications and *2N* additions are needed which makes this algorithm computational fast.

Goertzel algorithm compared to the DFT

The Goertzel algorithm has some advantages over the DFT under certain conditions. If $K < 4N/7$ [8], where *K* is the number of frequencies for which the coefficients are to be extracted, then the Goertzel algorithm is superior in computation speed to the DFT.

Another advantage of the Goertzel algorithm is that one can inspect a signal of length *N* without bothering that *N* is a power of 2 (case in which DFT is computationally fast by applying the Fast Fourier Transform).

Also, while inspecting the spectrum, based on the sampling frequency and the number of samples the DFT is prone to spectral leakage (e.g. if a signal has a sampling frequency of 12 kHz and the DFT has 1024 number of points, the frequency per bin of the DFT would be 12kHz / 1024 = 11.71 Hz per bin so that if the magnitude of the sinusoid with a period of 115 Hz, this would be leaked to the neighboring frequency bins of 117.1 Hz and 105.39Hz).

Generalized Goertzel algorithm

With a small computation expense, in [8] a generalized Goertzel algorithm is proposed that can use the extracted frequency which may be a real number, not only an integer. This is helpful in frequency-based fault diagnosis, because the frequencies computed will most probably be real numbers since formulas to compute these frequencies involve sine and cosine functions.

3. The proposed algorithm

For the EMD process to stop when a fault has been detected, the fault must be characterized by some features. Since the extracted information from each IMF is represented by the DFT coefficients for each of the monitored frequency, the features of these DFT coefficients will be represented by statistical information.

To have a comparison dataset, the proposed algorithm will have two main stages:

1. The training stage – done on data extracted for the system with no faults
2. The diagnosis stage – the algorithm will run continuously and compare the extracted features for new data with the stored features extracted during the training stage

In the training stage, statistical information will be extracted based on the distributions given by the DFT coefficients for each monitored frequency. The extracted statistical information will be represented by the mean, variance, and range for each of the distribution.

The mean of a dataset $X=x_1 \ldots x_n$ can be mathematically represented by:

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j \tag{8}$$

The variance of the dataset can be computed using the following equation:

$$\sigma^2 = \frac{1}{n}\sum_{j=1}^{n} (x_j - \bar{x})^2 \tag{9}$$

The range of the dataset is:

$$x_{range} = \max(X) - \min(X) \tag{10}$$

These 3 features are enough to characterize a distribution for the purpose of checking how a new sample affects the recorded distribution for each of the monitored frequencies.

The algorithm can be split into the following steps:

1. Create a list of monitored frequencies
2. Start the EMD algorithm on a training signal for a functional system
3. Once all the IMFs have been computed, let A be a matrix with m by n (m is the number of monitored frequencies and n is the number of IMFs) filled with the absolute values of the DFT coefficients extracted using the generalized Goertzel algorithm for the fault frequencies for each IMF signal

4. Each row in *A* represents a dataset for which the measures presented in equations (8), (9), (10) must be extracted and stored; besides these features, also the A matrix is stored
5. During the diagnosis stage, once every IMF is extracted, the DFT coefficients are computed for the monitored frequencies and their absolute values are added as a new column to the *A* matrix, forming a new *B* matrix
6. The statistical features are extracted from the *B* matrix and compared to the ones stored at point 4. If each of the feature values have changed with more than an a priori chosen coefficient, the fault has been detected and the IMF decomposition can stop

4. Results

For testing the algorithm proposed above, a set of recorded data was used. The data is provided by the Case Western Reserve University and it's available on [10]. The acceleration data was extracted from the bearings that support a two hp Reliance Electric (figure 1 left side) shaft.

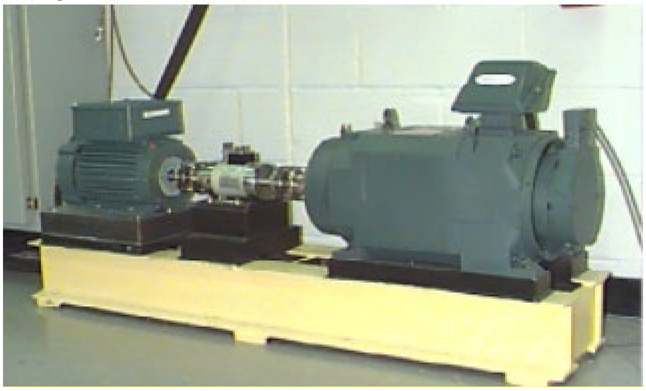

Fig. 1. Test stand from which the data for bearings was sampled [10]

Faults were made to the bearings using electro-discharge machining (EDM). Fault diameter on which the algorithm was tested is 0.017 mm and the data was sampled at a frequency of 12 kHz for the bearing being in a healthy condition and having faults on different components. The data was collected using accelerometers, which were mounted to the housing with magnetic bases. The accelerometers were placed on the top position at both the drive end and fan end of the motor housing. For the results, the data from the drive end was used.

Bearings are mechanical components for which fault frequencies can be easily computed if the rotation frequency of the rotating shaft is known and if the geometrical characteristics of the bearing are known. The used bearings in the experiment are deep groove ball bearings (6205-2RS JEM SKF) at a shaft speed of 1797 rotations per minute (rpm). The following components can be damaged in a bearing:

- Outer ring
- Inner ring
- Bearing balls
- Bearing cage

A localized bearing fault would normally produce periodic impacts, the size and the period being determined by the rotation speed of the shaft, the type of fault and the bearing's geometry. The successive impacts would produce a series of impulse responses, which may be modulated in amplitude given that the between the point of the impact and the vibration measurement point there are additional components which may modulate the frequency to their own resonance frequency. The spectrum of the measured vibration signal would be formed of harmonic series of frequency components that are spaced at the bearing defect frequency, having the highest amplitude around the resonance frequency. To detect the peaks at the fault frequencies, the instantaneous energy (or instantaneous envelope) of the signal is used because the impact components would increase the amplitude at the resonance frequency components [11]. The instantaneous energy can be easily computed as it is part of the Hilbert-Huang transform, using the Hilbert transform [2] which allows computing the envelope of a real signal.

The correspondent fault frequencies can be computed using the following equations [12]:

$$F_{outer} = \frac{nf}{120}\left(1 - \frac{B_d \cos(\beta)}{P_d}\right) \tag{11}$$

$$F_{inner} = \frac{nf}{120}\left(1 + \frac{B_d \cos(\beta)}{P_d}\right) \tag{12}$$

$$F_{ball} = \frac{f}{60}\frac{P_d}{B_d}\left[1 - \left(\frac{B_d \cos(\beta)}{P_d}\right)^2\right] \tag{13}$$

$$F_{cage} = \frac{f}{120}(1 - \frac{B_d \cos(\beta)}{P_d} \tag{14}$$

Where:

- n is the number of bearing balls
- B_d is the ball diameter
- β is the contact angle
- f is the rotation frequency of the shaft in rpm
- P_d is the pitch diameter computed as:

$$P_d = \frac{D_1 + D_2}{2} \tag{15}$$

Where D_1 is the diameter of the inner raceway and D_2 is the diameter of the outer ring raceway.

For the bearings used in the testing, the fault frequencies computed using the equations (11), (12), (13) and (14) are available in table 1.

Table 1

Fault frequencies for the analyzed bearing

Component	Outer ring	Inner ring	Bearing ball	Cage
Fault frequency [Hz]	107.36	162.18	141.16	11.92

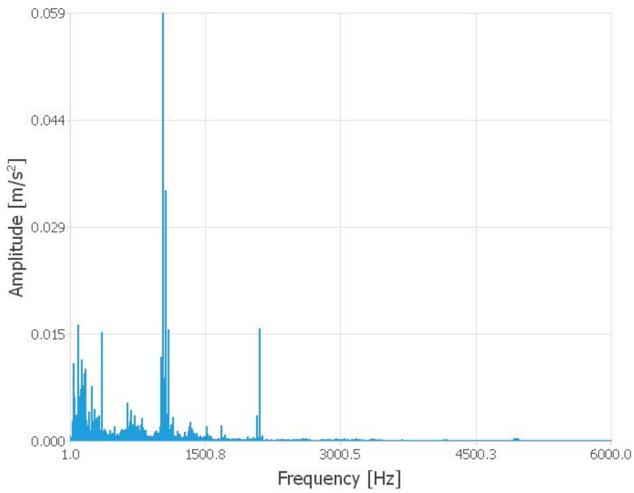

Fig.2. The spectrum of the vibration signal coming from a functional bearing

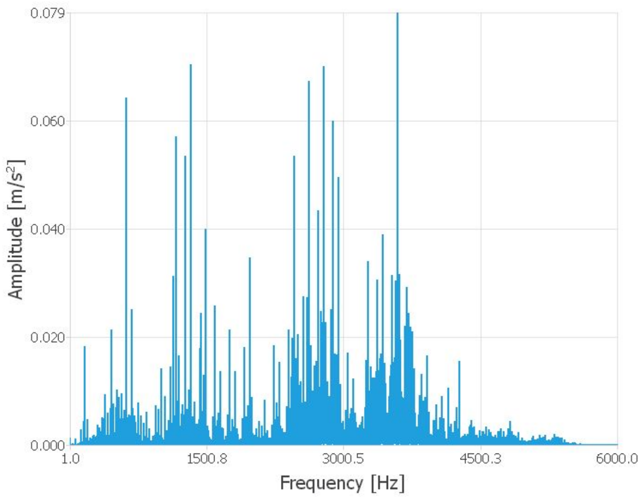

Fig.3. The spectrum of the vibration signal for a bearing with an inner ring fault

The spectrum of the vibration signals coming from a functional bearing and from a bearing with a fault present on the inner ring are presented in figure 2 and 3. The faulty bearing spectrum is clearly different than the functional's bearing spectrum. However, the fault frequency of the inner ring is not visible on the figure 3 chart and that's why signal processing is important when extracting features for a mechanical fault that are used by a condition monitoring system. Using the Hilbert-Huang transform allows extracting signal features in the form of IMFs and at the same time computes the envelope of each IMF and the instantaneous envelope for each extracted function and, as stated in [12], envelope spectrum is a good method to analyze and assert bearings condition.

Distributions for different data were extracted during the training stage. In figure 4 there can be observed the distribution of the DFT coefficients for the signal obtained for a functional bearing and the distribution for a bearing that has an inner ring fault. The distributions are specific to the inner fault frequency.

It is obvious from the figure that the distribution for the signal extracted from the faulty bearing has different statistical characteristics and that the features extracted by the algorithm are different, thus the EMD stopping from the first extracted IMF.

The coefficient mentioned at point 6 of the algorithm presented in section 3 was set to 0.7 during the training phase using an incremental approach. The coefficient represents the maximum offset allowed for any of the characteristics (in this case 70%). The method of finding it consists of first splitting the data for the functional bearing into two sets: training data and validation data. On the training data, the characteristics would be extracted using steps 1-4 of section 3. Afterwards, starting with a coefficient of 0.1, this was increased with a step of 0.1 until the algorithm would not detect a fault anymore for the validation data set (which is part of the functional data). This way, the coefficient can be chosen based on the functional data only, specific to each monitored system.

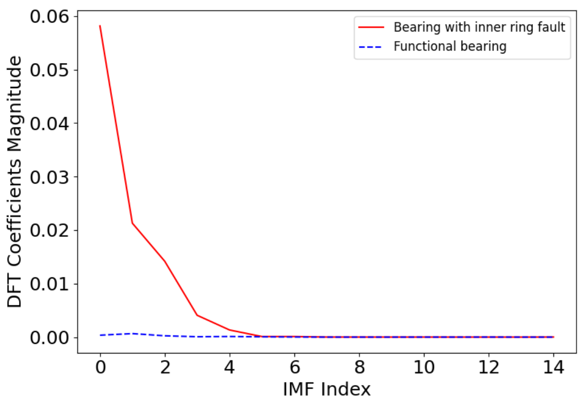

Fig. 4. Distributions of the DFT coefficients magnitude for a healthy bearing and for a faulty bearing at the inner fault frequency

Timings have been measured for applying the EMD on three signals for which the bearings have different faults: inner ring fault, outer ring fault and ball fault. The results are presented in table 2 and 3:

Table 2

Time for EMD completion on a 10s signal

Faulty component	**Time for EMD to finish without the algorithm implemented [s]**	**Time for EMD to finish with the algorithm implemented [s]**
Inner ring	64.1	5.2
Bearing ball	12.9	1.1
Outer ring	81.1	31.6

Table 3

Time for EMD completion on a 1s signal

Faulty component	**Time for EMD to finish without the algorithm implemented [s]**	**Time for EMD to finish with the algorithm implemented [s]**
Inner ring	0.7	0.1
Bearing ball	0.2	0.07
Outer ring	2.9	2.7

The algorithm improves greatly the finish timings, and it can also be used for diagnosis. Given that the Goertzel algorithm is dependent on the number of samples of the analyzed signal, it is important to have this number reasonable to avoid rounding errors.

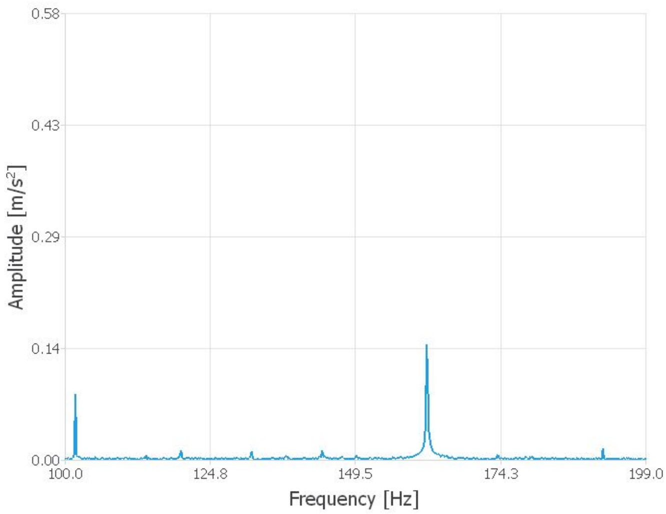

Fig. 5. The envelope spectrum of the IMF for which the EMD has stopped zoomed in the 100-199 Hz bandwidth

In figure 5 it is presented the envelope spectrum of the IMF for which the algorithm has stopped for a signal extracted from a bearing with an inner ring fault. The peak is around the fault frequency characteristic to this component, 162 Hz.

5. Conclusions

The results presented in the 4th section show that the algorithm can be successfully used for stopping the EMD at a preliminary stage if a fault has been detected. Obviously, the algorithm needs to be tested extensively on other data sets as well before being used, but it shows a computational optimization and if the fault diagnosis is run on a time-sensitive system, then any saved time is useful.

Besides finishing times, this algorithm can be also used to extract valuable information on what frequency component is dominant in a certain IMF. This can be helpful in a more complex fault diagnosis system which may detect and classify a fault, but it is important to check on which IMF index that fault is visible. These insights may help the signal to offer a probability on where an unknown fault can be coming from or to be able to differentiate an unknown fault from a known fault based on statistical data.

The tests were done in a custom application written in Python which can be used for signal analysis and diagnosis for research purposes. The EMD was run using the PyEMD library [13].

REFERENCES

[1] *Y. Lei*, Intelligent fault diagnosis and remaining useful life prediction of rotating machinery. Elsevier Inc., 2016.

[2] *N. E. Huang et al.*, "The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis," Proc. R. Soc. A Math. Phys. Eng. Sci., **vol. 454**, no. 1971, 1998, pp. 903–995.

[3] *A. Stallone, A. Cicone, and M. Materassi,* "New insights and best practices for the successful use of Empirical Mode Decomposition, Iterative Filtering and derived algorithms," Sci. Rep., **vol. 10**, no. 1, Dec. 2020, pp. 1–15.

[4] *L. Lin, Y. Wang, and H. Zhou*, "Iterative Filtering as an Alternative Algorithm for Empirical Mode Decomposition.", Advances in Adaptive Data Analysis, **vol. 1**, no. 4, Oct. 2009, pp.543-560

[5] *A. M. Jaber, M. T. Ismail, and A. M. Altaher*, "Empirical mode decomposition combined with local linear quantile regression for automatic boundary correction," Abstr. Appl. Anal., **vol. 2014**, 2014.

[6] *Q. Wu and S. D. Riemenschneider*, "Boundary extension and stop criteria for empirical mode decomposition," Adv. Adapt. Data Anal., **vol. 2**, no. 2, Apr. 2010, pp. 157–169.

[7] *A. M. Jaber, M. T. Ismail, and A. M. Altaher*, "Automatic Boundary Correction Using Empirical Mode Decomposition and Local Linear Regression," Far East Journal of Mathematical Sciences, **vol. 2013**, no. 6, Mar. 2014, pp. 563-576.

[8] *P. Sysel and P. Rajmic*, "Goertzel algorithm generalized to non-integer multiples of fundamental frequency," EURASIP J. Adv. Signal Process., **vol. 2012**, no. 1, Mar. 2012, p. 56.

[9] *G. Goertzel*, "An Algorithm for the Evaluation of Finite Trigonometric Series," Am. Math. Mon., 2006.
[10] "Bearing Data Center." [Online]. Available: http://csegroups.case.edu/bearingdatacenter/home. [Accessed: 02-Jun-2019].
[11] *N. G. Nikolaou and I. A. Antoniadis*, "Rolling element bearing fault diagnosis using wavelet packets" NDT E Int., **vol. 35**, no. 3, 2002, pp. 197–205.
[12] J. Mais and SKF USA Inc., "Spectrum Analysis, The key features of analyzing spectra," Science, 2012, pp. 16–18.
[13] "Intro — PyEMD 0.2.13 documentation." [Online]. Available: https://pyemd.readthedocs.io/en/latest/intro.html. [Accessed: 01-Apr-2021].