Frequency Measurements in Noisy Records using Vector-FFT

Herbert Weidner[A]

Abstract: The search for weak spectral lines with SNR ≈ 1 in noisy gravitational data is difficult, since the analysis methods used so far do not exploit all available information. The vector FFT presented here reduces the noise level and narrows the FWHM of all spectral lines by taking advantage of the phase information of the FFT.

Introduction

It is not particularly difficult to identify spectral lines, ie special frequencies in data series, as long as the signal-to-noise ratio exceeds 20 dB. The most commonly used method is called Fourier analysis, whereby the signal usually passes through several processing steps.

But how to discover unknown spectral lines in extremely noisy signals? Are all individual steps in a standard FFT necessary? Do they improve the result or are they even harmful if the SNR is very bad? Most of all: Do they destroy or ignore valuable information and therefore make the search more difficult?

The following analyzes and suggestions for improvement relate only to the search for extremely weak spectral lines in the frequency range around 60 μHz in noisy geophysical data series. There are no considerations as to whether they are universally valid and can be applied unchanged to any problem.

In the present question, some obstacles make it difficult to analyze the recorded data: The dynamic range is extremely large and amounts to about 120 dB. This means that the amplitude of the strongest tidal signals is at least 10^6 higher than the amplitude of the wanted spectral lines, the frequencies are similar. At irregular intervals, impulse-like noise peaks are recorded, which are generated by earthquakes somewhere on the globe. Each impulse creates another broadband spectrum that can hardly be modeled. Add to this the inherent noise of the measuring instruments.

Here is not discussed how the background noise can be minimized. This topic will be discussed separately. It is all about improving the FFT method so that spectral lines of constant frequency can be identified with greater certainty, even in very noisy data streams. Extensive series of experiments have shown that this goal can be approached by using the phase information, which is usually ignored in FFT. This will be discussed in detail below.

Comments on the usual procedural steps.

In a spectral analysis usually the following steps are performed:

Choice of a dataset that contains enough measurements. Accurate frequency measurement requires many (equidistant) readings because the Küpfmüller theorem must be satisfied. If you want to achieve a frequency uncertainty of Δf = 10 nHz, the formula $\Delta f \cdot \Delta t \geq 0.5$ implies that the entire measurement period must be at least two years. In the field of geophysics, there are measurement series that span ten years or more and show few interruptions. Shortening the measurement period reduces the accuracy. From a mathematical point of view, it does not matter if data points are taken at intervals of seconds or minutes. It requires a great deal of effort to eliminate obvious disturbances in the data record and to bridge data gaps in such a way that the noise floor is not substantially increased. Here is not the place to discuss the peculiarities of the known methods. The exceptionally

(A) 28. January 2019, email: herbertweidner@gmx.de

high frequency resolution of only 10 nHz may seem exaggerated, but it is appropriate because vector FFT reduces the half width of spectral lines. If the frequency resolution is too low, the probability of overlooking narrow spectral lines increases.

The next steps in a spectral analysis are elimination of the constant bias of the signal and the attenuation of the beginning and end of the measurement series by a pointwise multiplication with a bell-shaped function. If you do that, only the central part of the measurement series remains almost unchanged. This "windowing" means a loss of precious data points and a reduction of the frequency resolution. Another consequence is that the generated spectrum refers to the middle part of the data series. Details near the beginning and end of the sample are more or less ignored. This process can also be described in terms of telecommunications: "Windowing" is a (very low-frequency) amplitude modulation of all the signals contained in the data record. Each modulation *generates* additional frequencies which broaden each spectral line (sidebands). The sum of all additionally generated sidebands raises the noise level. However, as we are looking for extremely weak signals, any worsening of the SNR should be avoided. Therefore, "windowing" should be avoided. This fact is also presented in window disadvantages. When reading noise bandwidth, it must also be considered that all examples assume an SNR = 20 dB or higher. The following discussion, however, deals with finding spectral lines when the SNR is estimated to be 0 dB or even lower. We do not need any additional noise.

The impact of "windowing" is best illustrated by an example, the real recording of the superconducting gravimeter in Canberra in the years 1997 to 2000. The recording length includes 1,200,000 readings every 60 seconds. There are no significant earthquakes during this period. Using a "window", the spectrum preferably refers to the central part of the sample. Without a window, all measurement points contribute equally to the result. More data points mean higher frequency resolution. After the usual preparations, the sampling rate of the data is decimated to 3600 seconds, sufficient for spectral analysis in the frequency range below 100 µHz.

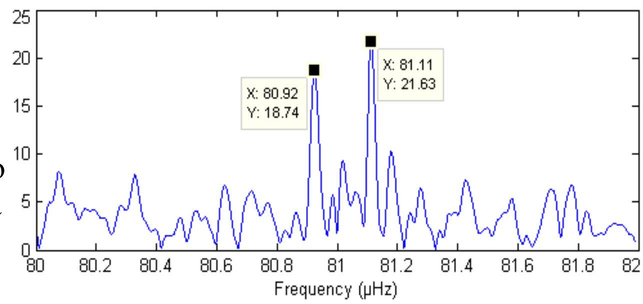

Fig 1: Spectrum of a sample of 20,000 data points that have passed through a blackman filter.

Figure 1 shows the spectrum of a narrow frequency range after the data has passed through a Blackman filter (Zero-padding was added to smooth the presentation). Two spectral lines can be safely located, the other peaks are probably part of the background noise. The amplitude comparison with the environment gives the SNR = 7.3 = 17 dB.

Turning off the Blackman window results in the spectrum shown in Figure 2. Since *all* data points contribute equally to the result, the half-width of the spectral lines decreases and it becomes easier to separate them. Now a third spectral line is clearly visible and the SNR has improved slightly, it is now 9.9 or 19.9 dB. Conclusion: If you look for weak spectral lines in very noisy time series, the SNR can be improved, if you do without "windowing".

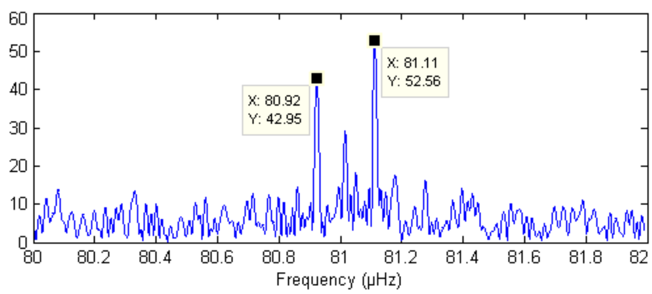

Fig 2: Spectrum as in Fig 1, but without blackman filter.

Vector-FFT

Every discovery of a spectral line stimulates the question: Are the corresponding oscillations invariable or can they only be measured during a certain period of time? In the above examples, all data points are processed as a single block, therefore no time dependence of the underlying oscillations (variable amplitude or frequency) can be determined. In order to discover a time dependence, the recording must be split into several smaller sections (called frames), which may overlap.

How many data points should one frame contain? If one chooses long sections, the time resolution decreases. If one chooses short sections, the frequency resolution decreases. Is it possible to keep these disadvantages down?

Yes. And the decisive hint can be seen by adding many individual spectra, usually called "stacking". The small number of data points in each section causes insufficient frequency resolution. Disturbing is the significant increase in the mean noise level, which worsens the SNR. Stacking the results gained by standard FFT is therefore unsuitable for detecting extremely weak spectral lines with SNR ≈ 1.

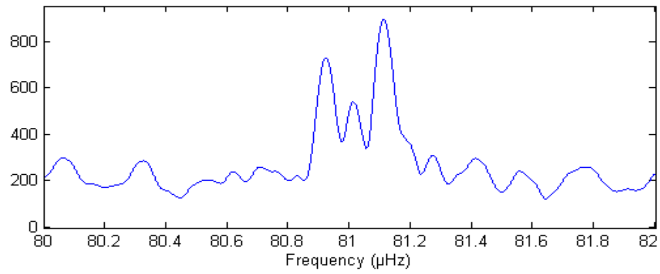

Fig 3: The sum of 700 time shifted single spectra. Each section contains 120 data points.

The spectra can also be summed up differently, because the FFT of each frame provides a complex result for each individual frequency. The real part indicates the amplitude of the COS oscillation, the imaginary part the amplitude of the SIN oscillation. In standard FFT stacking, only the magnitudes are summed up and the phase is ignored. As shown in figure 3, this simplification significantly reduces the signal-to-noise ratio and the frequency of the spectral lines can be more guessed than measured.

A quick test provides clarity: one feeds the "FFT machine" with noise, that is, with different sequences of random numbers, and notes the complex results, which are calculated successively for a freely chosen frequency. The left column of the adjacent table shows a possible result, a lot of complex numbers. The right column contains the magnitude of the neighbor left.

1	2
-56.8888 + 0.0000i	56.8888 + 0.0000i
-11.4095 + 58.9500i	60.0439 + 0.0000i
15.4662 + 22.2878i	27.1284 + 0.0000i
5.2836 - 10.4402i	11.7010 + 0.0000i
15.3674 - 10.3641i	18.5356 + 0.0000i
-17.2677 - 15.1931i	23.0001 + 0.0000i
1.0927 + 22.3864i	22.4130 + 0.0000i
-14.3679 - 1.4838i	14.4444 + 0.0000i
11.6721 + 15.8366i	19.6733 + 0.0000i
-9.2220 - 1.2110i	9.3012 + 0.0000i
7.9365 - 10.0402i	12.7982 + 0.0000i
10.9135 + 20.5977i	23.3103 + 0.0000i
[illegible]	[illegible]

After 1000 FFT transformations, the table contains 1000 rows. To get the mean amplitude of the chosen frequency, you have to sum the values of a column. If the signal consists only of noise, one would expect the result to be zero. This almost applies to the left column, but is far missed for the right column. The result of the right column is getting more and more away from zero – regardless of whether one feeds the "FFT machine" with pure noise or with a noisy signal. This is exactly what happens with the standard FFT method. It's hard to distinguish between noise and a weak spectral line.

This incessant increase in the sum (right column) is neither caused by the preprocessing of the data nor by the FFT method itself. The cause is the calculation of the absolute value (magnitude) of a complex number before the addition. That step should be avoided.

The result of the left column will not be far away from value 0 + 0i (mathematically precise: this is described by a random walk). The more rows the table contains, the more likely it will be that there are two arrows that nearly equal each other. Or that three arrows yield a small sum when arranged like a Mercedes star.

Why is this method called Vector FFT? One can represent any complex number as an arrow in the complex plane. Each arrow can be described either by its real and imaginary component *or* by its magnitude and the angle with respect to the positive x-axis (argument).

If you draw all the arrows of the left column, you can distinguish three cases:

1. If the signal is pure noise, you get a star-like, more or less symmetrical figure. There is no discernible preferential direction and all vectors sum up to a rather small result.
2. This changes when we have a weak signal buried in noise and the selected frequency *matches* the oscillation frequency of the signal. Then there is a preferred direction, the arrow lengths in some direction are longer than the arrows in other directions. It's almost as if a wind were trying to turn the arrows into a certain direction. The direction itself is not decisive, because at the very end the magnitude of the total is calculated. Summing up all vectors, the result devates markedly from zero.
3. If you feed a noiseless signal into the "FFT machine" and if you chose the proper frequency , all the arrows point into the same direction and you get the maximum possible total result.

It would be too nice if these simple cases fully described the vector FFT method.

Moving frames

Vector FFT requires that a long data record is divided into many short frames, which may overlap. For each frame, FFT calculates the complex amplitudes for a given set of frequencies. In contrast to the usual method, the complex amplitudes are summed *without magnitude calculation* in order not to delete the phase information. Only at the very end, when the spectra of all frames are added, the magnitude is calculated to determine the relative strength of each given frequency. A hurdle must be overcome for the procedure to work. If a frame is shifted by x data points, this means a phase shift Δp for each of the given frequencies. This is $\Delta p = 2\pi x \cdot T_s \cdot f$.

T_s is the time interval between successive measurements. In our case, 60 s or 3600 s.

f is the desired spectral frequency.

After FFT has calculated the complex amplitudes of all predefined frequencies, the phase shift Δp must be subtracted *before* the complex amplitudes are summed. If one imagines the complex amplitude as an arrow, each arrow must be rotated clockwise by the angle Δp. With complex numbers, this can be done very easily by multiplying with a unit vector, whose argument is Δp.

Reduction of half-width

The FFT algorithm works with precisely defined frequencies, which can be calculated from the sampling frequency and the length of a sample incl. zero-padding (for example 2^{15}). It does not calculate results for intermediate frequencies. However, the actual frequencies occurring in the input data sets will generally deviate from this set of predefined frequencies. If one observes how the phase changes over time, one can determine whether a spectral line coincides with or deviates from one of the given frequencies. Figure 4 shows an example in which not the amplitudes, but the phases of three adjacent frequencies are registered as a function of time.

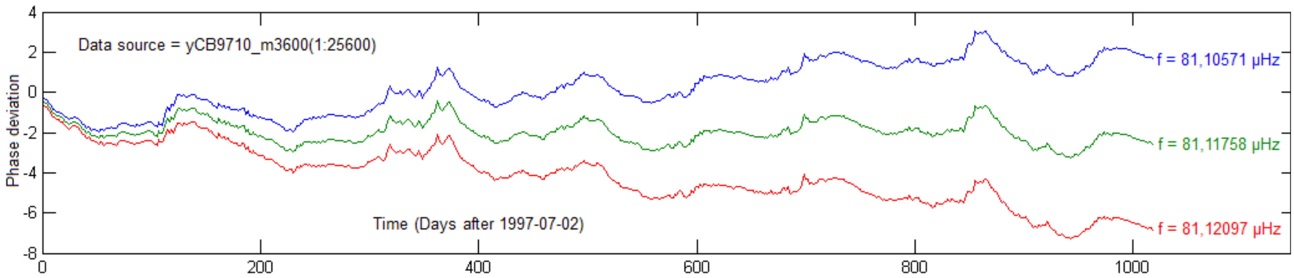

Fig 4: The phase change as a function of time for three closely adjacent frequencies

Obviously, all three individual frequencies belong to a single spectral line whose phase (and frequency) fluctuates slowly. The phase of the green curve has the lowest drift and is therefore the best estimate of the spectral frequency. Of particular interest is that the phase of the red curve has decreased by about 2π within 1000 days. During this period 612 individual spectra were calculated. Chaining all 612 red arrows in a complex plane results in a deformed but closed circle. So the vector sum of all arrows is approximately zero and in the graphical representation of the spectral line this means a zero, which leads to a remarkable reduction of the half width of that spectral line. This is a peculiarity of the vector FFT, because with the usual FFT analysis it is impossible to get the value zero by adding up the magnitudes of many non-zero vectors.

If one discovers longer periods of time in which the phase is approximately constant, this is a more convincing proof for a weak spectral line than a weak relative maximum in the usual spectral representation.

The next two figures show the increase in resolution and the reduction of the noise level, when standard FFT is replaced by Vector-FFT.