

## FFT method and how to calculate DFTs in Excel (branko@pecar-uk.com)

Let's make a few quick statements first. FFT stands for Fast Fourier Transforms and it is an algorithm, or a method, of calculating very quickly and efficiently a set of Discrete Fourier Transforms (DFT). So do not confuse the two. One is the method and the other one is the result. Needless to say, we could be interested in continuous Fourier transforms too, but this is not the topic of this tutorial. What else do we need to know about FFT? Not much, except that using this method you can calculate only  $2$  to the power of  $n$  ( $2^n$ ) of these transforms. This means that the number of these discrete Fourier transforms (DFTs) that you can calculate in Excel using the FFT method is 2, 4, 8, 16, 32, 64, 128, 256, ..., 4096, etc. In fact, Excel will allow a max of 4096 DFTs. Sometimes we use the word "coefficients" instead of DFTs, so do not be confused.

OK, what will the final result look like and how do we achieve this? Here is a graph (Fig. 1) of average monthly temperatures in Sheffield, South Yorkshire in degrees C between January 1977 and August 2019 (courtesy of: <http://www.sheffieldweather.co.uk/Averages/MONTHLYAIRAVERAGE.htm>). The graph is shown in time domain and it shows months on the x-axis and average temperatures on the y-axis. It consists of 512 temperature readings.

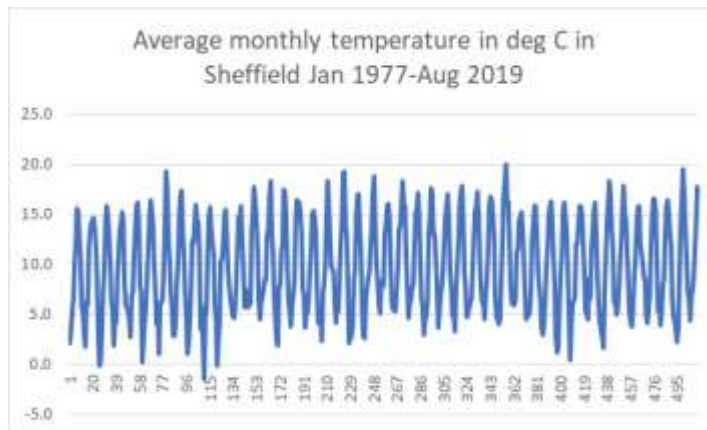


Fig. 1

The very same variable, i.e. the average monthly temperatures in Sheffield from Jan 1977 until Aug 2019 can be presented in the frequency domain. The graph below (Fig. 2) is the one we calculated. The x-axis now shows the frequencies and the y-axis shows the spectral densities for every frequency. We will explain how to calculate all of this, as well as what all the expressions mean.

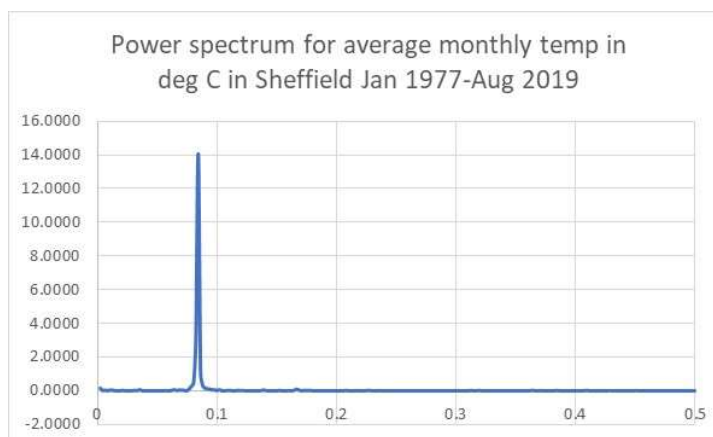


Fig. 2

We will now show a formula that is used to calculate the DFTs, but you do not have to worry about it. Excel will do it for you. Here it is, just for illustration purposes:

$$F_k = \sum_{n=0}^{N-1} x_n e^{\frac{-i2\pi nk}{N}} \quad (1)$$

Where:

$F_k$  = Discrete Fourier transform coefficients (DFT) expressed as complex numbers

$x_n$  = Observations (i.e. the function or the time series)

$k$  = Number of complex coefficients

$N$  = Number of observations

$n$  = Number of samples  $n = 0$  to  $N-1$  (same as  $k$ )

$i$  = imaginary number

This formula expresses discrete Fourier transforms as complex numbers. There are other ways to achieve the same, but we will stick to this one, because Excel relies on it. The only thing that you need to remember is that the result of this formula will be a series of numbers and that this series will be 2, 4, 8, 16, ..., 4096 long. The only other piece of information important is that that these coefficients are complex, which means that they consist of a real number and imaginary number, something like this:  $5-6i$ , or  $2.145+0.43i$ . If you do not understand this, get some quick lessons from the web, though you do not have to do anything with these numbers in this tutorial.

OK, just a few more points before we start the calculations. If, for example, we have 128 data points (like 128 temperature readings), we can calculate up to 128 DFTs, but not more. We could calculate less (64 or 32 for example), but that's not the point here. So, remember that the number of observations will determine how many DFT coefficients you can calculate.

Here is an important point. If you have 128 observations (or any other number), you do not actually need all 128 DFTs to get the correct picture. You can calculate only half of them. Exactly at half point, which is called the folding frequency, or Nyquist limit, the DFTs begin to show the mirror image of the first half. This means that it is sufficient to calculate only half the number of the coefficients. Fig. 3 illustrates the point. Once you passed the middle point, you do not get any new information, just the repeat of the first half.

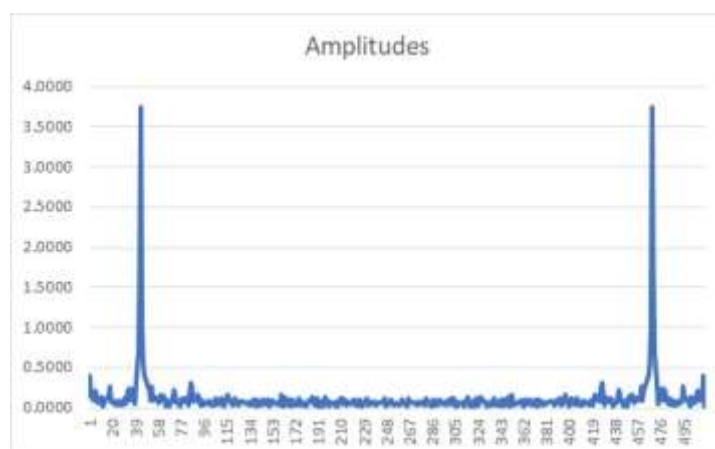


Fig. 3

And just one last point. What if you have only 123 observations and not 128 and you really want 128 DFTs? One option is to take only the last 64 observations and calculate 64 DFTs for them and ignore all the earlier observations. Unfortunately, this will give you only 32 meaningful coefficients (see above paragraph). However, what if you really want 128 DFT coefficients that will give you 64 meaningful DFTs? There is a workaround and it is called zero padding. In this case just add 5 zeros below your 123 observations so that your times series consists now of 128 observations. Despite these zeros artificially added to your time series, you will still get the correct values of the DFTs even if your time series is shorter than the number of DFTs you need.

So, what are we going to do in this tutorial? We will first calculate DFTs using Excel add-in for Fourier Analysis. After that we will calculate the amplitudes from every coefficient, i.e. discrete Fourier transform. The amplitudes will then be used to calculate the power coefficients. And finally, we will calculate the frequency bins and present our results in a graph. That's it. Let's start.

### 1. Calculate DFTs using Excel Fourier Analysis from Data Analysis in the Data tab

First of all, I will assume that you have installed Excel Data Analysis add-in. If you did, you will find it under the Data tab. Click on Data Analysis option, which will open a dialogue box to select Fourier Analysis option (see Fig. 4):

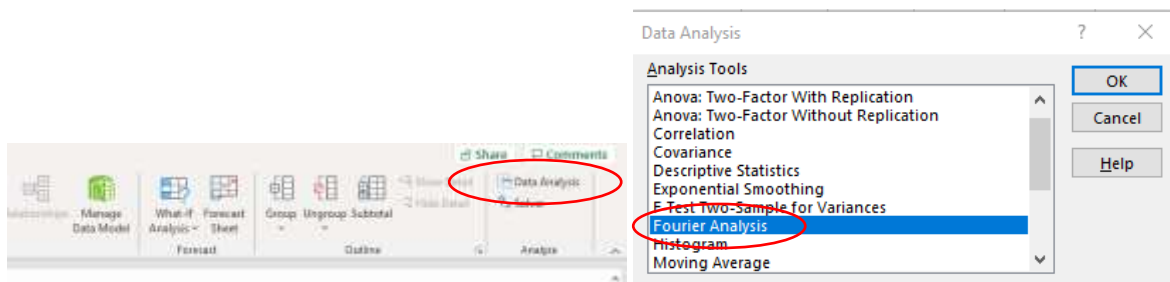


Fig. 4

The option presented in the next box is simple (see Fig. 5). Just select the time series (in our case cells C5:C516) and where you want the DFTs to be printed out, which in our case is cell D5.

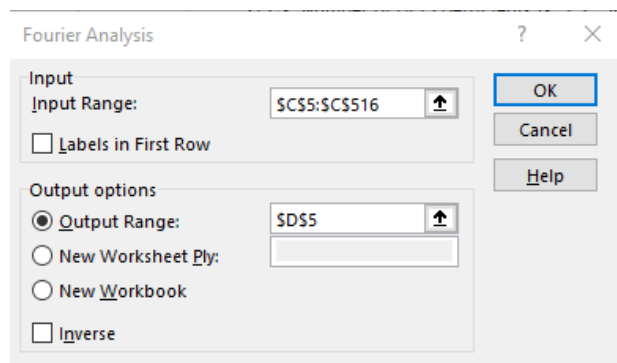


Fig. 5

Once you clicked OK, Excel will immediately calculate all 512 discrete Fourier transforms, or as we call them, coefficients. Fig. 6 below shows the results.

	A	B	C	D
1	Average monthly temperature in deg C in Sheffield, South Yorkshire			
2	<a href="http://www.sheffieldweather.co.uk/Averages/MONTHLYAIRAVERAGE.htm">http://www.sheffieldweather.co.uk/Averages/MONTHLYAIRAVERAGE.htm</a>			
3				
4	<b>Year</b>	<b>Month</b>	<b>Avg Temp</b>	<b>DFTs</b>
5	1977	JAN	2.1	5009.2
6		FEB	3.6	-111.241337890715+95.165798425201i
7		MAR	5.4	39.9293780611367+25.7866843139007i
8		APR	6.6	41.0666825545464+61.5204529307675i
9		MAY	10.0	-20.9891023681386+22.9046621188411i
10		JUN	12.2	-66.2938048939289+31.9307554472146i
11		JUL	15.6	55.2885118878217+54.2852348161773i
12		AUG	15.3	41.2455604802376+17.9560116942954i
13		SEP	12.8	-9.95655696142417-17.2864870279436i
14		OCT	10.7	-14.2853446173828+15.0904501339005i
15		NOV	5.7	40.6442948071919+17.4436301282536i
16		DEC	5.3	14.8762916388729+46.5078302232696i
17	1978	JAN	2.9	-0.146423312767434-5.46842847862301i
18		FEB	1.8	-20.3990282634228+6.17084033053154i
19		JUL	19.5	24.725332044646+16.8173834333594i
20		AUG	17.0	-20.3990282634227-6.17084033053167i
21		SEP	13.4	-0.146423312767407+5.46842847862261i
22		OCT	10.3	14.8762916388724-46.5078302232706i
23		NOV	7.4	40.6442948071917-17.4436301282535i
24		DEC	6.4	-14.2853446173828-15.0904501339003i
25	2019	JAN	4.4	-9.95655696142408+17.2864870279437i
26		FEB	6.9	41.2455604802377-17.9560116942959i
27		MAR	7.5	55.2885118878214-54.2852348161777i
28		APR	8.9	-66.293804893929-31.9307554472144i
29		MAY	11.5	-20.9891023681387-22.9046621188411i
30		JUN	14.1	41.066682554546-61.5204529307681i
31		JUL	17.8	39.9293780611365-25.7866843139011i
32		AUG	17.1	-111.241337890716-95.1657984252006i

Fig. 6

Column A are the years from 1977 until 2019, column B are the months, column C are the average temperatures and column D are newly calculated DFTs. Note that in Fig. 6 we are showing only the first 14 observations and the last 14 observations.

Also, take a look at the first DFT beyond the zero-frequency (we'll explain this shortly) in cell D6. What you see is a complex number that reads  $-111.24+95.16i$ . The second one in D7 is  $=39.93+25.79i$ , etc. Complex numbers are in fact two dimensional numbers that reside in a space defined by the real axis and the imaginary axis. Effectively, every number is a vector. This means that it has a direction and magnitude. How do we calculate a magnitude of a vector? Well, it is calculated as the modulus of the complex number.

$$|z| = \sqrt{a^2 + b^2} \quad (2)$$

Every complex coefficient in column D will consist of the real part  $A_k$  and the imaginary part  $B_k$ . This means that using equation (2) the magnitude for every complex coefficient can be calculated as  $M_k$ :

$$M_k = |F_k| = \sqrt{A_k^2 + B_k^2} \quad (3)$$

## 2. Calculate the amplitudes from every DFT

If the values in column D are DFTs, whose symbol is  $F_k$ , then in column E we can calculate magnitudes for every coefficient. However, because we are calculating the magnitudes for every frequency (to be explained shortly), we will not use the word magnitude, but we'll call them the amplitudes. The formula for amplitudes of a DFT is:

$$\text{For } k=0 \quad A_k = \frac{1}{N} |F_k| = \frac{M_k}{N} \quad (4)$$

$$\text{For } k>0 \quad A_k = \frac{\sqrt{2}}{N} |F_k| = \frac{M_k\sqrt{2}}{N} \quad (5)$$

Fig. 7 shows how to calculate the amplitudes from the DFTs.

	A	B	C	D	E	F
1	Average monthly temperature in deg C in Sheffield, South Yorkshire					
2	<a href="http://www.sheffieldweather.co.uk/Averages/MONTHLYAIRAVERAGE.htm">http://www.sheffieldweather.co.uk/Averages/MONTHLYAIRAVERAGE.htm</a>					
3						
4	<b>Year</b>	<b>Month</b>	<b>Avg Temp</b>	<b>DFTs</b>	<b>Amplitudes</b>	
5	1977	JAN	2.1	5009.2	9.7836	=IMABS(D5)/\$K\$1
6		FEB	3.6	-111.241337890715+95.165798425201i	0.4044	=SQRT(2)*IMABS(D6)/\$K\$1
7		MAR	5.4	39.9293780611367+25.7866843139007i	0.1313	=SQRT(2)*IMABS(D7)/\$K\$1
8		APR	6.6	41.0666825545464+61.5204529307675i	0.2043	=SQRT(2)*IMABS(D8)/\$K\$1
9		MAY	10.0	-20.9891023681386+22.9046621188411i	0.0858	=SQRT(2)*IMABS(D9)/\$K\$1

Fig. 7

We divide the first coefficient in D5 (which isn't a complex number) with the number of observations (512 in our case) and we get the zero-frequency amplitude. If we multiply the second one in D6 by the square root of 2, and all the subsequent ones, and divide each of them by the number of observations, we get the amplitudes for all other DFTs.

Two points to clarify. The first one is why do we multiply all non-zero frequency amplitudes with square root of 2? Because we only need half of them up to the folding frequency, or Nyquist limit, as we explained earlier. This is the formula to get the single sided spectrum (remember, the full spectrum consists of two mirror images of coefficients).

The second question is why are we using Excel function called =IMABS()? Because this is a dedicated function to complex numbers. Look at the equations (2) or (3) above and you will see that the modulus of a complex number is an absolute square root value of the squared sum of the two components. This is exactly what =IMABS() function does.

## 3. Calculate the power coefficients

Now imagine something. Think about our time series of 512 observations as a function. If we convert this function from the time domain into the frequency domain, which is what we are doing, then once we completed this transformation, the frequency domain will show us how much every frequency contributes towards the total energy (or power) of this function. Another way to say this is to ask yourself: what is the individual power of every frequency when compared to the total energy of the function? In other words: how is the total power distributed per frequency, or what is the power spectrum?

OK, this means that from our amplitudes we need to calculate the power coefficients. The formula for this is:

$$\text{For } k=0 \quad PSD_0(F_k) = \frac{A_k^2}{2} \quad (6)$$

$$\text{For } k>0 \quad PSD_x(F_k) = A_k^2 \quad (7)$$

Again, we treat the power for the zero-frequency differently from the rest of the spectrum. In Excel, the calculations are executed as in Fig. 8:

	A	B	C	D	E	F	G	H
1	Average monthly temperature in deg C in Sheffield, South Yorkshire							
2	<a href="http://www.sheffieldweather.co.uk/Averages/MONTHLYAIRAVERAGE.htm">http://www.sheffieldweather.co.uk/Averages/MONTHLYAIRAVERAGE.htm</a>							
3								
4	Year	Month	Avg Temp	DFTs	Amplitudes		Power	
5	1977	JAN	2.1	5009.2	9.7836	=IMABS(D5)/\$K\$1	47.8594	=(E5^2)/2
6		FEB	3.6	-111.241337890715+95.165798425201i	0.4044	=SQRT(2)*IMABS(D6)/\$K\$1	0.1635	=E6^2
7		MAR	5.4	39.9293780611367+25.7866843139007i	0.1313	=SQRT(2)*IMABS(D7)/\$K\$1	0.0172	=E7^2
8		APR	6.6	41.0666825545464+61.5204529307675i	0.2043	=SQRT(2)*IMABS(D8)/\$K\$1	0.0417	=E8^2
9		MAY	10.0	-20.9891023681386+22.9046621188411i	0.0858	=SQRT(2)*IMABS(D9)/\$K\$1	0.0074	=E9^2

Fig. 8

The power spectrum is in column G. As you look further down this column (not shown here, but check the spreadsheet), you will realise that we stopped calculating power coefficients in cell G261. Take a look at a neighbouring cell D261. You will notice that column D consists of complex DFT numbers until we hit D261. This is the first coefficient that it is not a complex number, but it is a real number. As it happens, this is the point exactly half way through the number of observations and it will be the point where the Nyquist limit applies, or as we call it, a folding frequency. Let's see now how to calculate the frequencies that are associated with the power coefficients we just calculated.

#### 4. Calculate the frequency bins

The power coefficients we just calculated need to be attributed to specific frequencies. How are these frequencies calculated? First of all, we need to start with how many observations we have (N) and what is the total time over which we collected these observations (T). From there, we can calculate the time increments ( $\Delta t$ ) and the sampling rate ( $f_s$ ). Once we have this, we can calculate the folding frequency ( $f_f$ ) and the frequency increments ( $\Delta f$ ). These frequency increments will tell us the width of every frequency bin.

Before we show these trivial calculations in Excel, let's see a series of very simple equations that explain mutual inter-relationships between all these parameters we just mentioned. The parameters are:

- N = Total number of discrete data points taken
- $\Delta t$  = Time between data points (time increments)
- T = Total sampling time
- $f_s$  = Sampling rate (how often do we take samples)
- $f_f$  = Folding frequency or Nyquist limit
- $\Delta f$  = Frequency increments that will create frequency bins

We'll start with the time increments  $\Delta t$  in which the time series is recorded as:

$$\Delta t = \frac{T}{N} = \frac{1}{f_s} \quad (8)$$

From there, we can extract any of the variables:

$$T = N\Delta t = \frac{N}{f_s} \quad (9)$$

$$N = \frac{T}{\Delta t} = f_s T \quad (10)$$

$$f_s = \frac{1}{\Delta t} = \frac{N}{T} \quad (11)$$

The folding frequency (Nyquist limit frequency) is defined as:

$$f_f = \frac{f_s}{2} = \frac{1}{2\Delta t} = \frac{N}{2T} \quad (12)$$

The frequency increments  $\Delta f$ , which will be used to create frequency bins are defined as:

$$\Delta f = \frac{1}{T} = \frac{1}{N\Delta t} = \frac{f_s}{N} = \frac{2f_f}{N} \quad (13)$$

As we already know, according to the Nyquist criterion, the maximum number of DFTs that should be calculated should not exceed  $f_s/2$ . This means that there is a number  $k$ , which we multiply with the number of frequency increments  $\Delta f$ , i.e.  $k\Delta f$ , which should equal to  $f_s/2$ .

We define  $k\Delta f$  as:

$$k\Delta f = \frac{f_s}{2} \quad (14)$$

From (14), we can extract  $k$  as:

$$k = \frac{f_s}{2\Delta f} = \frac{\frac{N}{T}}{2\frac{1}{T}} = \frac{N}{2} \quad (15)$$

From there we can provide alternative equations for (12) as:

$$f_f = \frac{f_s}{2} = \frac{N}{2} \Delta f = k\Delta f \quad (16)$$

Lot of equations, but the calcs are very simple. Let's see what we have in our spreadsheet.

	K	L	M	N	O	P	Q	R	S	
1	512	N=Number of observations							=COUNTA(C5:C516)	
2	512	k=Number of DFT coefficients (k=1,2,...,N-1) --> Has to be 2 <sup>n</sup>						=K1		
3	512	T=Total time (number of months in this case)						=K1		
4	1.00	Δt=Time increments in months						=K3/K1		
5	1.00	f <sub>s</sub> =Sampling rate per months						=1/K4		
6	0.5	f <sub>f</sub> =Folding frequency						=K5/2		
7	0.0020	Δf=Frequency increments per month						=(2*K6)/K1		

Fig. 9

The cells show:

$K1 = \text{COUNTA}(C5:C516) = 512$ ,       $N = \text{Number of observations}$   
 $K2 = K1 = 512$ ,                               $k = \text{Number of DFT coefficients}$   
 $K3 = K1 = 512$ ,                               $T = \text{Total time (number of months in this case)}$   
 $K4 = K3/K1 = 1.00$ ,                         $\Delta t = \text{Time increments in months}$   
 $K5 = 1/K4 = 1.00$ ,                         $f_s = \text{Sampling rate per months}$   
 $K6 = K5/2 = 0.5$ ,                          $f_f = \text{Folding frequency}$   
 $K7 = (2 * K6)/K1 = 0.0020$ ,          $\Delta f = \text{Frequency increments per month}$

We just used some of the equations (8) to (16) to do these simple calculations. The others are there if you need them.

OK, now we have  $\Delta f$ , we can calculate the frequency bins. In other words, the total spectrum of frequencies is in our case divided by 512 “bins”. The null frequency will be zero and every other frequency will be incremented by the value of  $\Delta f$ . What we mean is:

$\Delta f = 0.0020$        $f_0 = 0$   
 $f_1 = f_0 + \Delta f = 0 + 0.0020 = 0.0020$   
 $f_2 = f_1 + \Delta f = 0.0020 + 0.0020 = 0.0039$   
 $f_3 = f_2 + \Delta f = 0.0039 + 0.0020 = 0.0059$   
. . .  
 $f_{257} = f_{256} + \Delta f = 0.4980 + 0.0020 = 0.500$

Let’s show this in Excel (Fig. 10) and point out a few key elements. Column I contains the frequency bins calculated as above and we stop in cell I261, which is the point where our frequency bin reaches the value of 0.5. This is specified according to cell K6, i.e. the folding frequency.

Here is one very important point. If you take a look at the column G where we calculated the power coefficients, you will see that, excluding the power coefficient for the zero frequency, the largest power coefficient is in the cell G48 (the value is 14.0794). This power coefficient is associated with the frequency of 0.0840 (cell I48 in Fig. 10). If you take the inverse value of this number ( $=1/0.0840=11.91$ ), you get 11.91. If we round it up, we get the number 12. Effectively we have shown that our time series has the seasonality of 12 months, which is what we knew from the beginning and why we picked this time series for demonstration in the first place.



Why did we get 11.91 and not 12? Because we only have 512 frequency bins, i.e. not very fine resolution. If we increased the number of DFTs and the associated frequencies to 1024 or 2018, or even 4096, we would get closer to this whole number of 12.

	G	H	I	J	K	L
1					512	N=Number
2					512	k=Number
3					512	T=Total ti
4	Power		Frequencies		1.00	$\Delta t$ =Time i
5	47.8594	=E5^2)/2	0		1.00	$f_s$ =Sampli
6	0.1635	=E6^2	0.0020	=I5+\$K\$7	0.5	$f_f$ =Folding
7	0.0172	=E7^2	0.0039	=I6+\$K\$7	0.0020	$\Delta f$ =Freque
8	0.0417	=E8^2	0.0059	=I7+\$K\$7		
9	0.0074	=E9^2	0.0078	=I8+\$K\$7		
10	0.0413		0.0098			
11	0.0458		0.0117			
12	0.0154		0.0137			
13	0.0030		0.0156			
14	0.0033		0.0176			
15	0.0149		0.0195			
16	0.0182		0.0215			
17	0.0002		0.0234			
18	0.0035		0.0254			
19	0.0068		0.0273			
41	0.0314		0.0703			
42	0.0290		0.0723			
43	0.0027		0.0742			
44	0.0977		0.0762			
45	0.2574		0.0781			
46	0.4873		0.0801			
47	3.0576		0.0820			
48	14.0794		0.0840		11.9070	=1/I48
49	1.0233		0.0859			
50	0.2827		0.0879			
51	0.1582		0.0898			

Fig. 10

What if we wanted to express all these parameters not in terms of months but as fractions of a year? Well, we did this too. The figure below (Fig. 11) shows the calculations. Note that now we are expressing everything in years, rather than months. For example, T is now 42.67 (cell W3) which means that we collected data over 42.67 years. The time increment is now  $\Delta t=0.0833$  of a year (cell W4), which means that we collect data at the sampling rate of  $f_s=12$  per year (cell W5). This gives us the folding frequency of  $f_f=6$  (cell W6) and the frequency increments of  $\Delta f=0.0234$  (cell W7).

Although we changed the units, nothing else has changed. If we look again the frequency associated with the largest power in column G (Fig. 10), we still see that the highest power coefficient in cell G48, corresponding to yearly frequency in U48. The value of this frequency is 1.0078 (Fig. 11). Again, an inverse value of this number is 0.992248, which is close to 1. This means that our cycle repeats itself every year, which we know, but here we have a proof for this. It works regardless what units of measure you take, as long as you pay attention to it.

	U	V	W	X	Y	Z	AA	AB	AC
1			512	N=Number of observations					
2			512	k=Number of DFT coefficients (k=1,2,...,N-1) --> Has to be 2 <sup>n</sup>					
3			42.67	T=Total time (number of years in this case)					=W1/12
4	Frequencies		0.0833	$\Delta t$ =time increments in years					=W3/W1
5	0		12	$f_s$ =sampling rate per year					=1/W4
6	0.0234	=U5+\$W\$7	6	$f_f$ =Folding frequency					=W5/2
7	0.0469	=U6+\$W\$7	0.0234	$\Delta f$ =Frequency increments per year					=(2*W6)/W1
8	0.0703	=U7+\$W\$7							
9	0.0938	=U8+\$W\$7							
10	0.1172								
11	0.1406								
12	0.1641								
13	0.1875								
14	0.2109								
15	0.2344								
16	0.2578								
17	0.2813								
18	0.3047								
19	0.3281								
41	0.8438								
42	0.8672								
43	0.8906								
44	0.9141								
45	0.9375								
46	0.9609								
47	0.9844								
48	1.0078		0.992248	=1/U48					

Fig. 11

We could have used examples with Hz, which are the favourite measuring units of the engineering community. They are as easy to handle as other units, as long as you remember that 1 Hz means one sample per second. Let's take a trivial example with only 8 observations measured over 2 seconds to explain the principle:

$N = 8$  observations

$T = 2$  sec

$\Delta t = T/N = 2/8 = 0.25$  sec

$f_s = 1/\Delta t = 1/0.25 = 4$  Hz

$f_f = f_s/2 = 4/2 = 2$  Hz maximum frequency value for calculating power coefficients

$\Delta f = 1/T = 1/2 = 0.5$  a frequency increment in Hz

So, the frequency increments will be 0.5 Hz, starting from zero, and will go to a max value of 2 Hz.

OK, so let's show one more time our final result, which is the plot of the power spectrum, expressed per frequency, for our average monthly temperature measures for Sheffield over 42.67 years, or 512 months from January 1977 until August 2019 (Fig. 12).

Note that to create a graph as in Fig. 12 you need to put frequency bins in your spreadsheet in a column before the power coefficients. Then you use a scatter diagram to construct the graph, where the dots are connected by a line. This will give you the graph with frequencies on the x-axis and power coefficients on the y-axis. Also, always drop  $f_0$  and the associated power coefficient from this graph as it is just distorts the picture.

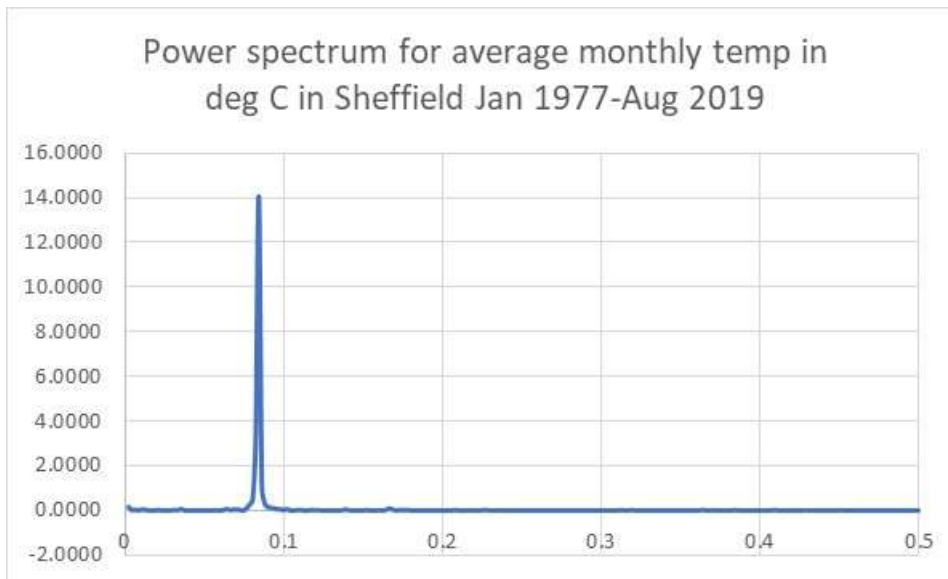


Fig. 12

We can see the power peaking just below 0.1, in fact at 0.084, which we know is  $(=1/0.084)$  11.91 or  $\approx 12$ , i.e. the periodicity of this time series. This could have been shown not as a function of frequency, but as a function of a period, in which case we would call it a periodogram.

One last cool thing. Imagine if we did not calculate the power coefficient for the zero frequency as the squared amplitude divided by 2. Let's say we calculated it as all other power coefficients, i.e. as just the squared value of the amplitudes. See cell G5 in Fig. 13 that we deliberately changed.

	A	B	C	D	E	F	G	H	I	J
1	Average monthly temperature in deg C in Sheffield, South Yorkshire						117.9135	=SUMPRODUCT(C5:C516,C5:C516)/COUNT(C5:C516)		
2	<a href="http://www.sheffieldweather.co.uk/Averages/MONTHLYAIRAVERAGE.htm">http://www.sheffieldweather.co.uk/Averages/MONTHLYAIRAVERAGE.htm</a>						117.9158	=SUM(G5:G261)		
3										
4	Year	Month	Avg Temp	DFTs	Amplitudes		Power	Frequencies		
5	1977	JAN	2.1	5009.2	9.7836 =IMABS(D5)/\$K\$1		95.7187 =E5^2	0		
6		FEB	3.6	-111.241337890715+95.165798425201i	0.4044 =SQRT(2)*IMABS(D6)/\$K\$1		0.1635 =E6^2	0.0020 =15+\$K\$7		
7		MAR	5.4	39.9293780611367+25.7866843139007i	0.1313 =SQRT(2)*IMABS(D7)/\$K\$1		0.0172 =E7^2	0.0039 =16+\$K\$7		
8		APR	6.6	41.0666825545464+61.5204529307675i	0.2043 =SQRT(2)*IMABS(D8)/\$K\$1		0.0417 =E8^2	0.0059 =17+\$K\$7		
9		MAY	10.0	-20.9891023681386+22.9046621188411i	0.0858 =SQRT(2)*IMABS(D9)/\$K\$1		0.0074 =E9^2	0.0078 =18+\$K\$7		

Fig. 13

If, after that, we sum up all these power coefficients (up to the folding frequency), then their sum is 117.91, as per cell G2 in Fig. 13. Take a look at cell G1 in Fig. 13. You will see the same number, but it is calculated as the normalised squared sum of all actual temperature readings in column C. In other words, take the square value of every temperature reading, add them all up and divide by the total number of them. You get the same number, i.e. 117.91. This shows you that in both the time space as well as the frequency space, the observations and their power coefficients are intrinsically connected. The power that the function (a time series) contains is the same whether presented in the time domain or in the frequency domain. Well, there is much more elegant mathematical explanation for this called Parseval's theorem, but it goes beyond this tutorial.

And finally, sometimes you will see the phrase "spectral density estimates", which most of the time is nothing but the power spectrum, as per our Fig. 12. In some cases, this means correcting the power coefficient values for the bandwidth capacity, but this is a typical engineering application and if you are not operating in these units, you can treat the two expressions interchangeably.

You might think that it is pointless discovering the periodicity of 12, especially as we already know that the monthly temperatures will repeat themselves every 12 months. However, very often you

will be unsure of the periodicity in advance, and the frequency analysis using discrete Fourier transforms will be of great help. Now you know how to do it.

Periodicity is not always so obvious, and it can hide a more complex patterns that only the harmonics from the power spectrum analysis will reveal. It remains to be said that Fourier transforms are the inverse function of the autocorrelation coefficients from the autocovariance function, so it does not matter whether you prefer the time or the frequency domain, the results will always yield the same conclusion.

This concludes our tutorial on how to calculate the DFTs, amplitudes, power spectrum and frequency bins using built in Excel functions.

Branko Pecar  
Autumn 2019