Frequency Analysis: Fast Fourier Transform

Fourier Series:

The possibility of a string vibrating in several of its harmonics at the same time was proven by **Daniel Bernoulli** in 1755 (Cannon & Dostrovsky 1981). Bernoulli developed his proof using the principle of the coexistence of small oscillations, which in modern terminology is referred to as the principle of superposition (Rao 2005, p.5). This principle made it possible to express any arbitrary function using an infinite series of sines and cosines. Initially, Bernoulli's work was not accepted by mathematicians **D'Alembert** and **Euler**. However, the validity of superposition was proved in 1822 (initially in 1807 and published in 1822) by **Joseph Fourier** who used the technique to solve the partial differential equation for the propagation of heat in solid bodies (University of St Andrews 1997). Prior to the findings made by Fourier there was no solution to the heat propagation equation for general conditions. However, particular solutions were known if the heat source behaved as a sinusoidal or co-sinusoidal wave. Fourier's work was based on the hypothesis that complicated heat sources could be modelled using the superposition of sinusoidal and co-sinusoidal waves, and that their effect could be modelled as the superposition of the resulting particular solutions. Although initially not well received by mathematicians such as **Lagrange** and **Laplace**, Fourier's work was eventually acknowledged and his superposition principle is now known as the Fourier series (University of St Andrews 1997).

Further development of Fourier's work has since provided an alternative to analysing signals in the time domain, namely frequency domain analysis.

Fourier Series: Any periodic function can be represented as the sum of a number of (co-)sinusoids

From: Lamb, M., 2011. *Monitoring the structural integrity of packaging materials subjected to sustained random loads* (Doctoral dissertation, Victoria University).

References:

Cannon, J & Dostrovsky, S 1981, *The evolution of dynamics: Vibration theory 1687 to 1742*, Springer-Verlag, New York

Rao, S 2005, *Mechanical vibrations*, SI Edition, Prentice Hall

University of St. Andrews 1997, *Jean Baptiste Joseph Fourier,* viewed 7 August 2008, < http://www-groups.dcs.st-and.ac.uk/~history/Printonly/Fourier.html>

Illustrative Example of Fourier Series:

By applying the Fourier transform, to extract the harmonics of a signal, it is possible to present the signal in a way that highlights its frequency content rather than its temporal content. If a third axis, which represents frequency, is introduced perpendicular to the time-amplitude plane each harmonic component can be displayed against its corresponding frequency. An example of the harmonic decomposition is given in *Figure 1*.

The harmonics shown in *Figure 1* can also be represented using a third axis which represents frequency as shown in Figure 2.

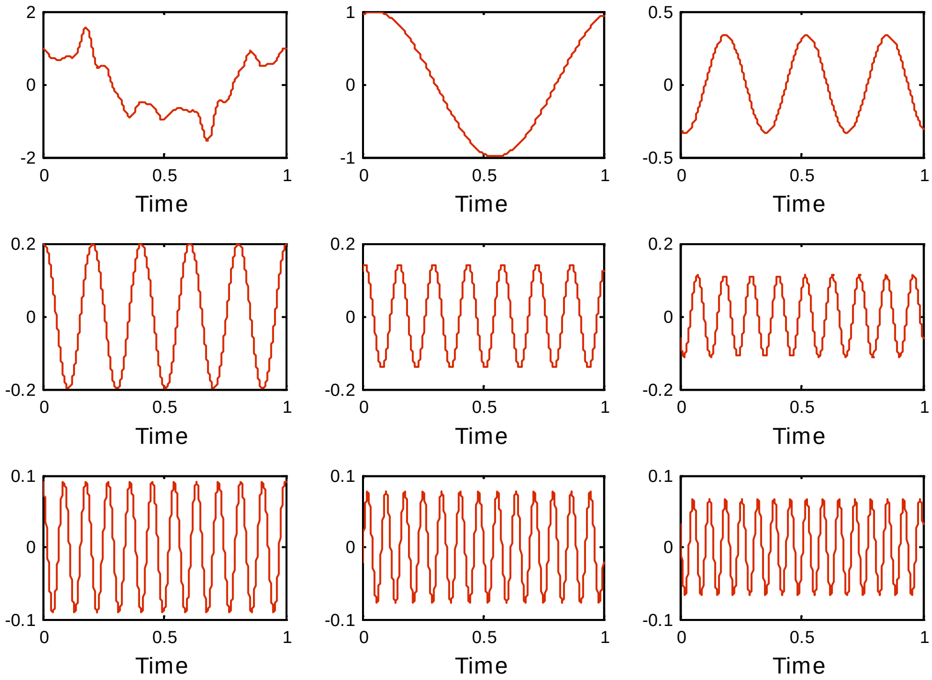

Figure 1: Random signal (top left) followed by its harmonic components.

When *Figure 2* is rotated so that the time axis is perpendicular to the page the signal is represented in the frequency domain. This procedure provides what is often referred to as a magnitude spectrum. The magnitude spectrum only displays the magnitude of the harmonic components of the signal, as shown in *Figure 3(a)*. The magnitude spectrum is a useful descriptor of random signals; however, both the magnitude and phase of the harmonic components are required to fully describe the signal. A phase spectrum, as in *Figure 3(b)*, is obtained by converting the complex signal obtained using the Fourier transform to its polar form and displaying the extracted phase values against their corresponding frequency.

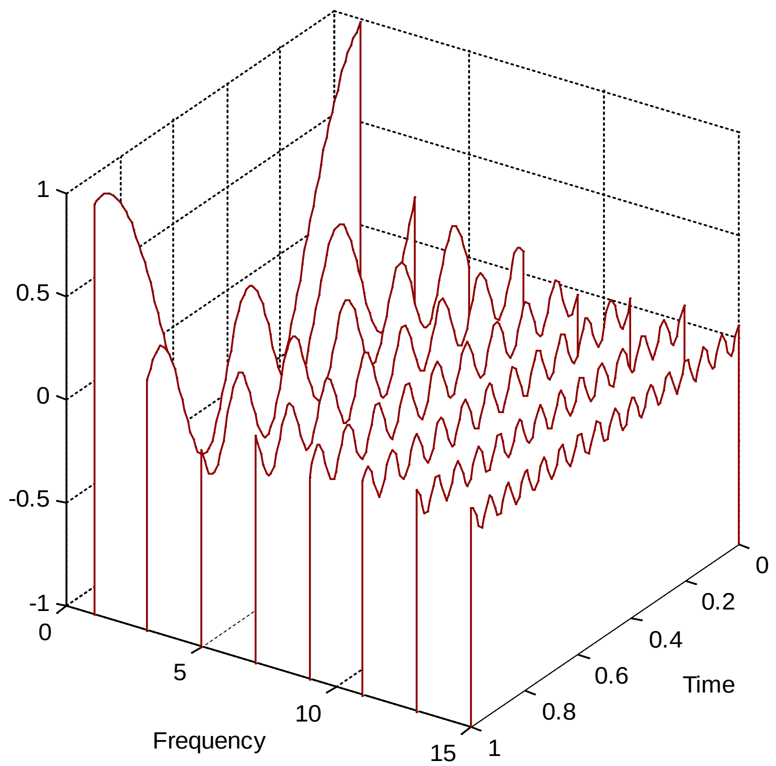

Figure 2: Time-Frequency-Amplitude representation of the harmonics of a random signal.

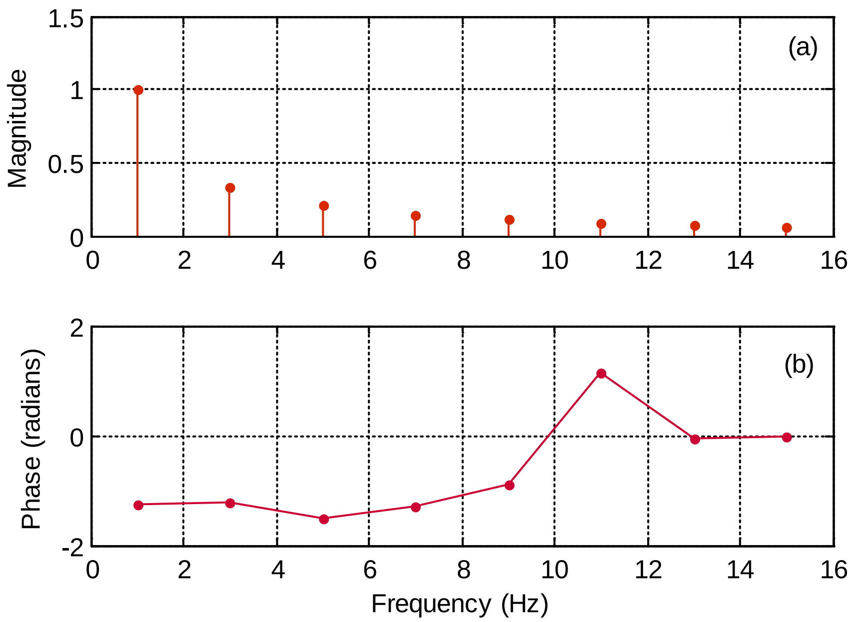

Figure 3: (a) Magnitude spectrum, (b) Phase spectrum, for the time signal shown in Figure 1.

Lab Exercise 1:

Develop an M-script capable of producing 6 different co-sinusoids with differing amplitudes (x_o), frequencies (f) and initial phases (ϕ_o) and sum them together to produce a new signal. You will need to develop a time vector {t } and 6 variations according to the equation $x(t) = x\cos(2\pi ft + \phi_o)$

Examples (Dr Michael Sek 2016)

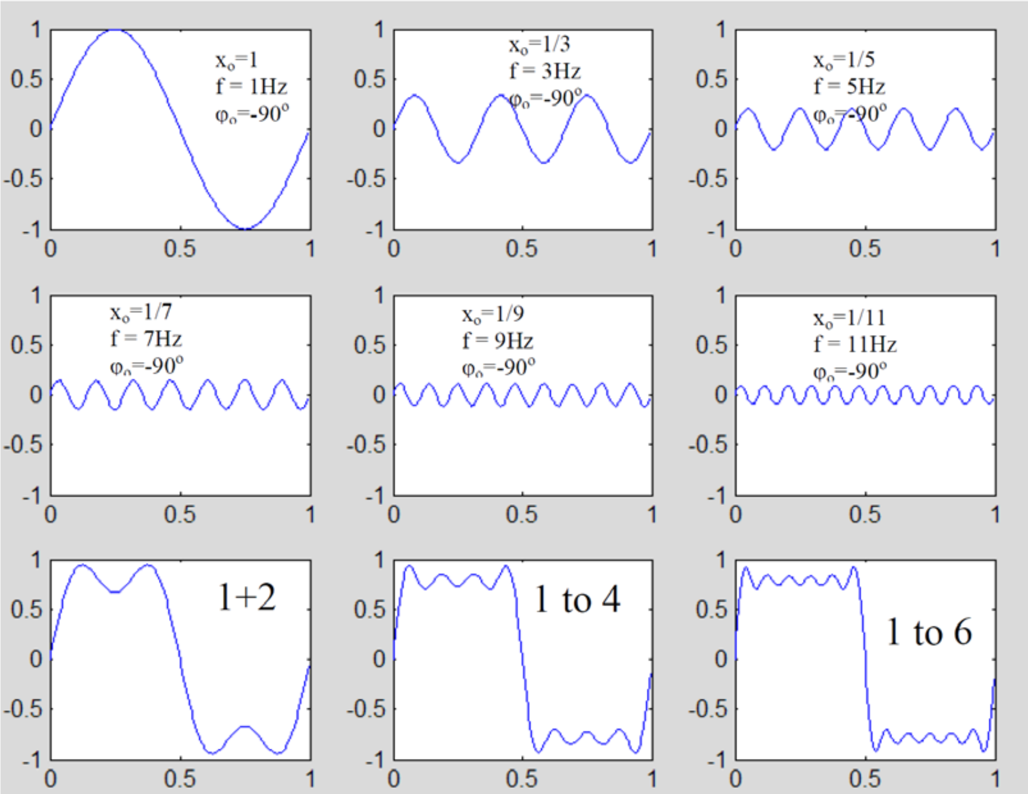

Figure 4: Developing a square wave from co-sinusoids. ϕ_o is -90° for all co-sinusoids. If the pattern is continued up to 100 co-sinusoids the wave will become almost a perfect square wave.

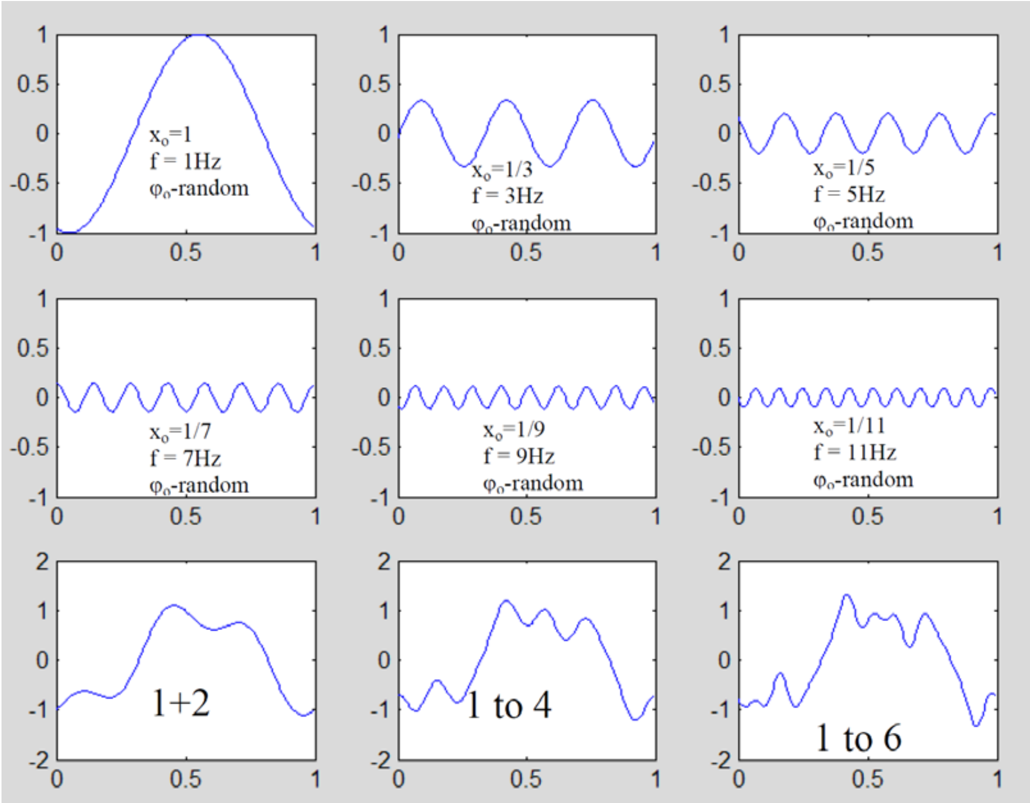

Figure 5: ϕ_o of each co-sinusoid varies and is allocated from the uniform random distribution between –π and +π radians (i.e. -180° and 180°), i.e. signals start randomly. The combination produces a random signal.

Application:

It is often beneficial to convert complex random signals (such as that developed in *Figure 5*) from the time domain to the frequency domain. Signals which can look extremely complicated in the time domain, can reveal very meaningful information when presented in the time domain. For this reason, Fourier analysis is regularly applied.

Example: An example of an automotive application of frequency spectrum is demonstrated in *Figure 6.* The roof of the car is known to vibrate excessively at certain driving conditions. The vibrations signal measured with an accelerometer is complex and contaminated, and it is not clear what the cause of this excessive vibration is. However the magnitude spectrum of acceleration reveals a peak at 30 Hz which, after some analysis, coincides with the RPM of one of the shafts in the gearbox which rotates at 1,800rpm at these driving conditions.

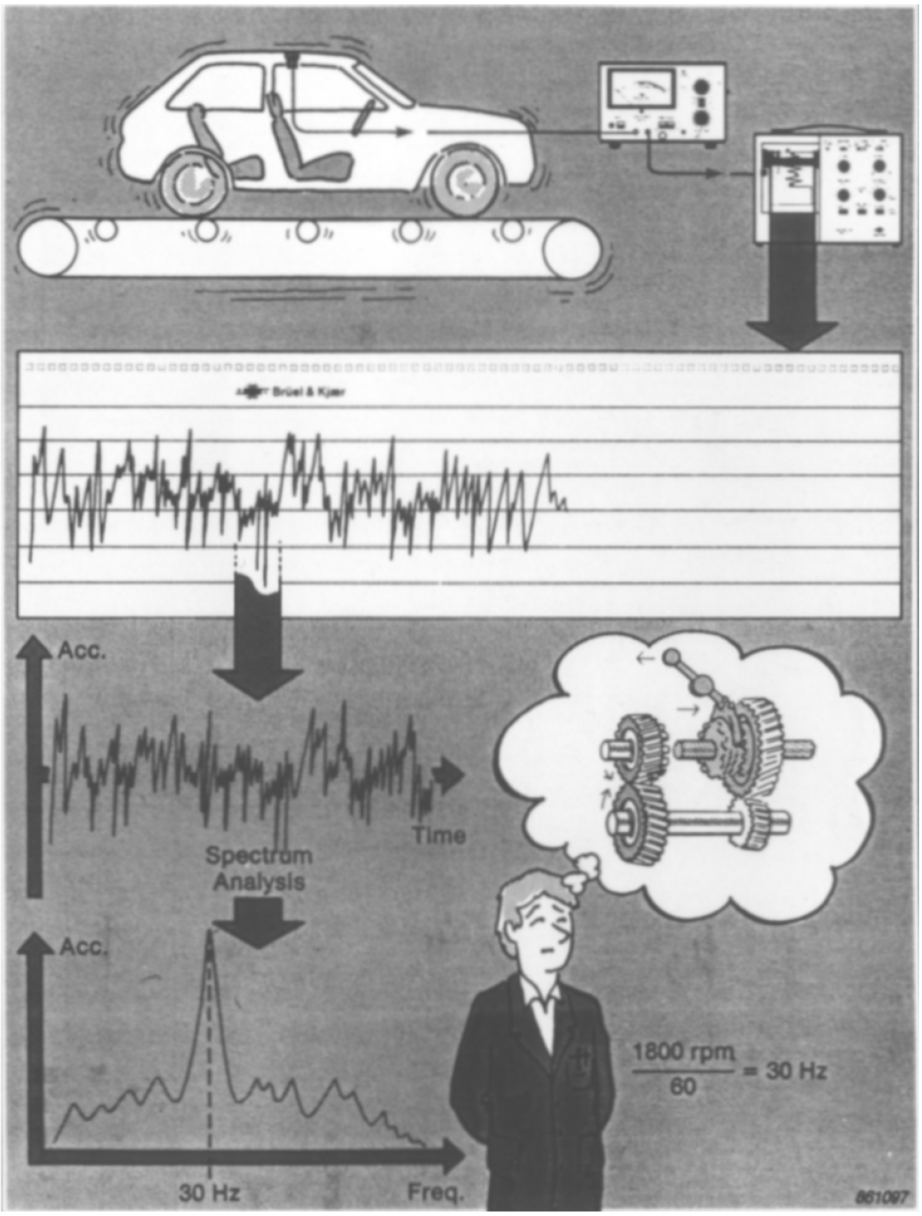

Figure 6: An automotive application of frequency analysis (Brüel&Kjær: Structural Testing)

Discrete Fourier Transform:

In engineering practice signals are usually not described by mathematical functions but come from measurement acquired with a selected sampling interval Δt (its inverse is the sampling frequency or rate fs). Thus the signal is not continuous but discrete. They may also be obtained from a simulation models (eg. in Matlab Simulink) which also produce discrete signals. The duration of a signal is finite and in most cases it will not be the same as the period required by the Fourier theorem. The signal may not be periodic at all.

The discrete Fourier transform for a signal (g) sampled at discrete times (t_n) and over the finite duration ($k\,\Delta t$) takes the form of:

$$G(f) \approx \sum_{n=1}^{k} g(t_n) e^{-j2\pi f t_n}$$

, where f is frequency in Hz.

When f= 0 the sum degenerates to the sum of the values in the signal which is real. In order to produce the mean value, the result needs to be divided by the number of points, k, in the signal. The same scaling needs to be applied to all frequency components of the spectrum in order to produce the correct estimate of the magnitude. At f= 0 the so called DC-value (or the mean value) of the signal is obtained. In many cases in engineering practice the DC value is, or should be 0 even if a measurement does not indicate so, e.g. the mean acceleration of a vibrating and stationary object.

Shannon's sampling theorem:

Measured signals generally contain various frequency components to be detected and often these frequencies are not known prior to performing the frequency analysis. What are the pitfalls of selecting a "wrong" sampling frequency? If the frequency to be detected is f and the sampling frequency is too low, i.e. $fs \leq 2f$, a lower erroneous frequency will be perceived, as illustrated in *Figure 7*. The phenomenon is called the aliasing.

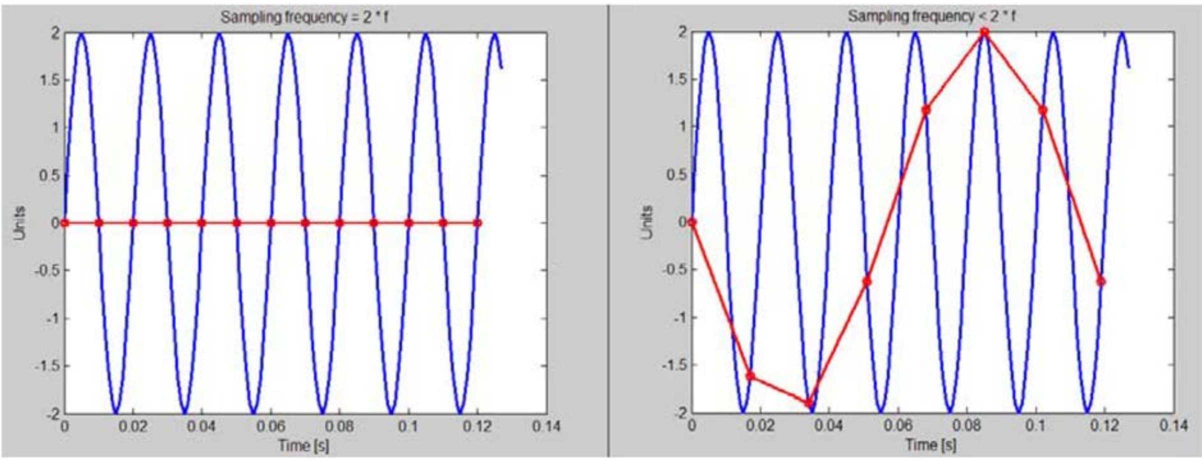

Figure 7: Aliasing (Dr Michael Sek 2016). LHS $f_s = 2f$. RHS $f_s < 2f$.

The frequency equal to half of the sampling frequency is the highest detectable frequency. It is called the Nyquist frequency. The most effective way of preventing the aliasing is by filtering it with a low pass filter with the cut-off frequency less than half of the sampling rate. Randall ("Frequency Analysis", pg. 156) recommends using a low pass filter set to allow past frequencies up to 80% of Nyquist and only display the frequencies unaffected by the filter.

Shannon's sampling theorem: The measured signal must not contain frequencies above half of the sampling rate.

Nyquist Frequency: For a signal sampled with the sampling rate *fs*, due to the limitation imposed by Shannon's theorem, components of the spectrum can only be extracted for frequencies between 0 and *fs/2* (Nyquist frequency).

Frequency resolution (***Δf***): Is the spacing between each magnitude estimate on the frequency domain spectrum (see Figure 3(b)). It is important to notice that the *Δf* depends solely on the length of the signal being analysed. This means that the longer the recording duration, the better (the finer) the resolution.

Fast Fourier Transform (FFT):

The Fast Fourier Transform (FFT) is an algorithm for calculation of the DFT first published in 1965 by J.W.Cooley and J.W.Tuckey. It has revolutionised the modern experimental mechanics, signal and system analysis, acoustics, and paved the way for the introduction of modal analysis.

The FFT return a set of complex numbers, with exception of the spectral which can be used to represent the magnitude and phase of the signal at each discrete frequency. The number of FFT elements returned is equal to the size of the time sample. <u>The second half of these complex numbers corresponds to negative frequencies and contains complex conjugates of the first half for the positive frequencies, and does not carry any new information</u>.

Implementations of FFT algorithm are commonly available and we will be using the functions available in Matlab

General Comments on FFT: In order to correctly interpret the results of FFT, one must understand how FFT returns the results of the transform. The following is common to most implementations of FFT including MS Excel® and Matlab®.

Assuming that the number of points *k*, being a power of 2, was sampled with frequency $fs = 1/\Delta t$ then

- The frequency resolution $\Delta f = fs/k$.
- The maximum frequency of the spectrum is the Nyquist frequency = $fs/2$.
- Since the FFT only does the summation of terms, the values returned by FFT must be scaled by dividing them by the number of points, *k*.

Folding FFT Spectrum into a One-Sided Spectrum:

The negative frequency components are there for mathematical reasons. For practical reasons we are interested in a presentation of the spectrum for positive frequencies only, the so called one-sided spectrum, similar to Figure 3. However we need to compensate for the effect of ignoring the negative frequencies, by multiplying all positive frequency components by 2. This process of converting from one-sided to two-sided spectrum is called folding. An algorithm for folding the FFT spectrum and calculating and plotting a
one-sided magnitude spectrum is depicted in Figure 8.

Lab Exercise 2a: Develop an FFT analyser (including folding, scaling and doubling) in Matlab and use it to analyse your signal from Exercise 1. Your M-Script should return both the magnitude and phase of the signal. Try varying the length of your signal and look at the influence it has on the available frequency resolution.

<u>Note:</u> The magnitudes that your algorithm returns may not perfectly match the amplitudes of the co-sinusoids in your original signal. These is due to an effect referred to as leakage which causes energy to spread to the adjacent frequencies. It is also a result of the discrete character of your signal and the lack of periodicity in the sample. We will look at corrections for this later.

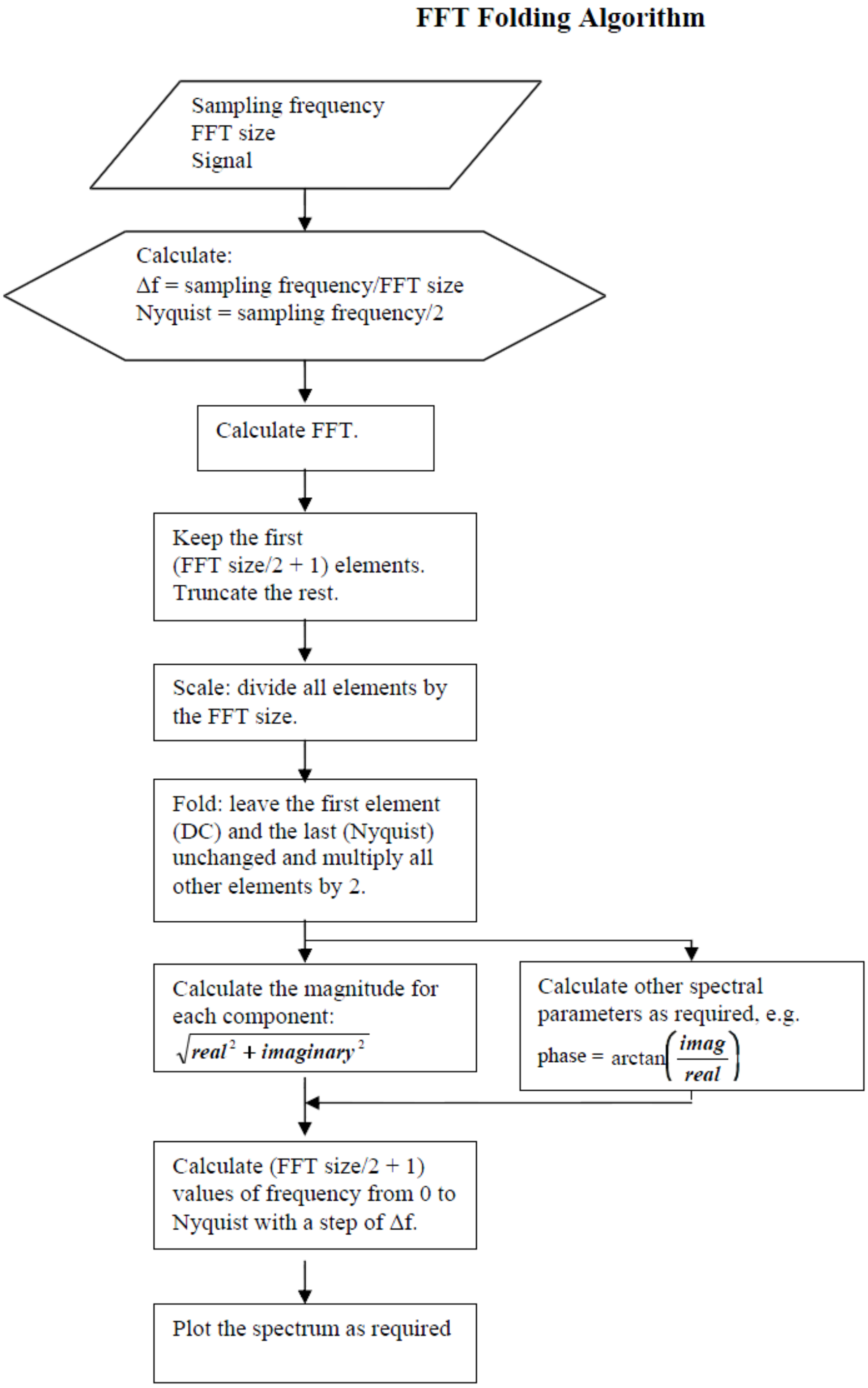

Figure 8: The FFT folding algorithm.