Linear Systems and Harmonic Response functions

David Jupp
Notes September/October 2018

1. Introduction

The examples in this discussion are 1D functions and systems. The theory uses functions but the examples must use discrete (sampled) data. This is also an introduction to differences between functions and sample data.

2. Revision

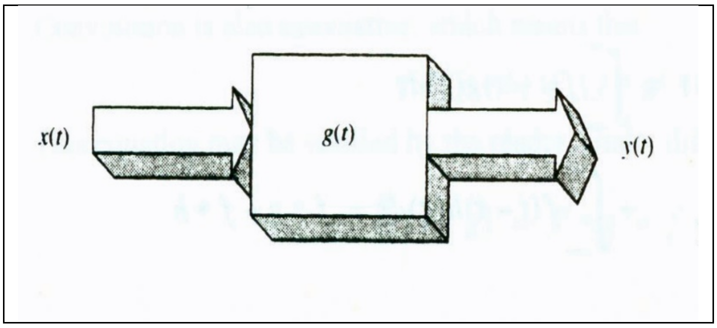
Chapter 9 of Castleman (2008) concerns linear systems. Examples of linear systems are time based electronic systems and space based optical systems. In all cases they can be expressed simply in the form:

$$y(t)=\int_{-\infty}^{\infty} g(t,t')x(t')dt'$$

If the systems are shift-invariant, then the system function is time independent and the system equation can be written:

$$y(t)=\int_{-\infty}^{\infty} g(t-t')x(t')dt'$$

In this form, the linear system is defined by a real function G that acts on the input signal x as a convolution to result is an output signal y.



Convolution is special form of cross product of the form:

$$y(t)=\int_{-\infty}^{\infty} g(t-t')x(t')dt' = g(t)*x(t)$$

It has convenient properties including:

$$\begin{aligned}
&f * g = g * f \\
&f * (g + h) = f * g + f * h \\
&f * (g * h) = (f * g) * h \\
&\frac{d}{dt}\left(f * g \right) = f' * g + f * g'
\end{aligned}$$

It is NOT the same as covariance (sometimes called cross correlation) operation defined as:

$$\text{cov}(g, x) = y(t) = \int_{-\infty}^{\infty} g(t + t')x(t')dt'$$

In convolution, the convolving function is "reflected" before multiplication with the input function. In covariance, the convolving function is centred at "t" and the functions G and x are integrated. The function G then moves forward. They are related as:

$$\text{cov}(g, x) = g(-t) * x(t)$$

They have some different properties that makes each of them better to use in different situations. For example, while $f * g = g * f$, cov(x,y) is not necessarily equal to cov(y,x). But cov(x,x) is a very important function – often called the auto-correlation function that will be used later.

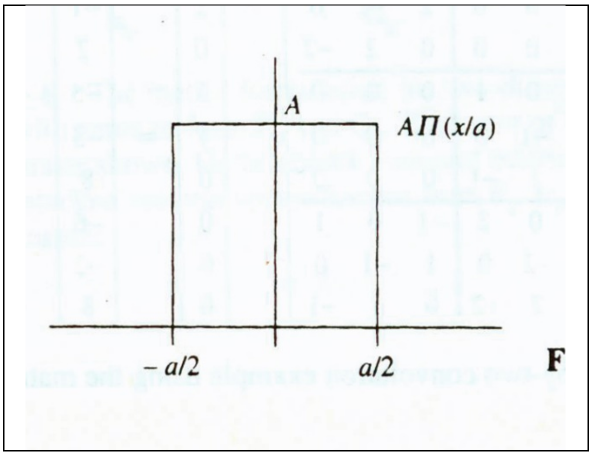
3. Impulse response

The function $g(t)$ is called the "impulse response" of the linear system. An "impulse" or delta-function (ie δ-function) is a special function that is not a part of the real number system and is defined by its operation on real functions and as the limit of well-defined functions. Its behaviour is defined as:

$$\begin{aligned}
&f(x_0) = \int_{-\infty}^{\infty} f(x)\delta(x - x_0)dx \\
&\int_{-\infty}^{\infty} \delta(x - x_0)dx = 1 \; \textit{for all } x_0
\end{aligned}$$

The delta-function can be approximated by many well-behaved functions including infinitely differentiable functions as illustrated by a simple step function (which is discussed in more detail later) called the Rectangular Pulse:

$$\Pi(x) = \begin{cases} 1 & -0.5 < x < 0.5 \\ 0.5 & x = \pm 0.5 \\ 0 & \textit{elsewhere} \end{cases}$$



The function $\frac{1}{a}\Pi(x/a)$ has integral 1.0 and by the mean value theorem:

$$\int_{-\infty}^{\infty}\frac{1}{a}\Pi((x-x_0)/a)f(x)dx = \frac{1}{a}\int_{x_0-a/2}^{x_0+a/2} f(x)dx$$
$$= f(\xi) \;\; for\ some\ \xi \in (x_0+a/2, x_0+a/2)$$

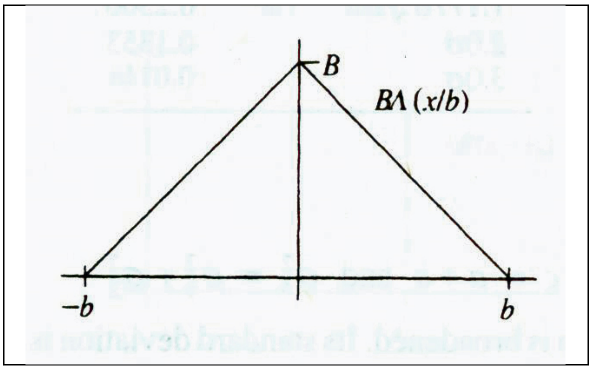
As the interval gets small the function behaves more and more like a delta-function. Applying the delta-function as input ($x(t)$) to the linear system the response is $y(t) = g(t)$. So if a linear system is defined as the application of a Rectangular Filter to a data set, its impulse response will be the Rectangular Pulse. This operation smooths the delta-function into a rectangular pulse.

4. Other Test Functions

Castleman section 9.4 provides the definitions of three useful functions. These will be used to illustrate ideas in many situations. One is the Rectangular Pulse (Π) above and the others are the Triangular Pulse and the Gaussian Pulse or Gaussian Function.

The basic Triangular Pulse (Λ) is:

$$\Lambda(x) = \begin{cases} 1-|x| & |x| \le 1 \\ 0 & |x| > 1 \end{cases}$$



The function $\frac{1}{b}\Lambda(x/b)$ has integral 1.0. Also, since $\Lambda(-x)=\Lambda(x)\geq 0$ it follows the Triangular Pulse is non-negative and symmetric. Such functions are ideal low pass filters.

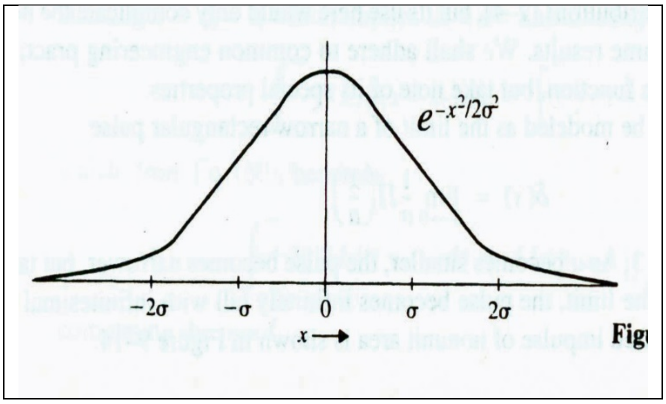
The Gaussian Function is defined as:

$$g(\sigma;x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-x^2/2\sigma^2}$$

The integral is 1.0 so we will call it the Gaussian Pulse. The Gaussian function has a very convenient convolution property as described by Castleman, Eq (55):

$$\begin{aligned} &Ae^{-(x-a)^2/2\sigma_1^2} * Be^{-(x-b)^2/2\sigma_2^2} = ABe^{-(x-c)^2/2\sigma_3^2} \\ &c = a+b \\ &\sigma_3^2 = \sigma_1^2+\sigma_2^2 \end{aligned}$$

That is, the convolution of two Gaussians is again a Gaussian. This will be very useful. It has some very interesting consequences such as the sum or difference of two normal distributions having a normal distribution.



Consider the convolution of a Rectangular Pulse with itself. That is:

$$\Pi * \Pi(t) = \int_{-\infty}^{\infty}\Pi(t-t')\Pi(t')dt' = \int_{-\infty}^{\infty}\Pi(t'-t)\Pi(t')dt'$$

Suppose that $t<-1$ or $t>1$ then the two pulses will not overlap so the result is zero. The result will also be symmetric so we will just then consider the case where $0\leq t\leq 1$:

$$\Pi * \Pi(t) = \int_{t-1/2}^{1/2} dt' = 1-t$$

Because of symmetry, we can write:

$$\Pi * \Pi(t) = \begin{cases} 0 & |t|\geq 1 \\ 1-|t| & |t|<1 \end{cases}$$

But this is $\Lambda(t)$ so the convolution of the Rectangular Pulse with itself is simply a Triangular Pulse.

[Class exercise: what is the convolution of a Triangular Pulse with a Triangular Pulse?]

From Castleman we also find that the convolution of a Gaussian with itself is another Gaussian. Specifically, suppose the Gaussian Pulse has standard deviation of σ, then from the formula provided by Castleman:

$$g(\sigma;t) * g(\sigma;t) = \frac{1}{2\pi\sigma^2} e^{-t^2/4\sigma^2}$$

The functions described here ca be used to smooth data as filters. For example, the Rectangular filter is a simple running mean filter. We will see how running mean averages are not "good" filters. A measure often used with smoothing filters is to compute the equivalent band width or "FWHM" defined as:

$$fwhm = \frac{\int_{-\infty}^{\infty} F(t)dt}{\max\left(F(t)\right)}$$

For the rectangular filter fwhm is a, for the Triangular Filter it is b and for the Gaussian it is $\sqrt{2\pi}\sigma$ so they all measure how much "smoothing" they carry out. These and other properties as well as 2D generalisations will be used in our later work.

The test functions can all be used to define linear systems that smooth data by convolution. In each case, the impulse response is simply the Pulse. Suppose the linear system smooths the input series by applying a rectangular pulse filter twice with suitable delay. The impulse response of the composite filter is a Triangular Pulse. The response of a linear system to various inputs is the way its behaviour is described and analysed. Castleman in Section 9.4.5 discusses how the response to a step function can be used to analyse a linear system. A step function is the integral of a delta-function. But our objective is to see how harmonic functions can display the properties of filters using Linear System theory.

5. Harmonic Response

Derivation

Castleman Ch 9.2 covers the way to characterise a linear system by its behaviour on harmonic functions. The primary harmonic signal is:

$$x(t) = e^{j\omega t} = \cos(\omega t) + j\sin(\omega t)$$
$$\omega = 2\pi f$$

Where $j^2 = -1$ is the primary complex number[1]. Harmonic inputs are usually complex and so a brief summary of complex numbers has been provided in an Appendix. The system response can be written:

$$y(t) = K(\omega)x(t)$$

Where K is known as the transfer function. Applying a harmonic input to the system and writing K in the polar form:

$$K(\omega) = A(\omega)e^{j\phi(\omega)}$$
$$K(\omega)e^{j\omega t} = A(\omega)e^{j(\omega t+\phi)}$$

The output from this system is real and so it follows:

$$y(t) = A(\omega)\cos(\omega t + \phi)$$

This leads Castleman to 3 conclusions:

1. A harmonic input to a linear system produces a harmonic output with the same frequency as the input;
2. The system is completely specified by its complex transfer function;
3. The transfer function only modulates the amplitude and/or creates a phase shift.

Harmonic responses of useful functions

In 9.4 of Castleman we found some useful functions that can be used to form smoothing filters. We will look at the linear systems these define and compute their harmonic response. There will be some differences between the properties of the continuous functions and the sampled data used in the calculations. They will be instructive as they are what will arise when we use discrete filters and discrete Fourier transforms.

We can first look at the function harmonic responses using integration. For reasons we will leave aside the system response can be defined only in terms of $\cos(\omega t)$. That is:

$$y(t) = \int_{-\infty}^{\infty} G(t-t')\cos(\omega t')dt'$$

First, suppose G is Π:

[1] If people have forgotten the basics of complex numbers a brief summary is included with name "Notes on complex numbers.pdf". DLBJ.

$$
\begin{aligned}
y(\omega;t) &= \frac{1}{a}\int_{t-a/2}^{t+a/2}\cos(\omega t')dt' \\
&= \frac{1}{\omega a}\left(\sin(\omega(t+a/2)-\sin(\omega(t-a/2))\right) \\
&= \frac{\sin(\omega a/2)}{\omega a/2}\cos(\omega t)
\end{aligned}
$$

So $A(\omega) = \dfrac{\sin(\omega a/2)}{\omega a/2}$

For the Triangular Pulse (Λ) it is a little more complex with:

$$
\begin{aligned}
y(\omega;t) &= \frac{1}{b}\int_{-b}^{b}\left(1-|t''|/b\right)\cos(\omega(t''+t))dt'' \\
&= \frac{\sin^2\left(\omega b/2\right)}{\left(\omega b/2\right)^2}\cos(\omega t)
\end{aligned}
$$

Observing that b and a are the same when the Triangular Pulse is obtained as the convolution of the Rectangular Pulse with itself, it appears that in this case:

$$
A(\omega) = \left(\frac{\sin(\omega a/2)}{\omega a/2}\right)^2
$$

That the modulation by the Triangular function is the square of the modulation of two Rectangular Pulses is no accident. What we are really looking at here are Fourier Transforms.

The function $\sin(t)/t$ is called the "Sinc" function and it will return often as we look into the Fourier transform and related filtering operations.

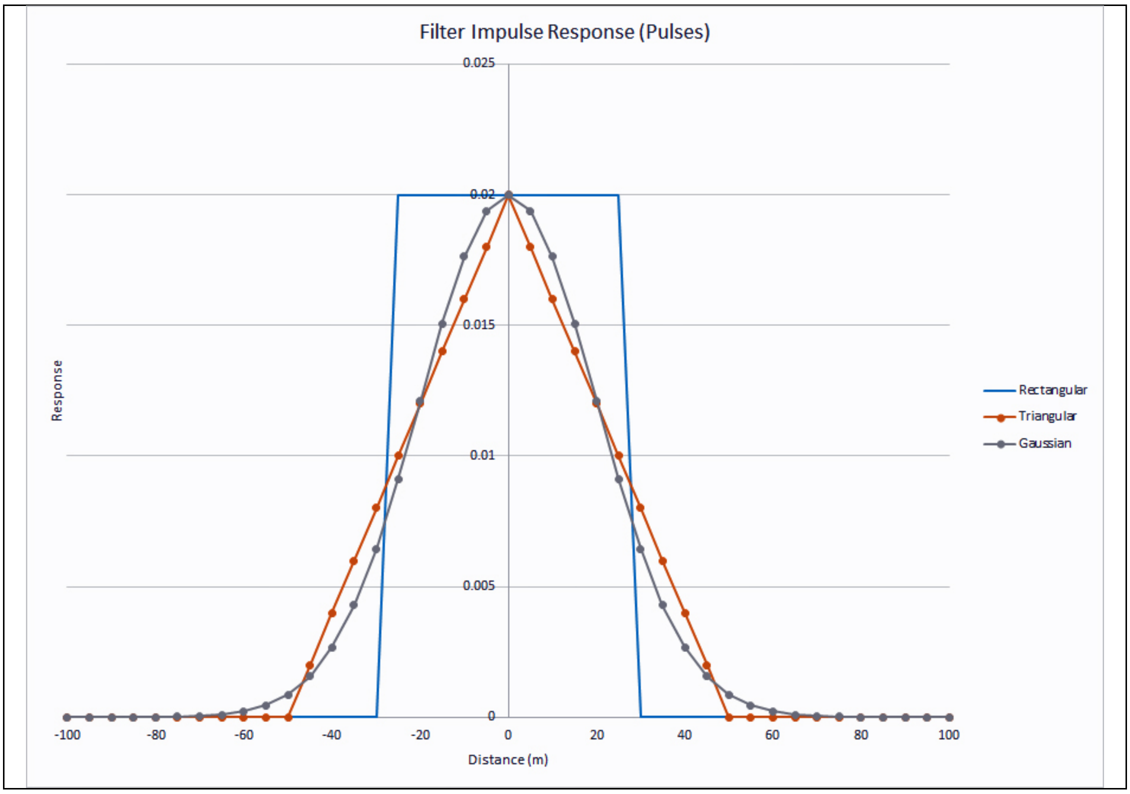
In each case, there is no phase shift. This comes about as the linear system functions are real, symmetric and non-negative. They are classic smoothing filters.

We will not derive the case of the Gaussian except to note that in this case:

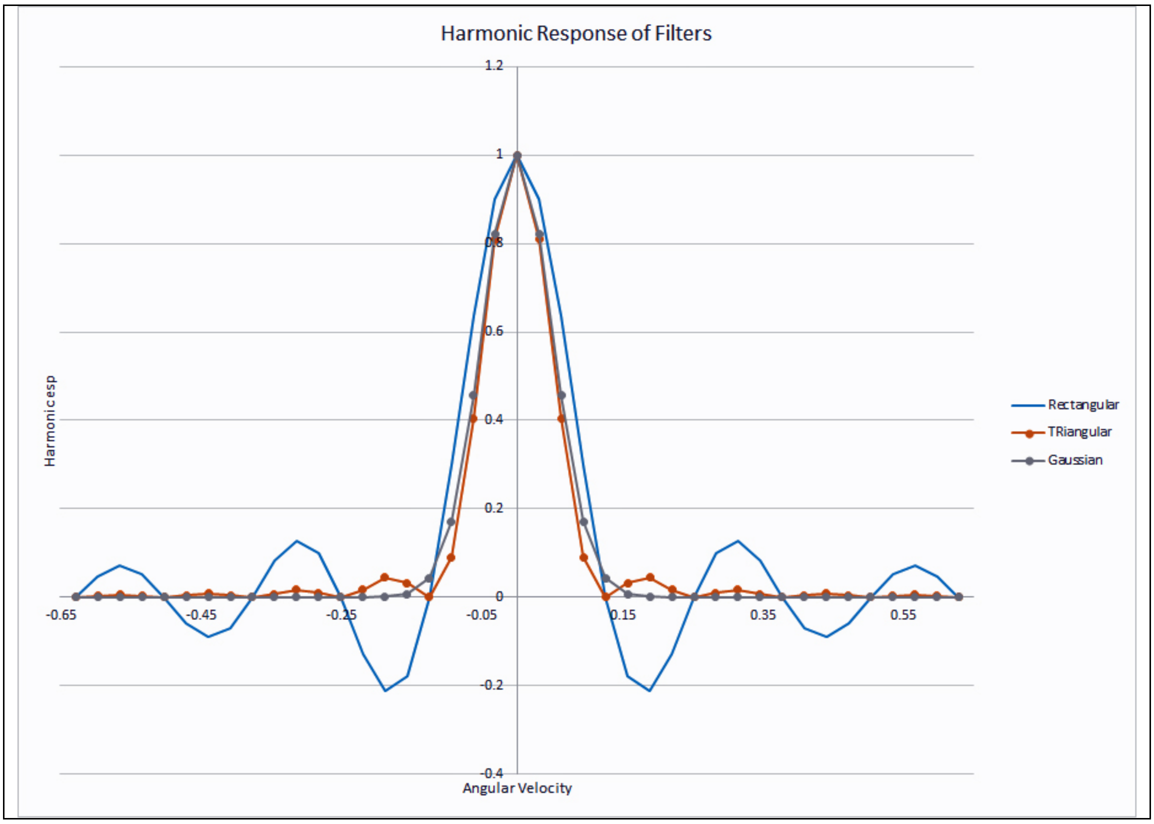
$$
A(\omega) = e^{-\sigma^2\omega^2/2}
$$

Proving this is much more complex than the other examples but will be visited again when Fourier transforms are reached.

Above we noted that a measure of the "width" of the filter was a for the Rectangular filter, b for the Triangular filter and $\sqrt{2\pi}\sigma$ for the Gaussian. So let the fwhm be 50m. A plot of the three functions (ie their impulse response functions) is as follows:



The Rectangular is the effective bandwidth for the other two functions in this case. Normally it is not as “neat”. They all have the same fwhm of 50m. The corresponding theoretical harmonic response functions for these pulses are as follows:



The horizontal axis is the angular velocity (ω) plotted with a range that will be discussed later. The behaviours of the modulation by the filter as a function of angular velocity (or frequency) has the form of the functions derived in class. Note the behaviour of the simple average filter which does not modulate the high frequencies anything as much as the others. Overall, the useful functions are smoothing functions that reduce the amplitudes of the high frequency components.

Harmonic Response in a sample data system

Data periodicity

To experiment with the harmonic response we will use a long series of $\cos(\omega t)$ sampled at spacings of 5 units (eg metres). The various operations will be applied to the series. It does not matter where the origin is so we will just say it is a sequence:

$$x(kh) = \cos(k\omega h) \qquad k = -N, N$$

The actual series continues outside of this range so that there are no edge effects with filters but this is the range to be analysed.

There are two parameters that are critical to consider when such a sample harmonic data series is being analysed. One is the data length and the other is the periodicity of the harmonic series.

The data length is $2hN$. It follows that the lowest frequency that can be represented is $1/2hN$ corresponding to angular velocity $\omega = \pi / hN$. This is the frequency at which one cycle just occurs in the data length.

The other represents the fact that the series is periodic with base range of:

$$-\frac{\pi}{h} \le \omega \le \frac{\pi}{h}$$

Suppose ω_p is in the range, it follows that for all integers m (positive or negative):

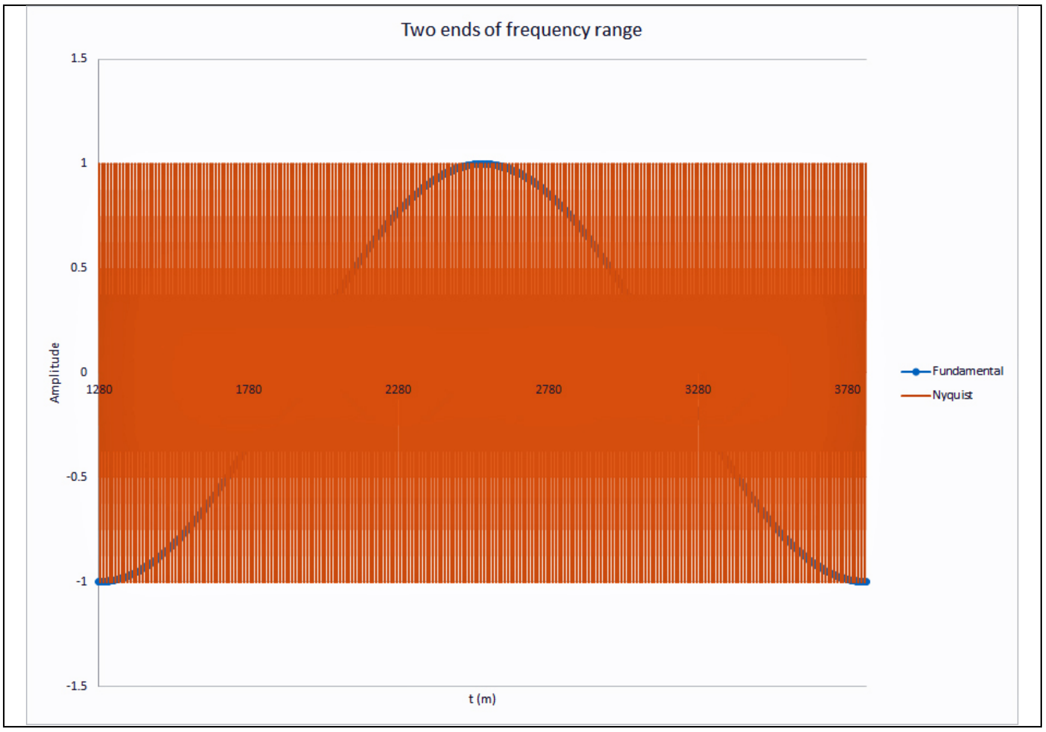
$$\begin{aligned} \cos(kh\left(\omega_p + m\frac{2\pi}{h}\right)) &= \cos(kh\omega_p + km2\pi) \\ &= \cos(kh\omega_p) \end{aligned}$$

This happens as above the "Nyquist" frequency, the higher frequencies are not representable using a series sampled at step h. It occurs because of the sampling and we will see how it affects filtering and Fourier processing later.

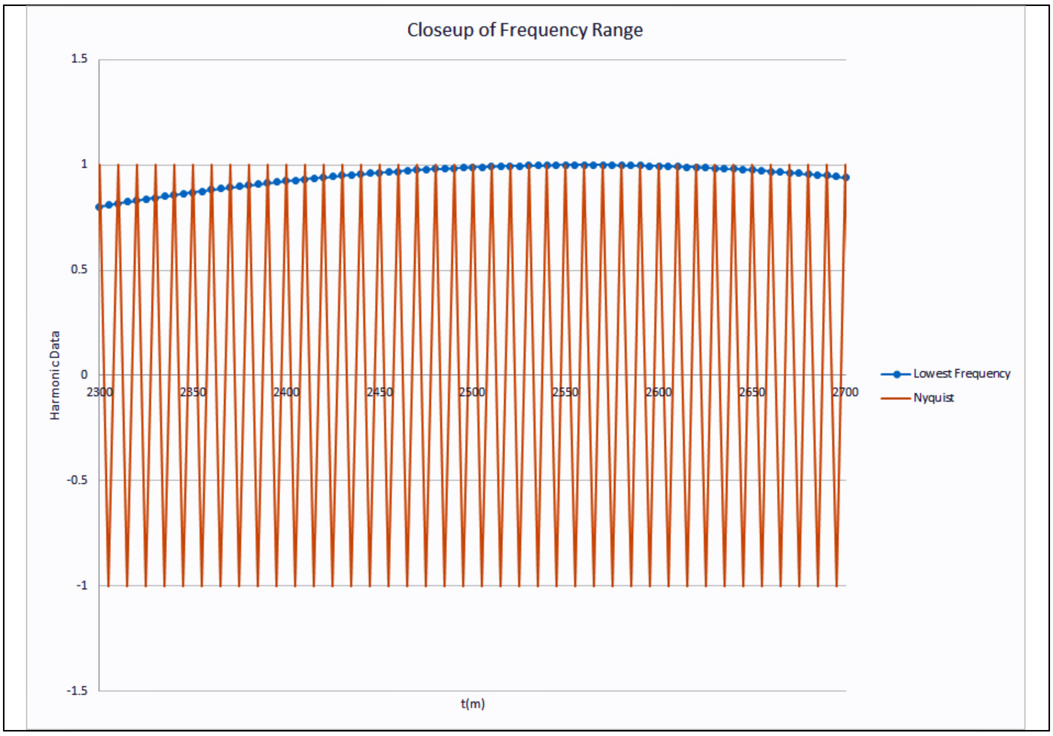
For now it means we will only consider the range $\left[\frac{\pi}{hN}, \frac{\pi}{h}\right]$ in our work. It can be seen that if hN extends “to infinity” and h to zero then all frequencies will be obtained and there will not be any periodicity. This represents the big difference between the continuous and discrete situations.

Results

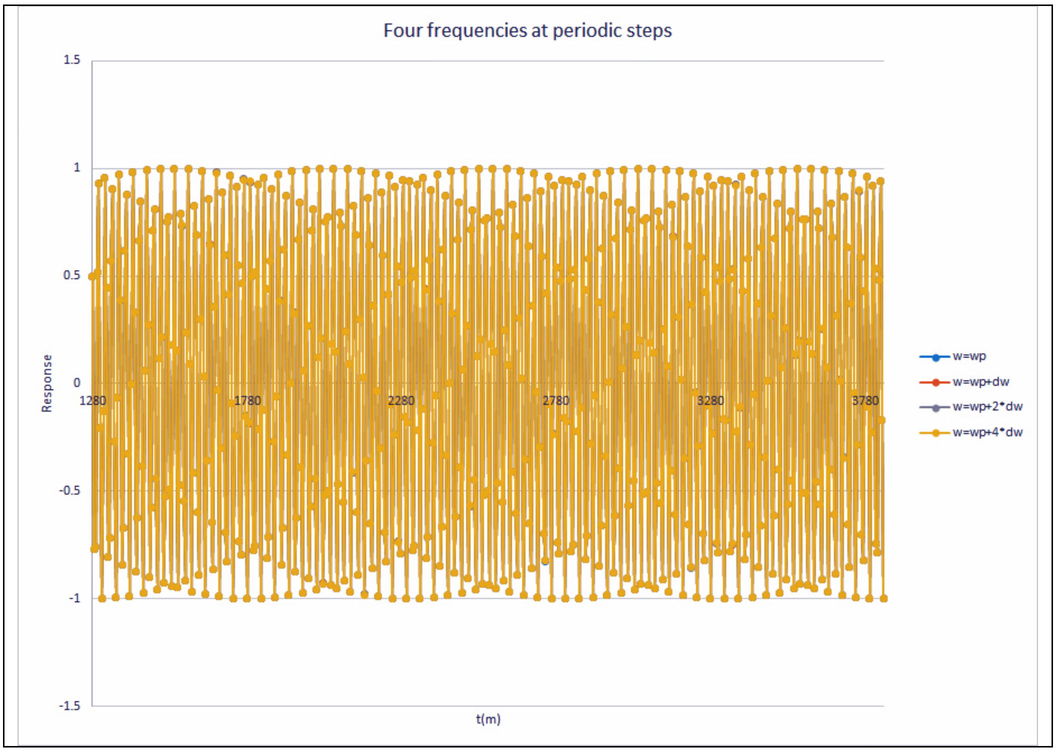
We first look at the harmonic series itself. The first plot shows the two cases where the frequencies (ω) are at the smallest representable frequency and the largest:



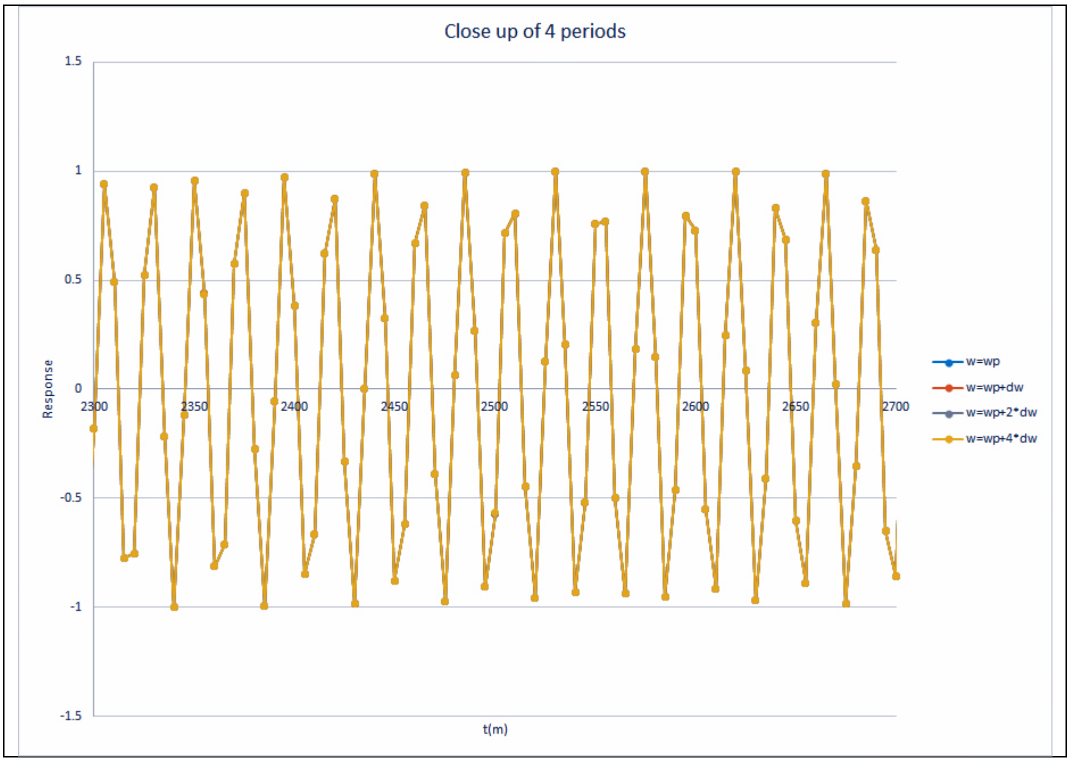
The highest frequency covers the page at the full data range and the background plot of the lowest frequency shows how it is one cycle within the data range. When you look closely at the plot, the highest frequency resolves to a function that has one cycle over 2 steps. Beyond this the sampling cannot represent the data.



The periodicity is demonstrated by taking frequencies of $\omega = 0.280616, 1.53725, 2.79389, 5.30716$ where $\omega = 0.280616$ is within the Nyquist range and $\omega = 1.53725, 2.79389, 5.30716$ are shifted to higher frequencies by 1, 2 and 4 times the periodicity of $\delta\omega = 1.256636$. The Nyquist range is $-0.628318 \le \omega \le 0.628318$. The following plot is of the four cases which are indistinguishable at the sample points:

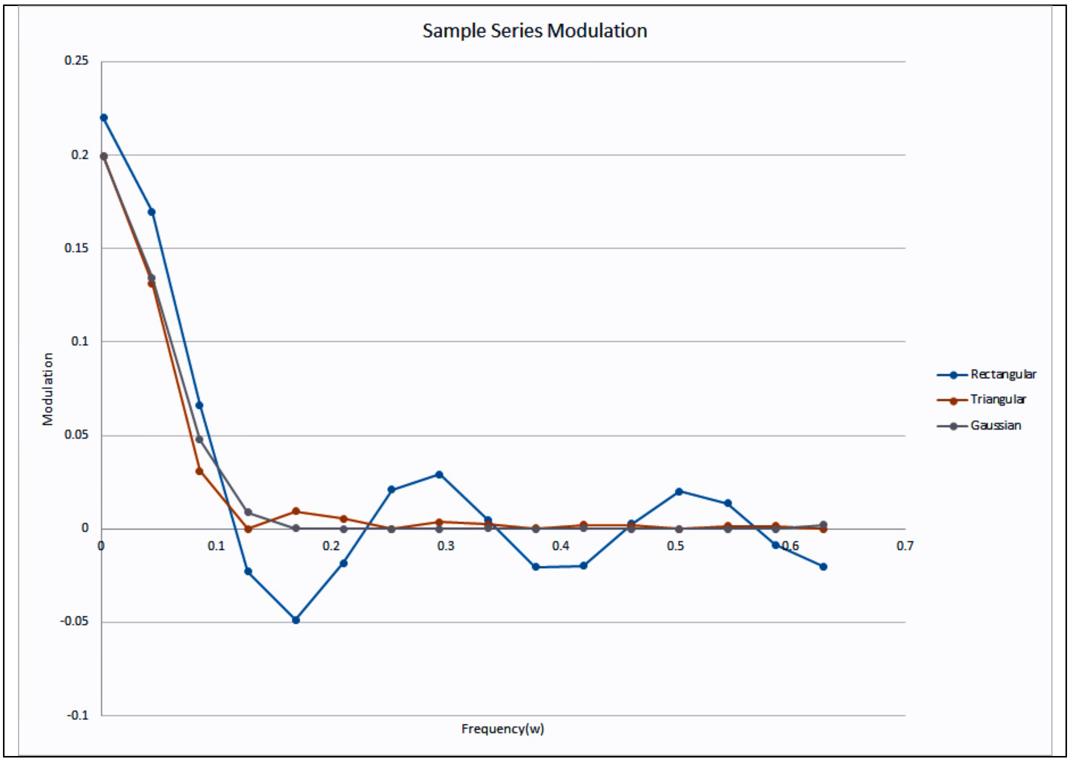


This set of four plots are not distinguishable. They show some clear aliasing due to the representations on Photoshop, Excel and Word! But the same plot close up as was done above for the basic harmonic data is much more as expected:



These effects are significant and so they were explained more carefully here so that later you will have some basis with which to understand aliasing and other discrete data effects.

Our main goal is to study the harmonic responses. To do this, the three filters were applied to the harmonic series and the model $A(\omega)\cos(\omega t+\phi)$ fitted to the data. In each case the phase shift was zero and the amplitudes match the theoretical responses within the differences that occur with sample data:



This is the end of the examples that arise from considering harmonic response. The Fourier Transform will take over as the main tool that can be used to analyse and design using these methods. The simpler frequency response example was introduced to see how the Fourier theory came about and give a simple introduction to the ideas that will recur in the course.

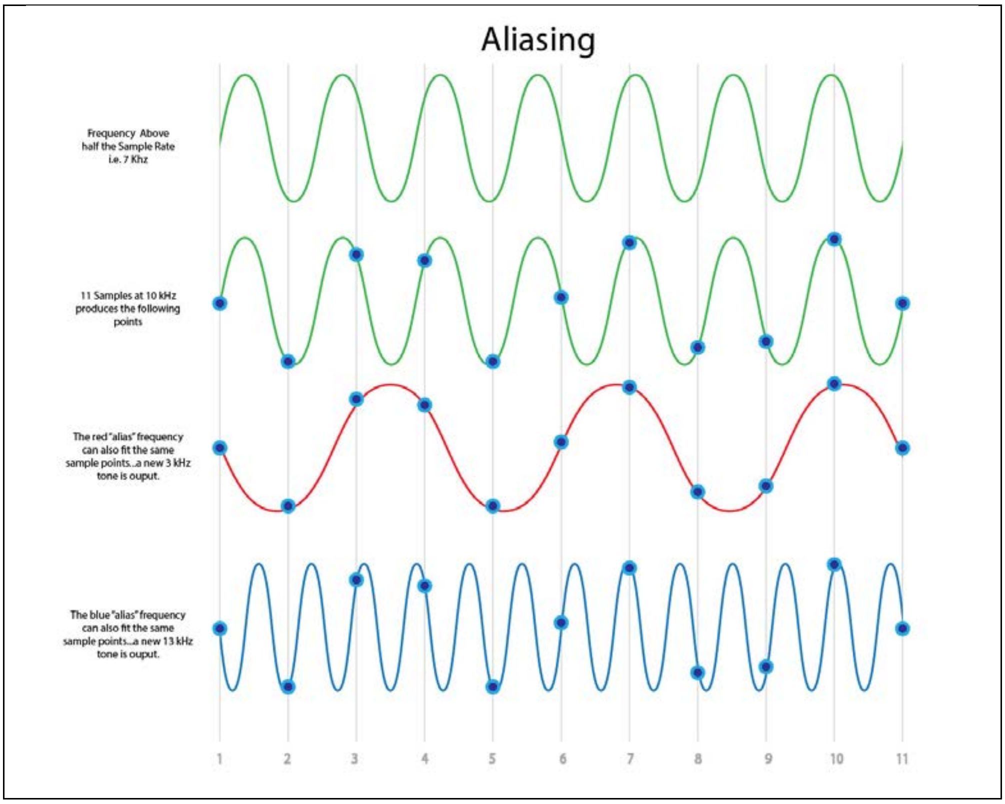
6. What we have covered?

In this discussion, we have looked at the "harmonic response" of a linear system. The linear system is one of the simple functions acting as a filter on the input data. The outputs from the linear system can be computed exactly or simulated by a sample data system input. The filters are the "impulse response" functions and we were looking at the difference between the impulse responses and the harmonic responses. That is, a linear system can often be characterised by its response to various controlled input functions.

In the case of the theoretical results the discussion was useful as next we will see that the harmonic response is basically the Fourier Transform – or rather its real part. The use of complex variables and separation of real and imaginary parts is involved but generally the filters are symmetric and so both the harmonic response and the Fourier transform are real.

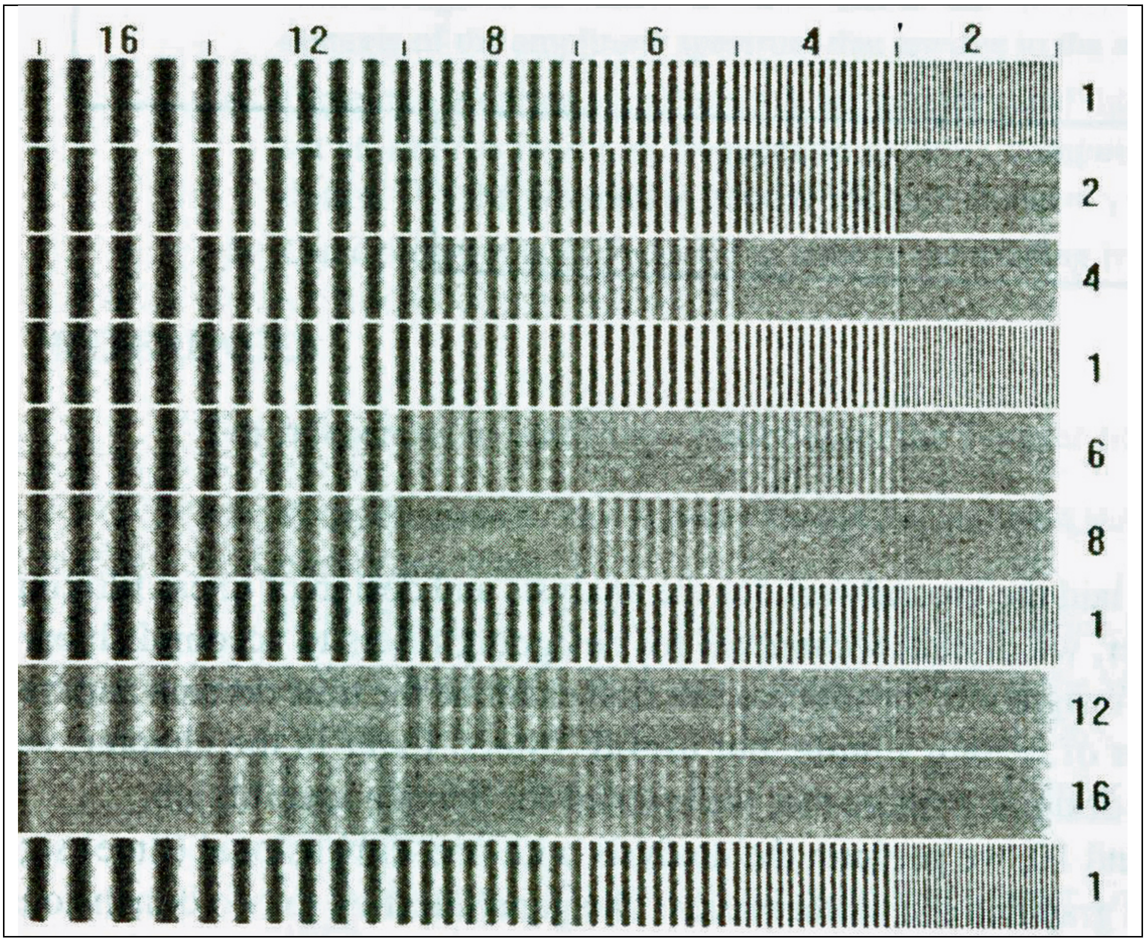
In addition, we looked at the results of simulating the harmonic responses with a sample data input stream of harmonic functions. This introduced some ideas basic to sample data systems. One was that the harmonic input is periodic with a period depending on the sampling step. It makes some differences to the harmonic response but these were not studied. The periodicity of the sampled harmonic function will be important later in the discussion of the discrete Fourier Transform (DFT).

Just as a diagram to keep in mind, we found that at frequencies higher than a base frequency by multiples of the period of the sample series the harmonic inputs were ***identical***. What has happened to the high frequencies? Later we will see that tis involves a phenomenon called "aliasing". For now the reason can be found in the following diagram but the outcome will be left until later.



7. Afterword - Frequencies and Lines

Another type of test function for linear systems – of special importance in optics – is the response to ruled lines. Parallel ruled lines of varying thickness and spacing are common test functions where line spacing is related strongly to frequency. A bar chart like the following is often used to describe the blur function (often called the MTF) of a lens system.



In this diagram from Castleman, the line thickness increases as spacing increases so that roughly the proportions of black and white are similar in the blocks of different line density. The numbers on the top indicate lines per unit length (inch or cm) in the block. The chart shows the effect of a simple average (rectangular filter) of width shown on the right hand column. A width of 1 is no averaging so all lines with value 1 are original unfiltered charts. Having one at the top and bottom and between “treatments” allows an easier comparison between unfiltered and filtered.

As the averaging filter gets longer what would you expect? If the averaging is over lines the average should reduce to a gray value. If the balance of black and while is about the same in each block the gray value will be similar. So as the width of the filter increases the gray bars should extend from right to left. This is shown in the lengths of 2 and 4 (lines 2 and 3 in the table). Line 2 has the finest line set gray and others blurring. Line 3 has two gray blocks and the blur extending. Line 4 is a comparison repeat of the unfiltered chart.

Lines 5 and 6 shown a strange behaviour. The leading gray bars extend as expected but behind them some blocks have lines appearing. These lines are not the same as the original lines but have ***opposite phase***! That is where there was a black line there is now a whiter line and vice versa. In lines 8 and 9 new areas are showing the phase reversed lines.

What is happening? The short answer is in the harmonic responses above. The rectangular filter has some negative lobes that are creating these effects. Such effects are not desirable but they are characteristic of the filter. There is more to a filter than width of averaging. These and other effects are best analysed and avoided in the Fourier domain and it is to the Fourier domain that we must now go.

DLBJ
October 2018