Over Clocked Digital Filters

© Malcolm Moore
02-Mar-2003

Impulse Response Filters

Having discovered that the simple moving average is in effect the output of a non-recursive digital filter, otherwise known as a finite impulse response (FIR) filter, and that the exponential moving average is in effect the output of a recursive digital filter, otherwise known as an infinite impulse response (IIR) filter; this opened the door to the whole world of digital filters, and the very powerful maths that has been developed over the last few decades.

I am excited! The above discovery was not ground breaking, but in another earlier life, I had spent several years engineering passive filters for communications and switch mode power supplies, using capacitors and coils. At a later stage I engineered active filters using operational amplifiers, resistors and capacitors, (and in some cases I included coils / inductors)!

Unfortunately I never got very far into digital filters, as most of the engineering work for Digital Filters was already 'done' by the manufacturer, leaving very little other than purchase the finished product, and fit it into the surrounding electronics!

All these filters use much the same technique in that the Electrical Engineer initially specifies the desired outcome usually in terms of as frequency response; maybe phase response and/or a time response.

This next stage is where the Electrical Engineer earns their value. They use their experience to identify a suitable cost-effective architecture filter and that lays the foundation for the mathematic model. The ensuing maths is essentially the same in that a computer is used to iteratively home in on the component values, and/or tap setting values for digital filters.

Once the modelling is completed the real filter is then constructed and commissioning tests then prove whether the design is adequate or not. This approach is a well-structured engineering discipline. Unfortunately very few people that use moving averages for technical trading have a real knowledge of what they are expecting before they apply.

I have yet to see anybody go much further than the "suck it and see how it works" approach, and most of the argument seems to be stuck on not even knowing what the transient response of the FIR or IIR filter will be, and why. This explains why techniques used in technical trading is a 'black art' for most people, and often scorned on by those who heavily favour fundamental trading.

Over Clocking and Computers

While waiting for a train I heard two computer enthusiasts talking about their computers just like rev-heads talk about their cars. Yep, the super model had been hotted up, and was going even better then when new. In computer-ese, these enthusiasts had set their computer chips to run at a higher speed than intended by the manufacturers.

They do this by tweaking some hardware settings on the motherboard, usually not realising that the time till failure actually has a similar time/fault relationship to that of an incandescent light globe (which is very highly dependent on applied voltage). The manufacturers know this and leave ample room in their error margin for commercial purposes, so that it should run a year or two without having a crash. Industrial computers are engineered with a larger error margin so an internally caused digital crash is not likely within a decade. When you are racing a car, you can expect to crash it. If you seriously over clock a computer it may run for a day without crashing!

This setting of the chips to run at a higher than manufacturers recommendations is called “overclocking” and it can be done with computer chips, remembering that each 4 deg C rise in temperature halves the long-term life of the chip. So the fans will have to be working overtime too!

With digital computer chips, the logic data and address words take a finite time to charge and settle, the latching (locking) registers have a finite time to lock and produce an output and the wiring takes a finite time to carry the new data and addresses.

By overclocking, the times for any or all of these physical functions are compromised and the clock can latch a transient (still settling) data or address word resulting in the computer crashing.

Because the clock is switching that much faster, the chip consumes power at a rate of least the square of the clocking frequency ratio, making permanent damage by burning out a real probability.

A Little About Sampling Theory

In 1923 Harry Nyquist[1] a Swedish electronics engineer working for Bell Labs in the USA came up with his hypothesis for a criterion that in decoding any quantised sampled sequence, the clocking (sampling) rate must be at least twice the maximum frequency of the sample being restored into an analogue waveform.

The theorem was formulated in 1928 ("Certain Topics in Telegraph Transmission Theory"), but was only formally proved by Claude Shannon in 1949 ("Communication in the Presence of Noise"[2]). The theory by Harry Nyquist put forward has profound ramifications in our daily world!

We all use CDs for music, and most of us know that the sampling rate is 44 kHz, which means (using the Nyquist criterion) that the absolute maximum audio frequency that can be reconstituted from this technology is 22 kHz.

In fact, for practical purposes the audio low pass filters in the playback circuitry of the CD player cuts off at about 18.5 kHz and has ‘notches’ at 22 and 44 kHz to obliterate the clocking pulses.

So we don’t even get up to 19 kHz and the audio is already severely attenuated! Don’t worry too much as most of us have hearing that ceases to be effective above about 15 kHz. We should all thank Harry for his insight into digital recording technologies in 1923!

[1] http://www.ieeeghn.org/wiki/index.php/Harry_Nyquist
[2] http://www.stanford.edu/class/ee104/shannonpaper.pdf

Sampling and the Stock Exchange

If we were to employ Harry Nyquists' criterion to the Stock Exchange then we gain a whole new appreciation of market technical analysis. By default, the standard sampling rate for Stock Exchanges is based on the end of (working) day (EOD) and that is a nominal 24-hour period.

In other words the standard (default) Stock Market sampling rate is based on a 24-hour cycle on working days, where the sample is taken at nominally 4 pm! This sample at the market close time is called the 'Close' price, and it is a standard reference used by most technical indicators in performing calculations.

With a little more 'Nyquist criterion' thinking, it becomes obvious that in the case of a Stock Market Close being sampled once every 24 hours, means that the maximum frequency (cyclic rate of change) in a security price cannot be detected if it cycles in two working days or less. And I can assure you that some indeed do!

In fact it is usually hard to see smooth changes in Close prices within a week, based on EOD figures! So the Nyquist criterion is as profound here as it is for music from CDs, telephony through digitally switched exchanges, modems in all their varieties, Televisions, and video monitors too!

Now it is time to think laterally and realise that the stock Market is actually open for 6 hours per day – nominally from 10 am through till 4 pm, so the 'timeslot' is really 6 hours, and the Close price is taken at about 4 pm (when the market closes for that stock).

During this timeslot, the price may vary considerably and because of this, the Close price may be very misleading, as it may be the result of a rush in buying or selling, and not near the average price during the day.

This should explain why there is a rush of trades as the market is opened, and usually another lesser rush in the last two hours of trading.

Over Clocking and Digital Filters

In another field of electronics – digital filters in fact – the term 'overclocking' has an entirely different meaning. When a digital filter produces an output, this output comes as a digital word based on the digital clock rate, and the clock rate is at least twice that of the maximum output frequency.

What is not realised by most people is that the digital words that come out of a digital filter are not a continuum of graduated digital words, like an analogue signal, but like a staircase or steps of different digital words, like a rapid-fire machine gun.

Graphically these steps are hidden from view as most graph packages draw a line that 'joins the dots' to form a piecewise linear curve that is pleasing to the eye, and soon all is forgotten, but that is why some results look strange – but are correct!

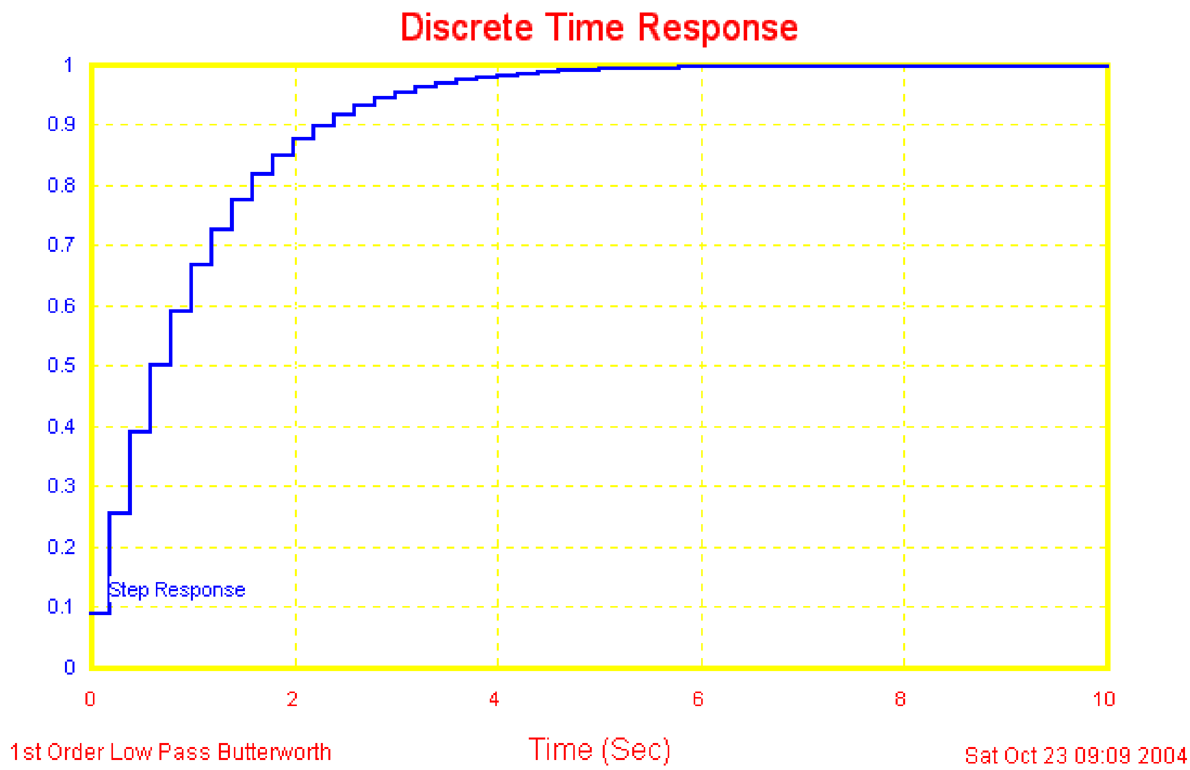
To illustrate the point, the two graphs below are the same, but the right hand one is shown in piecewise linear steps while the left hand one merely 'joins the dots' – and it ain't a pretty sight!

If the digital output came out in far smaller increments, and far more rapidly, then the overall output would have a far smoother transition and appear as a continuum, or flow, and not like a hail storm or a machine gun!

One way to minimise the size of the digital steps is to clock the analogue waveform at a much faster rate (frequency) and through that the whole process is speeded up, resulting in a much faster output clocking rate and much smaller differentials in values between the digital words of this output. This is not over clocking but merely speeding up the process.

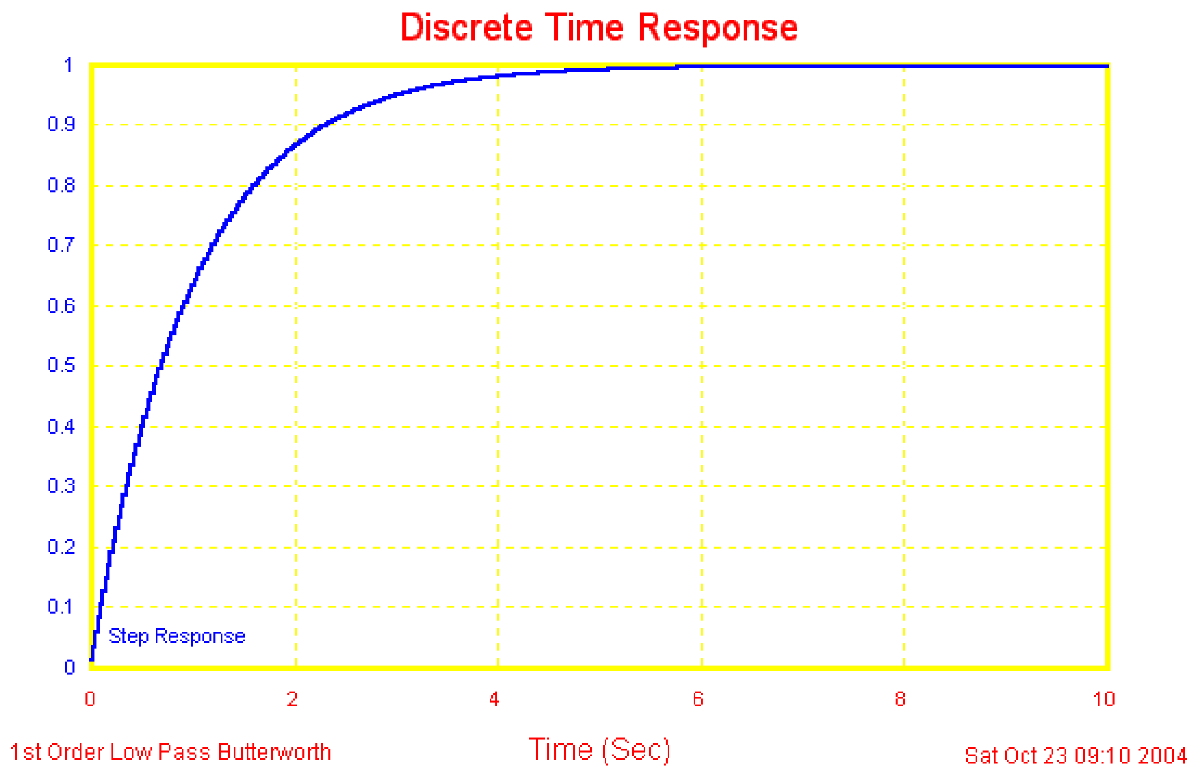
Another way is to accept that the incoming digital word rate is not under our control and that if the digital filter is clocked at a direct multiple of the input clocking rate, then the output data rate will be at this much faster rate.

We also have to assume that the changes between the original digital steps are piecewise linear. (With share prices this is a very big ask, as the prices in the most, meander like a lazy river during trading hours!) We also have to proportionally slow down the digital filter's constants so that the filter behaves as the original filter. An example of this is shown below:



The above time response is a 1st order digital response of a 1 Hz Low Pass Filter (Exponential Moving Average) being clocked at 5 Hz. Note the big steps that are coincidental with each sample.

Now, if we clocked the EMA (1st Order IIR Filter) at 50 Hz instead of 5 Hz, then the over-clocking ratio is 50 / 5 = 10 times. Now, if we multiplied the time constant of the IIR filter by 10, then the filter will have the same EMA general transient response, but there will be 10 steps coming out of this filter for every one step coming out of the previous filter. An example of the over-clocked output is shown below:



Over-clocking Applied to CD Players

CD players read the music digital words at a 44.1 kHz clock rate, and have digital signal processor included to convert the digital words to analogue sound, usually limited to about 19 kHz maximum.

In the normal case, the filter stop band has to have maximum attenuation by 22.05 kHz that is 1.1605 times the 19 kHz pass band cut-off limit, and that is not simple!

If the filter over clocking was at 88.2 kHz, then the digital step size would be halved (improving the inherent signal to noise ratio by about 6 dB and the filter would be much easier to engineer as the clock rate would be 2.32 times 19 kHz the pass band cut-off limit, and that simplifies the filter design!

The over clocking could be set to 8 times the signal reading rate (i.e. set to 352.8 kHz), reducing the apparent step size by 8 times, and improving the SNR by about 48 dB and the over clocking rate compared to the max audio signal would be about 17.64 times, making the filter design very simple!

But the catch is that the base clocking rate is 44.1 kHz and the filter would have to have a maximum attenuation by then and all its multiples too! Maybe this might explain why many audiophiles (extreme fanatics who critically listen to recorded music through over-expensive sound equipment) and some classical music listeners say that music from CDs ‘sounds metallic’!

In my opinion, understanding the actual filtering process is more important, because things ain’t what they seem!

Digital filters (particularly in digital signal processors) really hit their straps in telecommunications as echo cancellers and modem filters, in sound recording as echo and reverberation effects, in geophysical and medical instrumentation as ultrasonic image clarification, and for electric guitars as some effects units!

This same technique could be aptly used to 'smoothen' the graphing as used on EOD graphs, but the problem is that EOD data is not a simple shift from open price to close price, as the actual range of prices meanders over almost every trading day. Even using 30-minute timeslots is not good enough as the prices can meander in them.

This was a brief view on how everyday digital filters can be used apart from attempting to smoothen out the trading noise in price/candlestick graphs!

Higher Order Digital Filters

In analogue electronics, a higher order analogue filter usually consists of several components, in a 'ladder' formation (series, shunt, series, shunt…) from the input to the output, and the base number of reactive components (capacitors and coils / inductors) totals up to tell you the 'order' of the filter.

I found this to be a fascinating rule, and it worked with simplicity. Most filter manufacturers hated coils because they were usually labour intensive to manufacture, difficult to get the right component parts, prone to assembly error and therefore expensive, and manufacturers would often do anything to either minimise the number of coils, and this led to a whole range of alternative filter designs.

Such designs included crystal and ceramic lattice filters for radio and communications purposes, surface wave filters in televisions, switched capacitor filters in telecommunications, followed by digital filters in CD players and Hi-Fi systems and some service equipment.

The beauty about digital filters was that they could be programmed into what is called a digital signal processor (DSP), a small integrated circuit that was specifically designed to store and forward, multiply and divide with large digital words. Most DSP chips come with substantial analogue / digital conversion circuitry on board.

When delving into digital processors and digital filter technology with a view to using this technology for security price analysis, it suddenly struck that the clocking rate for digital filters well exceeds twice the maximum frequency (the Nyquist criteria) and more importantly that the number of stages in many filter algorithms can well exceed 100 stages.

Comments and Corrections are welcome