Effects of Sampling and Aliasing on the Conversion of Analog Signals to Digital Format

Ruwan Welaratna, Data Physics Corporation, San Jose, California

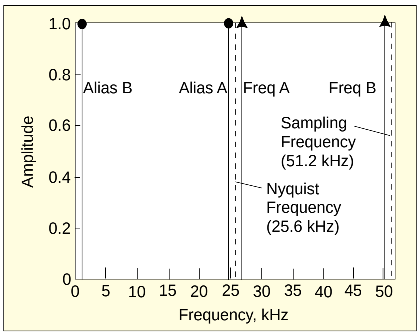

Figure 1. Frequency components above the Nyquist frequency of 25.6 kHz will create aliases.

A key step in any digital processing of real world analog signals is converting the analog signals into digital form. We sample continuous data and create a discrete signal. Unfortunately, sampling can introduce aliasing, a nonlinear process which shifts frequencies. Aliasing is an inevitable result of both sampling and sample rate conversion. It can be addressed with a properly designed anti-aliasing filter (AAF). An analog AAF must be applied before the initial sampling process. If any sample rate conversion, such as decimation, is performed, a digital AAF must also be applied. An AAF is one of the limiting factors in system performance. An improperly designed AAF, or improper application of one, can introduce distortion artifacts that may interfere with certain types of analysis. Modern AAF architectures eliminate these artifacts.

The Nyquist sampling theorem defines the minimum sampling frequency to completely represent a continuous signal with a discrete one. If the sampling frequency is at least twice the highest frequency in the continuous baseband signal, the samples can be used to exactly reconstruct the continuous signal. A sine wave can be described by at least two samples per cycle (consider drawing two dots on a picture of a single cycle, then try and draw a single cycle of a different frequency that passes through the same two dots). Sampling at slightly less than two samples per cycle, however, is indistinguishable from sampling a sine wave close to but below the original frequency. This is aliasing – the transformation of high frequency information into false low frequencies that were not present in the original signal. The Nyquist frequency, also called the folding frequency, is equal to half the sampling frequency f_s and is the demarcation between frequencies that are correctly sampled and those that will cause aliases. Aliases will be 'folded' from the Nyquist frequency back into the useful frequency range. Thus a tone 1 kHz above the Nyquist frequency will fold back to 1 kHz below, while a tone 1 kHz below the sampling frequency will appear at 1 kHz as shown in Figure 1. Frequencies above the sampling frequency are also folded back.

Aliasing is irreversible. There is no way to examine the samples and determine which content to ignore because it came from aliased high frequencies. Aliasing can only be prevented by attenuating high frequency content before the sampling process as shown in Figure 2. To prevent aliasing completely, we must apply a perfect filter that passes all energy from DC to the highest frequency of interest and rejects all energy at the Nyquist frequency and above. Unfortunately, perfect filters are not physically realizable in analog or digital form. Physically realizable filters must have variation in the passband, a smooth transition from the passband to the stopband, and finite attenuation in the stopband. Therefore, we must design a filter with unity gain and low variation in the passband and with the lowest tolerable attenuation in the stopband.

Note that finite attenuation means that you cannot eliminate aliasing, only reduce it. Suppose you sample a signal that contains a 1 V tone at 1 kHz and a 1 V tone at 39.9 kHz. You wish to analyze the data to 20 kHz, so the sampling frequency f_s is 40 kHz. If the AAF gain is –80 dB in the stopband (above 20 kHz) then the sampled signal will appear as a 1 V, 1 kHz tone and a 0.1 mV, 100 Hz tone (the 39.9 kHz tone aliases to 100 Hz and is attenuated 80 dB). The amplitude of the alias is dependant on the original amplitude of the out-of-band components and the amount of attenuation in the AAF. The effect is harder to analyze in the more realistic case of broadband energy that must be rejected. All of the broadband energy will fold back into the analysis band. In general, the AAF attenuation must be chosen considering the desired noise floor and the frequency content of the energy that needs to be rejected.

The next consideration is the width of the transition band. Consider designing a system with a useful frequency bandwidth of 20 kHz and 80 dB of alias protection. If the sampling frequency is 40 kHz, the AAF gain must change from 0 dB at 20 kHz to –80 dB at just over 20 kHz. We must increase the sampling frequency to make the filter realizable. Consider a sampling frequency f_s = 51.2 kHz. The Nyquist frequency is 25.6 kHz, which means that frequencies above 31.2 kHz will fold back into the band of interest. Therefore, the AAF gain must go from 0 dB at 20 kHz down to –80 dB at 31.2 kHz (see Figure 3). The region between the highest useful frequency and the Nyquist frequency is known as the guard band. Frequencies in this range will be attenuated and may suffer from aliasing, and are usually discarded in the presentation of spectral results.

So far we have considered the rejection band attenuation and guardband width as performance limits in AAF design. Three more error sources are passband variation, dispersion, and channel to channel match. These error sources are in the passband, and are thus very important in determining the overall performance of the system. Passband variation creates absolute accuracy errors, while dispersion, or nonconstant group delay, spreads out signals over time. Channel to channel match is important when making cross channel measurements. Low variation and low dispersion are both desirable, but are hard to achieve with high order analog filters. Channel to channel match is compromised by analog component variations. In the above example, sampling at 51.2 kHz and keeping 20 kHz of information required an 8th order elliptic filter. This filter has high variation and dispersion across the passband and is difficult to fabricate.

If we could increase the sampling rate, we could make the filter less aggressive, thus reducing the variation and dispersion and making it easier to manufacture. We could even sample at a very high rate, then perform digital filtering (Figure 2). It is much easier to make low variation filters digitally, and dispersion can be essentially eliminated. Digital filters are trivially duplicated across channels. This technique is used in a class of analog to digital converters called delta-sigma ADCs. For a 51.2 kHz sample rate, these converters often sample at 3.2768 MHz, making the Nyquist frequency 1.6384 MHz. A 3rd order Butterworth filter provides sufficient anti-alias protection in this situation. A Butterworth filter is maximally flat in the passband, has maximally low dispersion, and is easier to fabricate than an elliptic filter. Additionally, the lower order means there are fewer parts, improving the channel to channel match. The rest of the filtering and sample rate conversion is performed digitally inside the ADC with digital filters that can be designed with essentially zero dispersion.

The advantage of the Butterworth anti-alias filter in dispersion performance is significant. Dispersion corresponds to filter delay that varies with respect to frequency. For instance, a pure delay corresponds to a linear phase shift. If the phase is not a straight line, the delay will be different for different frequencies. Thus, broadband signals, like transients, will be spread in time, or dispersed. The derivative of the phase with respect to frequency provides a measure of the delay. If we examine this function for the two proposed AAFs shown in Figure 4, we see that the 8th order elliptic filter has over

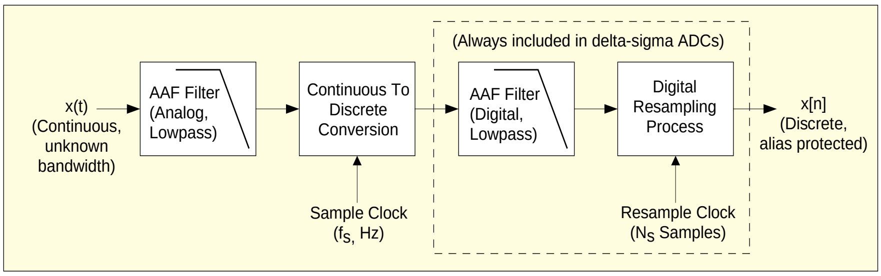

Figure 2. An anti-aliasing filter is required before sampling or sample rate conversion.

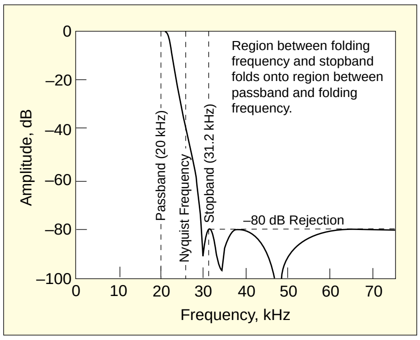

Figure 3. 8^{th} order elliptic analog AAF for 51.2 kHz sampling, 20 kHz useful bandwidth.

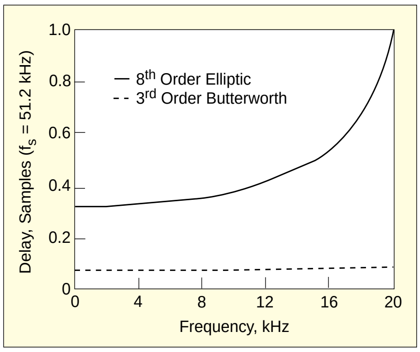

Figure 4. Delay versus frequency over 20 kHz passband for two proposed AAFs.

0.6 more samples of delay at 20 kHz than when near DC. However, the 3^{rd} order Butterworth filter has less than 0.015 samples of delay difference. This filter will not noticeably change the shape of any time domain pulses.

What if we want to sample at a lower rate? The examples we've discussed so far use 51.2 kHz sampling to acquire 20 kHz information. What if we want to analyze data to 10 kHz? We could sample at 25.6 kHz to keep the ratio of passband to stopband the same. Then the Nyquist frequency would be 12.8 kHz for normal sampling, and 819.2 kHz for delta-sigma ADCs. The AAF must be scaled with the sampling frequency. In fact, as the sample rate is reduced, the aliasing band might encounter more energy than initially thought. For instance, 819.2 kHz is in the middle of the AM radio band, exposing the delta-sigma ADC to a strong source of aliasing energy. If we go to very low sample rates, the aliasing energy may include mechanical and acoustic signals, which will definitely be present in the acquired data. It is very important that some AAF be used, whether purely analog or a mix of analog and digital. To simplify analog design and manufacturing, the ADC can be used at a fixed sample rate. Then digital filters are implemented in a DSP (Digital Signal Processor) for further sample rate conversion.

The question might arise, should the acquisition system use an AAF at all? After all, oscilloscopes do not include any filtering. Oscilloscopes, however, are capable of significantly higher sample rates than typical frequency domain oriented instruments. A state of the art scope boasts GHz sample rates. It is quite easy to start at the high end of the sample rate range and work down, observing if the character of the signal changes. Dynamic signal analyzers and vibration controllers, on the other hand, have sample rates that do not go much above the frequencies they are used to measure, so aliasing is a larger concern.

Certain analysis techniques, such as shock response spectrum calculations, have also historically not included filtering. Suppose the signal is inherently within the frequency band of interest. An AAF with high dispersion will tend to alter the time domain shape of the signal, spreading it out. However, today's AAF architectures create almost no dispersion, so the time domain shape will be unchanged. In this case, using the AAF makes no difference in the final result. Now suppose the signal has energy beyond the AAF cutoff. If you use the AAF, this energy will be attenuated, providing a more correct answer. If you do not use the AAF, this energy will be folded into the analysis band, contaminating the answer in a manner that is impossible to characterize.

As we have discussed, the purpose of an AAF is to attenuate (not eliminate) frequencies that may alias into the analysis band in the course of sampling. Aliasing scrambles frequency content and is impossible to correct after sampling. Important characteristics for an AAF are minimal variation in the passband, linear phase shift for low dispersion, width of the transition band, and attenuation in the stopband. A combination of analog and digital filters, with appropriate sample rate conversions, are used in current designs of high performance sampling systems. The use of an AAF is always recommended when sampling dynamic signals.

The author can be contacted at: rwelaratna@dataphysics.com.