Digital Sampling According to Nyquist and Shannon

Thomas L. Lagö, Jonkoping University, Jonkoping, Sweden

Today, a digital representation of the world around us has become almost mandatory. This makes the conversion of an analog signal into a digital representation a frequent requirement. When performing this conversion, it is of interest to be sure that it is possible to correctly reproduce the analog signal from the record of digital data samples collected. The sampling procedure has been covered in many text books, especially in electrical engineering and for associated analog time data. It is important to remember that the strict rules described in early papers by Nyquist and Shannon must be followed. If they are violated, it might not be possible to find the real world signal in the sampled data set. The sampling rules will be covered in this article and the result may surprise some groups, especially nonelectrical engineers.

When converting an analog signal into a digital representation, discrete samples are collected often at equidistant points or steps. The number of samples needed to fully represent an analog time signal becomes a crucial question to answer. It is obvious that if one took an infinite number of samples, it would be easy; but if too few are collected, it is not possible to reconstruct the analog signal. A typical ADC system is illustrated in Figure 1 and suggested sample points are indicated along the analog time signal axis.

Figure 2 shows that it is possible to draw any number of other signals through the recently collected data points. Therefore, it is not possible to be sure that a reconstruction of our original signal is valid, unless some limitations to the analog signal and the corresponding sampling rate are defined. This sampling technique or requirement is often referred to as "The Sampling Criterion," "The Sampling Theorem," "Nyquist Sampling Criteria," or something similar. There are many names and descriptions for this process and this article will discuss the validity of these criterions and their assumptions.

When studying Figure 2, it is intuitive to differentiate the three signals in terms of frequency content. The one that "moves the most" between the data points must have a larger bandwidth. The one that "moves the least," must have the lowest bandwidth or frequency content. This leads us to believe that the sampling rate should be coupled to the frequency bandwidth or frequency content of the signal.

In many text books and articles, especially books on Mechanical Engineering, it is common that the sampling theorem is stated as: *the sampling frequency must be larger than twice the maximum frequency.*[3,4,5,6] Many other texts could have been included in this list, but those cited were selected as examples only, and there is no intent to discredit their otherwise great content, but to highlight the widespread misconception of their findings. It is also common to refer to Shannon when making this sampling selection statement, thus validating its accuracy. The sampling formulation above is unfortunate and could easily lead to a misunderstanding or incorrect interpretation of how a suitable sampling rate should be selected.

An article in *Sensors Magazine*, July 1999,[7] brought up the sampling selection issue and some of the misconceptions were discussed in order to clarify the matter. In the text, a discussion on why Shannon used the terminology "maximum frequency" could easily lead a reader to wrong conclusions. The article stated: "If you need to measure a 1 kHz component, Shannon's theorem would demand that you sample at >2× the component (see the sidebar)." The sidebar then states, and refers to Shannon: "if a function $f(t)$ contains no frequencies higher than W cps (Hz), it is completely determined by giving its ordinates at a series of points spaced $1/2W$ sec apart. This has come to be known as Shannon's sampling theorem, or simply the sampling theorem, often misnamed the Nyquist theorem." The formulation in this article is unfortunate and could easily lead to a wrong understanding or interpretation of how a correct sampling rate should be selected. It also illustrates how widely the "twice the maximum frequency" sampling formula has been disseminated.

Both Nyquist and Shannon fully understood how to select the sampling rate and why. However, they talked about different types of signals, *general* versus so-called *baseband* and Shannon explained this difference clearly before stating his well known *Theorem 1*, that will be presented later. This means that Shannon assumed the signals started at zero Hz and had signal information content up to f_{max} Hz, and Nyquist *did not.* Nyquist talked about the general definition using the bandwidth W and never discussed the restriction using baseband signals. Figure 2 illustrates the difference in spectral content for the two cases, in which the sampling rate is the same, but the maximum frequency is different.

Shannon mentions Nyquist in "A Mathematical Theory of Communication" from 1948,[2] but makes reference to another article for a more thorough discussion.[1] Shannon points out in his article that Nyquist assumed the time signals were strictly limited to a time period T but still had a bandwidth W.[2] Such signals have, however, unlimited bandwidth and thus it is not possible to fulfill both criteria simultaneously. By making reasonable restrictions in time and frequency this can be handled – like the time function being very small outside the time period T. Shannon's article discusses bandwidth, but sometimes he hides this by using wording like "components up to." This is unfortunate. If a signal contains frequency components up to a certain frequency limit, within the dynamic range of interest, the sampling rate as discussed earlier would be correct. This assumes that "up to" implies that the starting frequency is zero Hz. The problem with most of these statements is that the zero to f_{max} discussion, as a definition of the bandwidth, is very often hidden by improper wording. Both Nyquist and Shannon stated the sampling theorem in the same manner, but used different wording. So there is no real conflict and/or difference between them.

Several books are explicit. Most of them are mathematical signal processing books. It is rare to find a more practical book to be correct, but Hewlett-Packard's[8] is a good example. However, application notes [9,10] use a less correct sampling rate statement, unless the statement assumes a baseband signal, which is not clearly expressed. It is interesting to note that the instrumentation products that the application notes are written for do not use the maximum frequency for sampling determination, but the bandwidth. Thus, there seems to be confusion in text books and articles whether the maximum frequency or bandwidth should be used when determining the sampling frequency. Also, the definition of bandwidth is not expressed. It is possible to find the filter bandwidth used as the bandwidth definition for sampling. That is not correct and will yield a very low signal-to-aliasing distortion ratio.

It is of interest to revert back to the original text written by Shannon. He said the following:[2]

II. The Sampling Theorem. Let us suppose that the channel has a certain bandwidth W in cps starting at zero frequency, and that we are allowed to use this channel for a certain period of time T. Without any further restrictions this would mean that we can use as signal functions and functions of time those

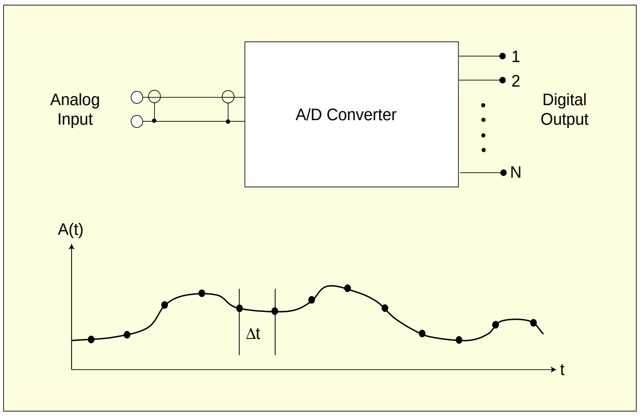

Figure 1. Illustration of a typical ADC system and suggested sample points indicated along the analog time signal. The A/D converter can be of many different types but that is not of importance for this article.

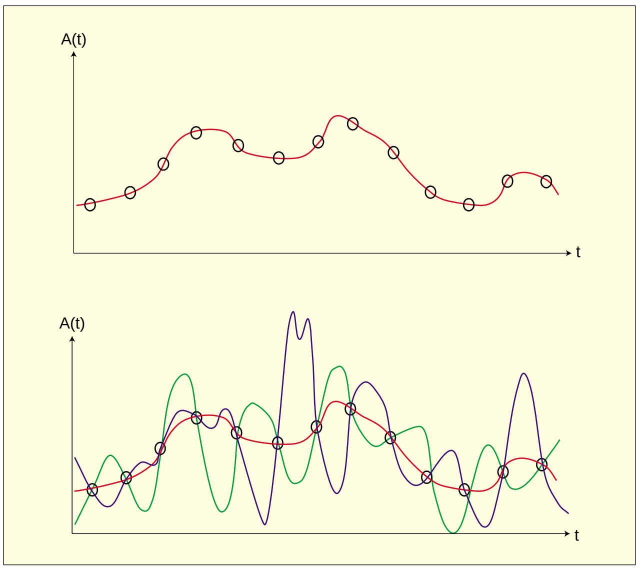

Figure 2. Illustration of a sampled signal and how it is possible to superimpose many different analog signals that will go through the exact same sample points. This fact makes the sampling process ubiquitous and that is not good. By adding some restrictions to the analog signal and making sure the sampling rate is in harmony with these restrictions, we avoid this unfortunate situation.

whose spectra lie within the band W and whose time functions lie within the interval T. Although it is not possible to fulfill both of these conditions exactly, it is possible to keep the spectrum within the band W, and to have the time function very small outside the interval T. Can we describe the functions which satisfy these conditions in a more useful way? One answer is the following:

Theorem 1. If a function $f(t)$ contains no frequencies higher than W cps, it is completely determined by giving its ordinates at a series of points spaced $1/2W$ sec apart. It is possible to interpret *The Sampling Theorem and Theorem 1* incorrectly. Shannon talks about the "band W." He assumes, as stated in the beginning, that the starting frequency is zero Hz. Thus, the bandwidth W is equal to the maximum frequency denoted f_{max} and given by W. This is the case for baseband signals only, and not a more general statement where the band W must be used. He explains this very clearly, except in the theorem, and thus many readers have had difficulty understanding the difference between f_{max} and W. Maximum frequency f_{max} and the bandwidth W may imply the same sampling frequency f_s for baseband signals, but *only* for baseband signals. Thus, there is no difference between Nyquist's explanation and Shannon's. They are both based on the bandwidth W, which is the correct way of selecting sampling rate.

The use of maximum frequency does not apply for a bandlimited signal around some carrier *frequency* f_c. The con-

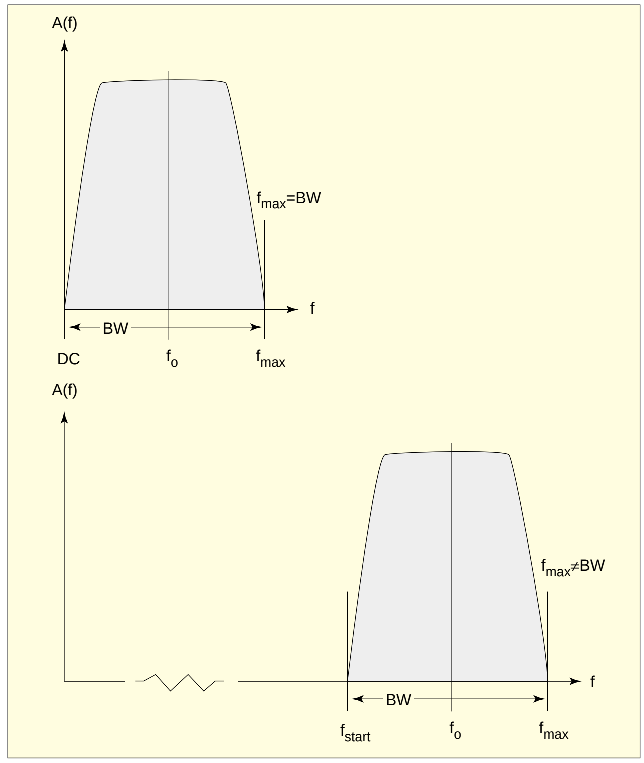

Figure 3. Illustration of the difference between a baseband signal and a generalized signal with frequency content in a band and the starting frequency separated from DC. The bandwidth BW is the same in these two cases, but the maximum frequency can be totally different. The bandwidth could be 1 kHz, but the maximum frequency, if we use an ultrasonic transducer, could be in the MHz region. Please note that the maximum frequency notation used is not equal to the one commonly used for filter applications.

tinuous-time signal $x_r(t)$ representing an ultrasonic signal for example, is typically band limited. However, there is often a center frequency, e.g., 1 MHz. It is only the signal *information* centered around the carrier frequency that is of interest. This provides a very good opportunity to "under sample" relative to the carrier frequency but not relative to the sampling theorem. Therefore, it is possible to fulfill the Nyquist sampling theorem for the narrowband signal that represents the ultrasonic signal, yet sample with only $2\times W$, where W is the bandwidth of the signal around the center frequency f_c. It is not a contradiction, due to the limited *bandwidth*. Also, a correct sampling based on the bandwidth W is sparse versus the frequency content which is most important when performing FFT analysis (Fast Fourier Transform). If this is not the case, very large block lengths must be used for the FFT analysis and/or a resampling of the data must be performed before the FFT calculation can take place. Both methods require computer power and complicate the analysis. A common selection of sampling rate in FFT analyzers is 2.56× the bandwidth (*not frequency* if they can do real zoom analysis).[9,10] For this case only 20% of the frequency lines are wasted whereas in an analyzer with a 10:1 sampling rate selection, 80% of the lines are wasted.

Example. Assume an ultrasonic transducer has a 1 MHz center frequency and a bandwidth of 1 kHz. If f_{max} is used as the rule for selecting sampling rate, f_s >2 MHz. But using the bandwidth W, f_s is only >2 kHz. There is a 1000:1 decrease in sampling rate without loss of information. If we are interested in the frequency information within the 1 kHz band, an FFT block size of 2,097,152 (1024×2048) samples must be used for $f_s=2\times f_{max}$, but $f_s=2\times W$ only leads to 2048 samples, a great saving of storage and complexity. The computation time for the 2 M sample FFT is enormous in comparison to the 2048 block FFT. Most of the frequency lines when using f_{max} sampling must be disregarded for a situation like this. This is why most FFT analyzers sample versus the bandwidth and not f_{max}, despite what their application notes may declare.[9]

There is one disadvantage when selecting the sampling rate versus the bandwidth W for the example above. The time signal does not look 'right' since the "carrier frequency" is gone due to bandwidth sampling. The time signal can be reconstructed and thus the original signal can be reconstructed, but this implies interpolation and filtering. For example, oscilloscopes commonly display the time samples only, without interpolation, even for bandlimited signals and in this case one must sample versus f_{max} and not the bandwidth. This has nothing to do with Nyquist and/or Shannon. It is simply one method to avoid a more complex handling of the data reconstruction. If you know that the signal will be "looked at" in the time domain, this could be a good strategy.

Rule of Thumb Sampling Example

When deciding the sampling rate, it is clear from the previous discussion that we need to define the bandwidth W for the signal to be sampled. All real systems have background noise and thus, in reality the actual bandwidth is unlimited. The analog to digital converter's bit number will also set the quantization noise floor. If we then tie the expected dynamic range from the analog to digital converter directly to the bandwidth as described by Figure 3, it is possible to develop an easy "rule of thumb" for selecting the sampling rate as long as the following applies:

- The analog noise floor is below the quantization noise floor
- The maximum input spectrum is "white noise." This implies that the spectrum from the sensor signal must not increase with frequency. If the input signal decreases with frequency, the sampling rate will be conservative.
- The anti-aliasing filter is attenuating the signal several octaves above the cutoff frequency.
- There is low noise after the anti-aliasing filter.
- The anti-aliasing filter must have a linear slope in dB.

If this applies, we can formulate a rule of thumb as:

$$f_s = f_{max} + f_g \qquad (1)$$

The sampling rate should thus be selected by adding the cutoff frequency of the anti-aliasing filter (–3 dB, or the point specified by the filter designer as the filter bandwidth) and the frequency –X dB down as specified by the anti-aliasing filter. Figure 4 illustrates this.

Example. Assume the anti-aliasing filter used has a cutoff frequency of 10 kHz. The filter used has a slope of 80 dB per octave. We are using a 16 bit ADC and we are expecting a dynamic range of at least 80 dB. The filter's frequency bandwidth 80 dB down will thus be 20 kHz (one octave equals a doubling of frequency). The sampling frequency should thus be 10+20 kHz = 30 kHz. If we would like a 96 dB dynamic range, we must increase the sampling rate or increase the slope of the filter.

Summary

For baseband signals, it is correct to interpret the sampling theorem as $2\times f_{max}$ since f_{max} and the bandwidth W, will be the same as the maximum frequency, but for this case only. Both Nyquist and Shannon understood this and both used the bandwidth as the key parameter in their mathematical analysis of sampling rate selection. However, Shannon talked about a bandwidth W starting at zero Hz and that could be interpreted as the maximum frequency. Without the restriction starting at zero Hz, it would be a different statement. Many authors have studied Shannon's text, taken W cps and then drawn the false conclusion that this is what should be used when selecting the sampling rate, even for a generalized signal. This leads to a limitation of the Sampling Theorem. For some applications this can lead to more complicated systems and A/D converters with a higher power consumption than necessary. There are many more implications. A correct choice of sampling rate will lead to an efficient ADC system and be helpful when further analysis is to be performed, especially for FFT analysis. The real truth of the sampling theorem is that it is the bandwidth W that is the key to a correct sampling rate selection process. This implies that, for many cases (such as ultrasonic and telecommunication signals), it would be incorrect to use f_{max} as a sampling criteria.

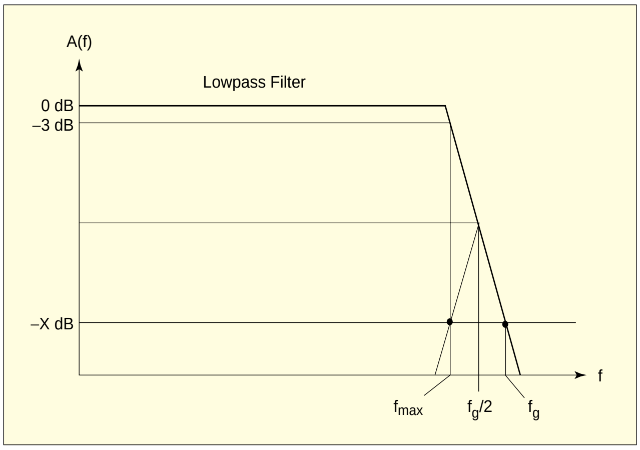

Figure 4. Illustration of a typical anti-aliasing filter for an ADC system. The –X dB represents the desired dynamic range. The –3 dB point (or the filter cut-off definition) is used to determine the maximum frequency the filter allows to pass undistorted, from an amplitude point of view.

Another misconception is that when performing sparse sampling, as given by the Sampling Theorem,[1,2] it is not possible to fully reconstruct the analog signal. This misconception can also be described as: *it is not possible to find the values in between the samples*. However, it is most important to keep in mind that the sampling theorem states that it is possible to fully reconstruct the analog signal. This implies that any point in between the digital samples is known (the analog signal is completely determined). However, one must reconstruct the data and interpolate with a correct algorithm. A correct digital FIR (finite impulse response) filter or an analog filter needs to be used to find these signal values depending on whether more samples are needed in the sampled domain or in the analog domain.[11,12] When not done properly, there will be an error. This is not a limitation of the sampling theorem, but of the reconstruction filter used. A correct sampling based on Nyquist's or Shannon's theorems, together with a good reconstruction filter will lead to a good analog signal that is 'equal' to the analog signal we started to sample. A good example of this is a high-performance CD or DVD player. By using the bandwidth W as a selection criterion, we know that we will have the lowest possible sampling rate while still maintaining the signal information.

References

1. Nyquist, H., "Certain Topics in Telegraph Transmission Theory," AIEE Trans., vol. 47, pp. 617-644, 1928.
2. Shannon, C. E., "Communication in the Presence of Noise," Proceedings IRE, vol 37, pp. 10-21, 1949.
3. Broesch, James D., *Digital Signal Processing Demystified*, HighText Publications, 1997.
4. Zölzer, Udo, *Digital Audio Signal Processing*, John Wiley & Sons, 1995.
5. Dorf, Richard C., *Electrical Engineering Handbook*, CRC Press, 1993.
6. M. P. Norton, *Fundamentals of Noise and Vibration Analysis for Engineers*, Cambridge University,1990.
7. National Instruments, "Data Acquisition," *Sensors Magazine*, Page 45, July 1999.
8. White, Robert A., *Electronic Test Instruments, Theory and Applications*, Hewlett-Packard Professional Books, Prentice Hall, 1993.
9. Application note AN-243, "Fundamentals of Signal Processing," Hewlett-Packard Company, 1981.
10. Application note AN-243-1, "Effective Machinery Measurements Using Dynamic Signal Analyzers," Hewlett-Packard Company, 1991.
11. Lagö, Thomas L., Davis, Larry, and Preti, Giuseppe, "Time and Frequency Domain Analysis Considerations In Automotive Applications," Proceedings of International Conference SIAT-2001, Society for Automotive Engineering, SAE, India, 2000.
12. Proakis, John G., and Manolakis, Dimitris G., *Introduction to Digital Signal Processing*, Macmillan, New York, 1998.

The author may be contacted at: thomaslago@msn.com.