

Modern Sampling: A Tutorial

JAMIE A. S. ANGUS, *AES Fellow*

(j.a.s.angus@gmail.com)

University of Salford, Salford, Greater Manchester, M5 4WT, UK

High-resolution audio has started to use modern sampling principles to further enhance the quality of digital audio reproduction. This tutorial seeks to explain how these new methods are an improvement over traditional Shannon sampling. The tutorial first reviews the relevant parts of traditional sampling and then goes on to introduce the modern method based on splines, which are one of many possible approaches. It shows that it is possible to provide better signal reconstruction with practical filters compared to traditional sampling. Where possible a non-mathematical approach is used.

1 INTRODUCTION

Sampling is a critical process in digital audio. Sampling is the process of converting a continuous time signal that exists for all time values, into one that exists only at discrete time values. Continuous time signals must be converted to discrete time ones in order to provide a list or sequence of signal values that can be processed or stored.

In audio this is usually achieved by taking snapshots, or samples, of the signal at regular time intervals T_s and therefore at a constant sampling frequency F_s . The process can be described as the product of a series of sampling impulses and the original signal, as shown in Fig. 1.

The most obvious application of sampling is at the analog to digital interface, where the analog signal must be sampled so that it can be quantized into a digital word. Likewise at the digital to analog output the sampled signal must be converted back, or reconstructed, into a continuous time one. However, sampling, or re-sampling, also occurs implicitly in any process that changes the sample rate of the incoming digital signal. In these sample rate change processes the sampled signal must be converted to a continuous, or very high sample rate, signal and then re-sampled at the new output sample rate.

Thus sampling processes may occur both within an audio processing system, as well as at its analog inputs and outputs. It is also important to note that sampling, the discretization of the signal in time, is independent of the process of quantization, which is the discretization of the input signal in amplitude. Although in many analog to digital converters the two operations are entwined, they remain distinct processes theoretically.

Sampling can be, in principle, “lossless” in that, under the right conditions, a finite bandwidth continuous signal

can be sampled and reconstructed without error, as shown by Shannon, Nyquist, Whitaker, and Kotel’nikov [1–4]. On the other hand, quantization always causes some error to the signal, as either distortion or noise and, thus, is never “lossless.”

This tutorial concerns itself with only the sampling process. The reason for this is that modern sampling approaches are being used in some modern high-resolution audio formats such as MQA [5]. There has also been research [6] that seems to show that wider bandwidth, and thus higher sampling rate, material may be preferred over 44.1 kHz or 48 kHz sampled material. High-resolution audio is concerned with obtaining the highest possible audio quality and modern sampling approaches allow better reconstruction of the audio waveform with finite length filters compared to the traditional approach.

This tutorial will look at modern theories of sampling, and explain them, in as much of a non-mathematical way as is possible. Where there are equations, they are there to add some detail and rigor to the explanations. However, they can be skipped over without compromising the explanation and be revisited, if required, at the reader’s leisure.

We will first examine the traditional sampling process and the classical theory behind it. Then the modern techniques will be discussed. It will show that sampled audio, when properly reconstructed, preserves more of the spectrum of the original signal.

2 TRADITIONAL SHANNON-NYQUIST SAMPLING: THE FREQUENCY DOMAIN EXPLANATION

Multiplying a time domain signal by a series of regularly spaced impulses results in copies or aliases of the

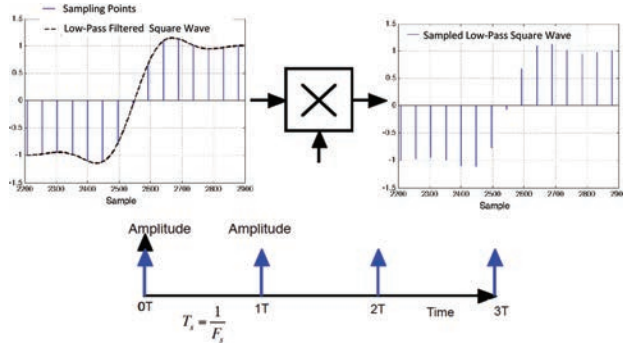


Fig. 1. The sampling process.

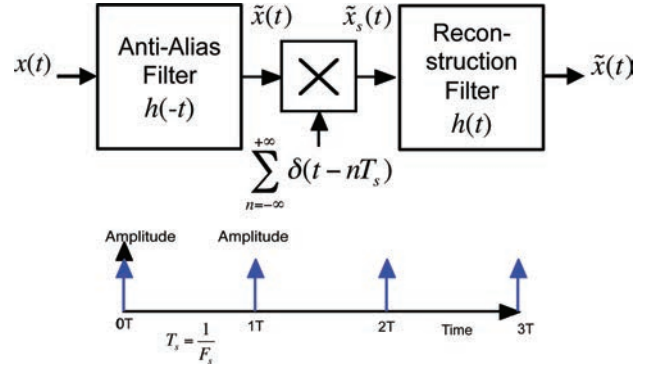


Fig. 4. A sampling process including the anti-alias filter.

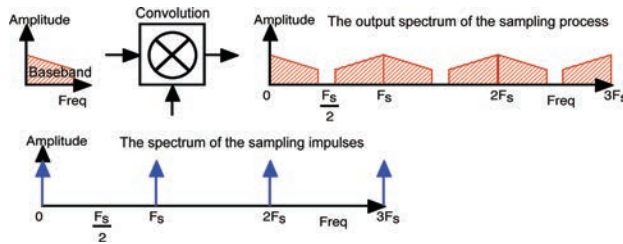


Fig. 2. The spectrum of a sampled signal.

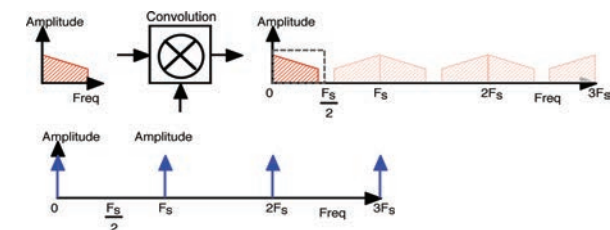


Fig. 3. Using a brick-wall low-pass filter to reconstruct the sampled signal.

original baseband signal at multiples of the sampling frequency, as shown in Fig. 2. Traditional Shannon-Nyquist sampling uses a brick-wall low pass filter to perform the reconstruction of the continuous signal from its sampled version, as shown in Fig. 3.

Providing the aliases do not overlap, the original baseband signal can be separated from the aliases and reconstructed. Of course if the aliases do overlap then the original baseband signal cannot be separated from the aliases and cannot be reconstructed. In order to prevent any overlap occurring it is usual to place a matching brick-wall anti-alias low pass filter before the sampler, as shown in Fig. 4.

Unfortunately, as brick-wall low pass filters have infinitely long impulse responses, perfect separation is not possible in practice and some reconstruction error will occur.

Although the frequency domain explanation is straightforward it only focuses on the frequency domain effects of the sampling process and so hides some important details.

Another way of looking at the sampling process is to consider, as Shannon did, the sampling process in the time domain.

3 TRADITIONAL SAMPLING: THE TIME DOMAIN EXPLANATION

For the time domain explanation we will assume that the sampling system contains a pre-filter and a reconstruction filter, as shown previously in Fig. 4. In order to understand the time domain explanation we need to realize two essential aspects of signals and filtering.

1. Any continuous time signal $x(t)$ can be described as a succession of samples with a sampling time equal to zero seconds, or an F_s of ∞ Hz.
2. Any linear filtering process can be described as the convolution between the input signal and the filter's impulse response, which represents the response of the linear filter to a single impulse applied to its input.

These assumptions allows us to visualize the effect of the various filters in the sampling process via the convolution theorem, which states that the output of a linear filter can be described as the convolution of the input signal and its impulse response, as shown in Eq. (1).

$$y(t) = \int_{-\infty}^{\infty} x(t - \tau)h(\tau)d\tau \tag{1}$$

Where: $y(t)$ = the convolution result

$x(t)$ = the input signal

$h(\tau)$ = the impulse response

t = time

τ = the time index for the impulse response

If a train of impulses is convolved with an impulse response then the impulse response is both duplicated at each impulse and scaled in proportion to each impulse's amplitude. It is also added to any overlapping duplicated impulse response arising from the other impulses, as shown in Fig. 5. Note that in this example a simple triangular impulse response is shown to aid clarity. In practice the filter's impulse response would be more complex. Eq. (1) essentially describes this process assuming that the time distance between the train of impulses tends to zero seconds as per our continuous signal defined in aspect 1.

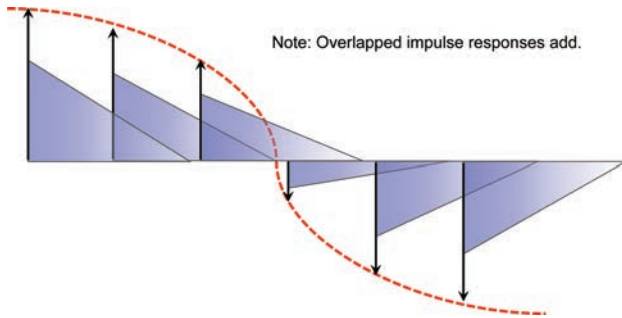


Fig. 5. Convolution of a filter impulse response with samples.

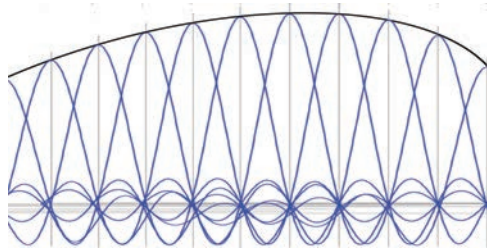


Fig. 6. Summing $\sin(x)/x$'s together.

For the brick-wall low pass filters used in traditional Shannon-Nyquist sampling the impulse response is a $\sin(x)/x$ function. Unfortunately, as discussed in Sec. 2, this impulse response is acausal because it extends over infinite time. So, for practical filters, the impulse response is truncated, using a window [7], and a finite time delay is imposed. This has the effect of limiting the roll-off rate and stopband performance of practical filters. For pedagogical purposes we will assume ideal brick-wall filters, as Shannon did in his original paper [1], in the following discussion of traditional sampling in the time domain.

3.1 The Role of the Reconstruction/Synthesis Filter

Filtering is convolution; thus what the reconstruction filter does is replace each of the weighted Dirac delta sample values with its impulse response. Shannon proposed a brick-wall lowpass filter, which has an impulse response equal to a $\sin(x)/x$ function as described in Eq. (2).

$$h(t) = \frac{\omega_0}{\pi} \frac{\sin(\omega_0 t)}{\omega_0 t} \tag{2}$$

Where: $h(t)$ = the impulse response
 t = time
 ω_0 = the cut-off frequency of the filter, $\frac{2\pi F_s}{2}$

This filter would therefore take each sample value and replace it with a $\sin(x)/x$ impulse response.

It is possible to show that when these weighted $\sin(x)/x$ functions are added together a perfect facsimile of the original signal is produced, as shown in Fig. 6.

What are the characteristics of a $\sin(x)/x$ function that allow this remarkable thing to occur?

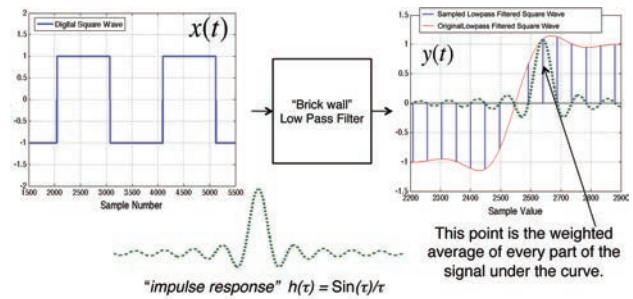


Fig. 7. The effect of a brick-wall low-pass filter on the continuous input signal.

$$y(t) = \int_{-\infty}^{\infty} x(t - \tau)h(\tau) d\tau$$

1. If a constant value is sampled and each sampled value is replaced by the $\sin(x)/x$ function then the resulting continuous signal is also constant.
2. The $\sin(x)/x$ function is orthogonal to itself for time shifts equal to multiples of the sampling period $T_s = \frac{1}{F_s}$ as described in Eq. (3).

$$c(k, l) = \int_{-\infty}^{\infty} \left(\frac{\sin\left(\frac{\pi}{T_s}(t - k)\right)}{\frac{\pi}{T_s}(t - k)} \right) \times \left(\frac{\sin\left(\frac{\pi}{T_s}(t - l)\right)}{\frac{\pi}{T_s}(t - l)} \right) dt \tag{3}$$

Where: $c(k, l)$ = the cross-correlation result

k, l = integer time shifts of the waveform equal to multiples of T_s

T_s = the sampling period

t = time

In Eq. (3) $c(k, l)$ is equal to the Dirac delta function. That is it is zero whenever k is not equal to l ($k \neq l$) and one when k is equal to l ($k = l$).

These two properties allow the reconstruction filter to reproduce the original waveform without error, providing there are no frequencies that are above $\frac{F_s}{2} = \frac{1}{2T_s}$. This is because the addition property (characteristic 1) means you get a constant output for a constant input, and the orthogonality property (characteristic 2) means that each sample value does not affect any other sample values.

In order to enforce this restriction the input signal is usually pre-filtered by a brick-wall low-pass filter with a cutoff frequency of $\frac{F_s}{2}$, as shown previously in Fig. 4.

In the time domain the effect of this filter is to replace every instantaneous value of the signal with a value that represents the convolution of the input signal and the $\sin(x)/x$ impulse response, as shown in Fig. 7. Because of the orthogonality properties of $\sin(x)/x$, one can consider this filtering operation to be a linear transformation from the space of non bandlimited signals into the space of bandlimited signals. This is analogous to the Fourier transform, which

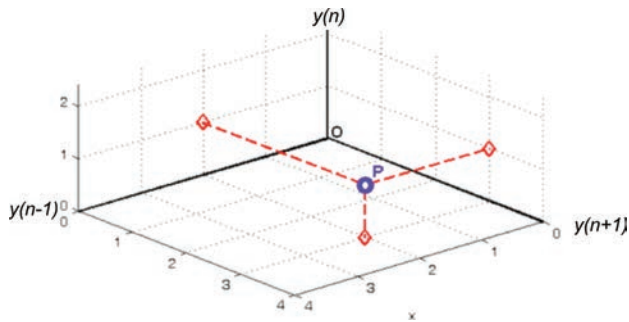


Fig. 8. The $\text{sinc}(x)/x$ transform of the continuous signal.

linearly transforms functions in time space, to functions in frequency space. Although it should be noted that this is only strictly linear if the original signal before the filter was already band limited to $\frac{F_s}{2}$. If the signal spectrum extended above $\frac{F_s}{2}$ then the frequencies above it would be lost and therefore the reconstruction would be in error. Each filtered sample time point in the continuous space now represents a coordinate on an axis in the band-limited space, as shown in Fig. 8 for three adjacent samples.

In Fig. 8 the point P represents a unique value that consists of all the samples in the signal and contains all the information within it, except for any frequencies that have been removed by the pre-filter. Thus the sampling operation selects the coordinate values from the band-limited space, and these coordinate values uniquely describe the (filtered) continuous input signal, and the reconstruction filter takes these coordinates (samples) and uses them to produce the original signal. For filtered noise these samples will be independent.

This is the essence of what Shannon was trying to say in his paper. That by pre-filtering a signal such that its highest frequency is $\frac{F_s}{2}$ one is left with a band limited signal whose information content is completely described by F_s samples, and that the missing bits between the samples can be completely reconstructed using the same $\text{sinc}(x)/x$ impulse response that was used to pre-filter the original signal.

3.2 What's the Problem with Shannon-Nyquist Sampling?

The problem with using a $\text{sinc}(x)/x$ function, as the impulse response of a filter to reconstruct the signal, is that while it works perfectly in theory, it doesn't in practice! This is due to the infinite extent of the $\text{sinc}(x)/x$ function in time, that is, it requires *infinite support*. In order to realize practical anti-alias and reconstruction filters we have to truncate the time extent of the impulse response of the filter. This has the following consequences.

1. The filter no longer has an infinite rate of cut-off and thus needs a *guard band* between the upper frequency of the continuous signal and the lowest frequency of the first alias.
2. More subtly, because the impulse response is now finite in extent, it is impossible to realize a stop-

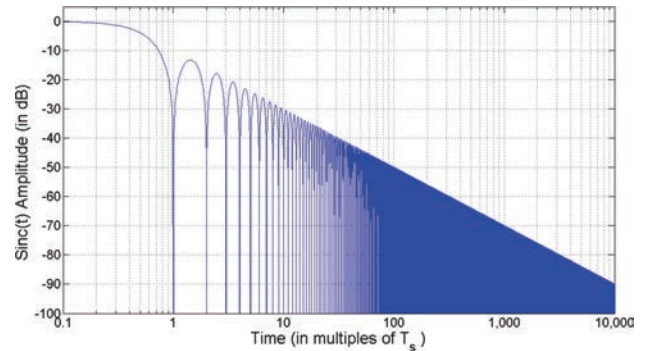


Fig. 9. The slow time decay of the magnitude of the $\text{sinc}(x)/x$ impulse response.

band response of zero ($-\infty$ dB of attenuation) over the whole frequency range of the stop band.

3. In fact, unless the stop band achieves infinite attenuation at infinite frequency, there will always be some alias components present even if the sampling frequency is increased to infinity.
4. The truncated $\text{sinc}(x)/x$ functions no longer add up to a constant value when the sampled continuous signal is constant. This means that there is some difference between the reconstructed signal and the original signal.
5. The truncated $\text{sinc}(x)/x$ functions are no longer orthogonal for a time shift equal to multiples of a sample period. This means that the samples are no longer projected properly onto a sampleable space and therefore the samples will have leakage into each other (alias distortion), even if the continuous signal was white noise.

For these reasons the reconstructed signal no longer matches the original continuous signal because it has errors between the sample points. These errors are due to the truncation of the infinite $\text{sinc}(x)/x$ impulse response. They are often called alias errors as they can be considered the result of the aliases due to sampling overlapping with the continuous signal's spectrum.

In order to reduce these aliasing errors to an acceptable level reconstruction filters are often made with extremely long impulse responses. This is because as well as needing *infinite support* the $\text{sinc}(x)/x$ function's amplitude decays very slowly, proportional to $1/t$, as one moves away from its central point, as shown in Fig. 9. As many low pass filters are in some way related to a $\text{sinc}(x)/x$ impulse response this means that they have to have long impulse responses to achieve low levels of error. Fundamentally, sampling methods based on brick-wall low-pass filtering, as proposed by Shannon and others [1–4], can never be perfect!

4 ALTERNATIVE SAMPLING STRATEGIES

As suggested by Oñativia and Dragotti [8] Shannon Nyquist sampling is so pervasive that we forget that, like any theory, there are terms and conditions attached. In particular

Shannon’s sampling theorem gives us some sufficient but not necessary conditions for perfect reconstruction. In other words, his theorem does not claim to be the only way to sample and reconstruct a continuous signal. So what would we want from an alternative sampling strategy?

Before we look at these conditions let us consider what we would want from any sampling process.

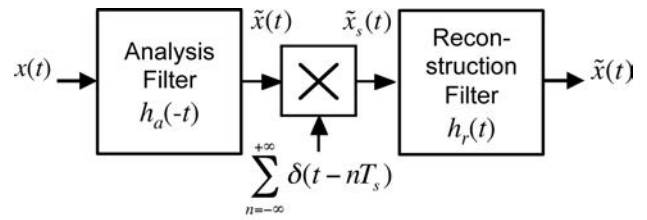
For this process to work we would need to be able to reconstruct the input waveform without error from the sample values. To be able to perfectly reconstruct the input waveform in practice requires the following four conditions to be met.

1. The individual samples should produce a constant continuous output for a set of constant input samples. This is known as the *partition of unity condition* in the literature [9]. Although not usually called this, this condition is familiar to us in audio as it is the condition that a Linkwitz-Riley crossover meets when the magnitudes of the separate crossover regions all add up to form a unity gain response over the full audio frequency range [10]. It is also the condition that a quadrature mirror filter’s magnitude response meets, as it also has a uniform frequency response over its combined frequency range [11, 12].
2. The individual samples should be independent from each other. That is, each sample will represent a part of the signal that will be unaffected by the variations in any of the other samples. They should also be linearly related to the original continuous signal. These conditions mean that the samples should be *orthogonal* and *orthonormal* to each other. Again, the concept is familiar to us in audio. For example, quadrature mirror filters have been used in audio coding because they allowed us to split a frequency band into two and subsample it, and then reconstruct it without distortion because the bi-orthogonality of the quadrature mirror filters, as well as their “partition of unity” property allowed perfect alias free reconstruction of the input signal.
3. We would like any reconstruction error to tend to zero as the sampling period tends to zero.
4. Finally we would like the associated anti-alias and reconstruction filters to have *compact support*. That is, they would be realizable using time limited impulse responses.

Unfortunately there are very few impulse responses, or *kernels*, that satisfy these criteria. The $\sin(x)/x$ kernel satisfies the first three criteria but, unfortunately, does not have *compact support*. In fact it requires *infinite support* due to its infinitely long impulse response.

4.1 Modern Sampling

However, there are two filters involved in the sampling process: the anti-alias (*analysis*) filter and the reconstruction (*synthesis*) filter, as shown in Fig. 10. Although Shannon sampling assumes so, there is no requirement that they



$h_a(-t) \neq h_x(t)$ But they form a bi-orthogonal pair

Fig. 10 A modern sampling process.

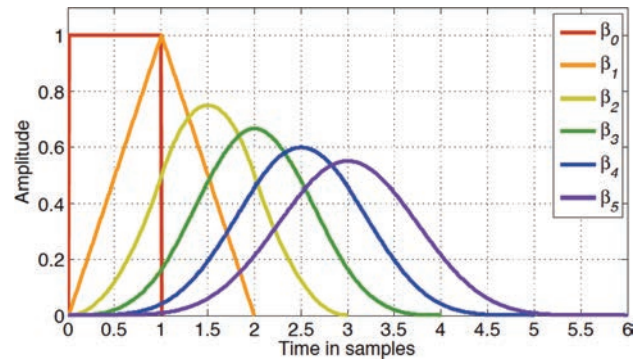


Fig. 11. The family of basic β -splines shown normalized to area but displaced for clarity.

be the same. They can be different and, providing they meet the criteria listed earlier, can provide error free reconstruction. One of the main issues being the need for the filter impulse responses to be orthogonal when shifted in time. This can be achieved by using two complementary filters that are designed to be *bi-orthogonal*. The second requirement of them forming a *partition of unity* can be realized by choosing an appropriate reconstruction *kernel*, or impulse response. Thus in modern sampling *different* but complementary filters must be used at both the input and output of the sampling process.

The question then becomes, what are suitable impulse responses, or *kernels*, that can be used as alternatives to brick-wall filters for the sampling process?

4.2 Using β -Splines for Sampling

One such possible *kernel* is the family of functions known as basic β -splines, which were proposed by Unser et al. [13–20].

These functions have been used in interpolation as they mimic the effect of the draughtsman’s spline ruler for filling in the gaps between points on a curve. They do this by having requirements on the derivatives as well as the values at the points that are being interpolated between. They are formed by successively convolving zeroth order spline, a simple rectangle of width one to yield the higher order β splines, as shown in Fig. 11. Mathematically this can be expressed as:

$$\beta^n(x) = \beta^0 \otimes \beta^0 \dots \otimes \beta^0(x) \{ \text{using } n\beta^0 \text{ spline terms} \} \quad (4)$$

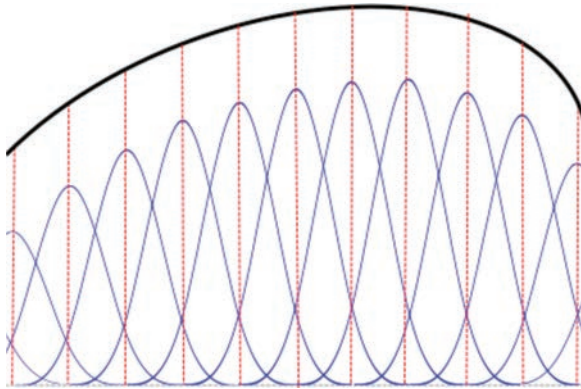


Fig. 12 Reconstruction using splines. Note the spline amplitudes are different to the sample amplitudes because of the interaction between the splines.

Where: β^n = the n th order β spline
 β^0 = the zeroth order β spline (a square pulse)
 n = the order of the β spline.

Basic β -splines have a lot of advantages for reconstructing the signal. As well as having the desirable properties mentioned earlier, they also produce no “ringing” or “overshoot” on the reconstructed waveform.

However, there is one serious problem with β -splines and that is that, except for β^0 , they all have non-zero values at sample times away from the current sample. Because of this adjacent sample values interact, and one cannot simply use the current sample value for the size of the β -spline for its reconstruction. Fig. 12 shows how this might happen in practice; note that the amplitudes of the splines are not the same as the sample amplitudes due to this overlap.

Thus calculating the *coefficients* (β -spline values for reconstruction) for the β spline curves is complex, requiring a matrix solution [21]. In the past this precluded their application due to the processing cost. However, Unser et al. [13–20] have shown that it is possible to do this calculation via a convolution, or filtering, operation that has considerably less processing cost.

They recognized that the act of using a spline *kernel* was analogous to filtering using the β -spline as an impulse response.

Their idea was to use a pre-filter whose impulse response is the *dual* of the β -spline reconstruction filter’s response, prior to sampling, as shown in Fig. 13, for a third order spline.

This pre-filter must then meet a bi-orthogonality condition when correlated with the reconstruction spline filter. Mathematically:

$$c(k, l) = \int_{-\infty}^{\infty} \varphi_{prefilter} \left(\frac{1}{T_s}(t - k) \right) \varphi_{reconstructionfilter} \left(\frac{1}{T_s}(t - l) \right) dt \quad (5)$$

Where: $c(k,l)$ = the cross-correlation result
 k,l = integer T_s time shifts of the waveform
 $\varphi_{prefilter}, \varphi_{reconstructionfilter}$ = the filter impulse responses
 T_s = the sampling period
 t = time

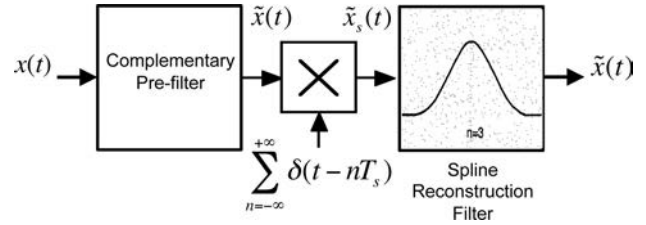


Fig. 13. Using a pre-filter instead of a matrix solution for spline reconstruction for a third order spline. (impulse response from [17]).

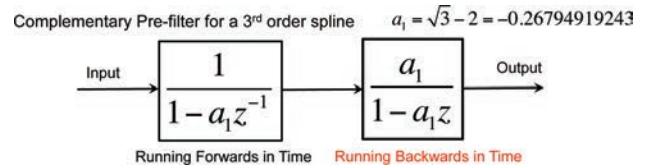


Fig. 14. Implementing the pre-filter as a cascade of two IIR filters.

In Eq. (5) $c(k,l)$ is equal to the Dirac delta function. That is it is zero whenever k is not equal to l ($k \neq l$) and one when k is equal to l ($k = l$).

Unser et al. showed that the transfer function $H(z)$ of this filter could be readily calculated. An example for a third order β spline is shown in Eq. (6) below.

$$H(z) = (\beta_1^3(z))^{-1} = \frac{6}{z^{+1} + 4 + z^{-1}} \quad (6)$$

Although simple, implementing the transfer function $H(z)$ described by Eq. (6) is a problem because it has positive powers of z in it and is therefore unrealizable as it requires knowledge of the future!

Unser’s brilliant leap [17] was to realize that, for the basic β splines, the pre-filter can be simply realized by the application of two recursive (IIR) filters in cascade. One filter runs in the normal time direction and the other runs in the reverse time direction, as shown in Fig. 14 for our 3rd order β spline example.

The z -transform of these filters, for a third order spline, is given by:

$$\begin{aligned} (\beta_1^3(z))^{-1} &= \frac{6}{z^{+1} + 4 + z^{-1}} \\ &= 6 \left[\left(\frac{1}{1 - a_1 z^{-1}} \right)_{forward} \left(\frac{-a_1}{1 - a_1 z^{+1}} \right)_{reverse} \right] \end{aligned} \quad (7)$$

Where: $(\beta_1^3(z))^{-1}$ = the required β spline pre-filter
 z^{+1}, z^{-1} = integer sample time (T_s) advances, delays of the signal respectively
 a_1 = the pre-filter’s coefficient for $\beta^3 = (-2 + \sqrt{3}) \beta^3 = -0.26794919243$

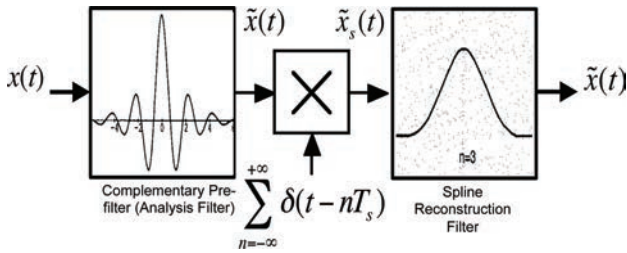


Fig. 15. Spline reconstruction with the complementary pre-filter in the time domain (impulse responses from [17, 20]).

Using this filter one can implement the necessary pre-filter, as shown in Fig. 15 for our 3rd order β -spline example.

Unfortunately, while the impulse response of the spline reconstruction filter is finite, the impulse response of the pre-filter is both acausal and infinite. As we have already seen the first problem can be removed by filtering forwards and backwards in time. This is a good solution for pictures or mastering, as the whole picture, or song, is already stored in memory. Thus running backwards over the samples is not an issue provided one is careful with the boundary conditions, due to the infinite nature of the impulse response, as discussed in [13, 20]. Providing these conditions are met it is possible to reconstruct the original signal without error, providing it was bandlimited in the first place. In practice the signal will need to be oversampled, as discussed later in Sec. 5.

For real time applications, however, the pre-filter would have to be implemented as a finite impulse response (FIR) filter. This would have to be windowed to limit the length, just like the brick-wall filter that the Shannon sampling requires. However, truncating this infinite impulse response also introduces errors (aliasing) into the reconstructed signal! So are we any better off?

The answer is yes, because the pre-filter’s impulse response’s amplitude is now falling off exponentially, rather than linearly, with time. Therefore, one can use a shorter filter for the same level of reconstruction error compared to a “brick wall” filter. Even when the effects of coefficient quantization are considered the pre-filter should perform better because the faster decay will minimize the effect of very small coefficient values.

4.3 The Frequency Responses of the Two Filters

The frequency responses of the pre-filter, the reconstruction filter, and their composite effect, for our example 3rd order β -spline are shown in Fig. 16.

The basic β -spline reconstruction filter has a very gentle cut-off but an ultimate roll-off rate of $(n + 1) \times -20\text{dB/decade}$, where n is the order of the spline ($n = 3$ in the Fig. 16). However, it does have the advantage of not having any overshoot in the time domain, which would make the design of the post D/A analog circuitry easier. The pre-filter on the other hand has a peak at the cut-off frequency and a slower roll-off rate of $(n - 1) \times -20\text{dB/decade}$. This peak could cause some issues with overload margins and consequential dynamic range loss in analog implementa-

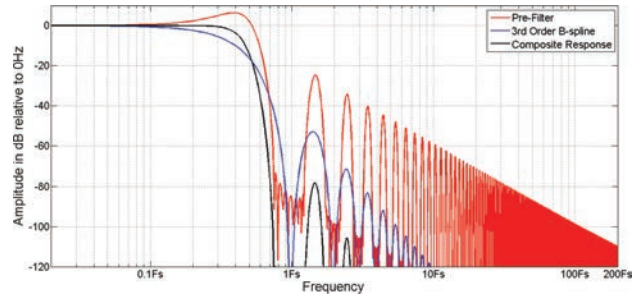


Fig. 16. The frequency responses of the pre and post filters for β spline sampling for $n = 3$.

tions. Fortunately, there is usually little signal energy at these frequencies in audio signals.

The composite response is much better, as the two filters are complementary and the resulting response is flat in the passband and has a roll-off that is much faster. It should be noted that the magnitude response does not tell the whole story, because the time responses of these filters are bi-orthogonal and thus also have joint properties that are hidden by looking at just the magnitude response.

4.3.1 Other Spline Orders

This tutorial has focused on 3rd order splines, both for convenience and because it is the order most used in video processing. Audio, on the other hand is, in many respects, a much more challenging signal with regards to the perception of error. There is no restriction, in principle, to having higher orders. Aldroubi et al. [20] have proved that the method is extendable to all orders up to infinity, at which point it converges to the ideal Shannon case.

Higher order spline *kernels* will still have the advantage of reduced error for finite length pre-filters and a reduced requirement on the amount of oversampling required for a given level of error.

4.3.2. Other Spline Kernels

Providing the bi-orthogonal, and partition of unity, conditions are met there are other filter pairs, or *kernels*, that can be used for modern sampling schemes.

Of these the *cardinal spline* pair may be the one with the most utility. Like the basic β spline filter, these filters are compact in time and approximately twice as long. Their matching pre-filters also have infinitely long impulse responses that decay exponentially in time and so can be implemented as shorter filters compared to the brick-wall lowpass filter for a given level of error. Cardinal splines have the advantage that they share the same property that the brick-wall Shannon lowpass filters have of being zero valued at every other sample to the current one. This type of spline has application in the interpolation of sampled signals and so may be useful for sample rate change processors.

It should be noted that the spline functions are only one of a myriad of bi-orthogonal *kernels* that could be used. There may be others that are superior for audio applications.

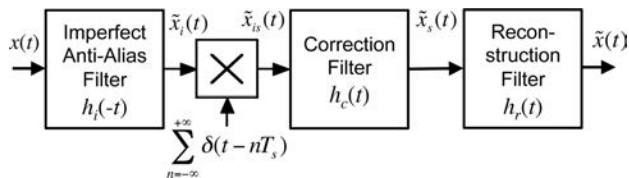


Fig. 17. Sampling for non-ideal analog acquisition devices.

4.4 Compensating for Imperfections in Real Filters

When converting between the analog and digital domains it may not be possible to exactly replicate the necessary pre-filter and post filter impulse responses. In these cases Unser [18] proposes that a digital “correction” filter as shown in Fig. 17 could compensate for imperfections within the analog domain.

The correction filter is designed to enforce the necessary bi-orthogonality relationship between the analog pre-filter and reconstruction filter; more details can be found in [18].

5 BUT WHAT ABOUT ALIASING?

In Shannon’s original paper [1] he showed that the information rate in an analog signal was no more than twice the highest frequency. In his paper he was primarily referring to baseband signals that extended from 0 Hz up to some maximum frequency B Hz.

In 2002 Vetterli et al. [22] introduced a new class of signals called signals with *finite rate of innovation* (FRI) that includes both bandlimited signals and non-bandlimited signals. Instead of characterizing a signal by its frequency content, FRI theory describes it in terms of the *innovation* parameters per unit of time. Bandlimited signals are therefore a subset of this more general definition. In general very sparse signals may require sophisticated sampling and post processing schemes in order to reconstruct them and they are outside the scope of this tutorial; for a good easy to read overview see [23]. However, as discussed in [24], audio signals are naturally bandlimited and so can be examined using this theory. The *innovation* of a signal is a function of its information content and can be considered the “surprise value” of the continuous signal. Again, this links back to Shannon who suggested the concept of change representing information. For the case of bandlimited signals with a bandlimit of B Hz the rate of innovation turns out to be $2B$ Hz, in line with Shannon’s theory and with the time domain explanation outlined previously in Sec. 3. In Sec. 3 we showed that that white noise bandlimited by a brickwall low-pass filter could be sampled at two times the bandwidth for perfect reconstruction, corresponding to its finite rate of innovation. Unfortunately audio signals are not sparse in the time domain that precludes the direct time domain application of sparse sampling systems to the audio signals. However they are sparse in the frequency domain, a fact that is exploited by audio coding systems to reduce the bit rate. The techniques described in this tutorial are applicable to high resolution audio because the short impulse responses of the spline based reconstruction filters may offer some temporal advantages.

In modern sampling schemes the (anti-alias) pre-filter and the reconstruction filter are designed to be bi-orthogonal. This means that the pre-filter puts constraints on the rates of change and higher order derivatives of the continuous signal, thus limiting its rate of innovation. This, combined with the smoothing provided by the reconstruction filter, results in the rejection of the possible alias components in the output due to the bi-orthogonal relationship between the two filters.

Note, it is vital that the filters are designed together, as discussed in Hummel [25]; failure to do this results in inferior reconstruction. But so long as the pre-filter and the reconstruction filter are matched together then aliasing should not be an issue because there is no, or minimal, error in the reconstruction.

Notwithstanding this, it is evident that the sampling frequency required will be greater than the usual audio rates. Unser [18] presents graphs that suggest that 3 to 4 times oversampling might be needed, depending on the order of the spline and the level of residual error required. However he was primarily interested in video signals with high bandwidths. For audio a sampling rate that is high enough to capture the naturally band limited spectrum of approximately 90 kHz, as discussed in [24], should be sufficient.

6 MODERN VERSUS TRADITIONAL SHANNON SAMPLING

The pros and cons of modern sampling methods compared to traditional Shannon sampling are as follows.

6.1 Pros

Modern sampling has reconstruction filters that have short impulse responses whereas traditional sampling requires infinitely long impulse responses for perfect reconstruction.

For real time use modern sampling has a lower reconstruction error for a given anti-alias filter impulse response length when compared to traditional sampling methods.

6.2 Cons

Perfect reconstruction is only achievable when the whole piece of music is already available for modern sampling methods. On the other hand traditional sampling requires infinitely long impulse responses to achieve this.

7 CONCLUSION

This tutorial has examined both traditional Shannon and modern approaches to sampling. We have shown that, despite its perfection in theory, a Shannon sampled signal in practice will always have some aliasing errors when reconstructed. We have described alternative approaches based on *splines* that offer the possibility of a lower level of error for a given filter length as well as reconstruction filters that are very compact in time. The challenge for the future will be to develop circuits that allow us to achieve these advantages in practice. Thus allowing us to capture

precious audio heritage and performances in better, or even perfect, quality in the future.

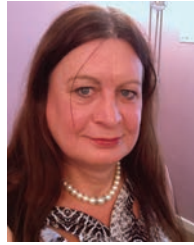
8 ACKNOWLEDGMENTS

The author wishes to thank the reviewers whose detailed comments greatly improved this tutorial. This work was also partially supported by JASA Consultancy.

9 REFERENCES

- [1] C. E. Shannon, "Classic Paper: Communication in the Presence of Noise," *Proc. IEEE*, vol. 86, no. 2, pp. 447–457 (1998), <https://doi.org/10.1109/JPROC.1998.659497>.
- [2] H. Nyquist, "Certain Topics in Telegraph Transmission Theory," *Trans. Amer. Inst. Elect. Eng.*, vol. 47, pp. 617–644 (1928), <https://doi.org/10.1109/T-AIEE.1928.5055024>.
- [3] E. T. Whittaker, "On the Functions which Are Represented by the Expansion of Interpolating Theory," *Proc. R. Soc. Edinburgh*, vol. 35, pp. 181–194 (1915), reprinted in *Proceedings of the IEEE*, vol. 88, no. 4 (2000 Apr.).
- [4] V. A. Kotel'nikov, "On the Transmission Capacity of 'Ether' and Wire in Electrocommunications," *Izd. Red. Upr. Svyazzi RKKA (Moscow)* (1933).
- [5] J. R. Stuart and P. Craven, "A Hierarchical Approach for Audio Capture, Archive and Distribution," *J. Audio Eng. Soc.*, vol. 67, no. 5, (2019 May) DOI: 10.17743/jaes.2018.0062 (this Issue).
- [6] J. D. Reiss, "A Meta-Analysis of High Resolution Audio Perceptual Evaluation," *J. Audio Eng. Soc.*, vol. 64, pp. 364–379 (2016 Jun.), <https://doi.org/10.17743/jaes.2016.0015>.
- [7] F. J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83 (1978 Feb.), <https://doi.org/10.1109/PROC.1978.10837>
- [8] J. Oñativia and P. L. Dragotti, "Sparse Sampling: Theory, Methods and an Application in Neuroscience," *Biological Cybernetics*, vol. 109, no. 1, pp. 125–139 (2015 Feb.). <https://doi.org/10.1007/s00422-014-0639-x>.
- [9] "Partition of Unity," Wikipedia, https://en.wikipedia.org/wiki/Partition_of_unity.
- [10] S. H. Linkwitz "Active Crossover Networks for Non-coincident Drivers," *J. Audio Eng. Soc.*, vol. 24, pp. 2–8 (1976 Jan./Feb.).
- [11] A. Croisier, D. Esteban, and C. Galand, "Perfect Channel Splitting by Use of Interpolation/Decimation/Tree Decomposition Techniques," *Proceedings of the International Symposium on Information Circuits and Systems*, Patras, Greece (1976).
- [12] S. K. Agrawal and O. P. Sahu, "Review Article—Two-Channel Quadrature Mirror Filter Bank: An Overview," *SRN Signal Processing*, vol. 2013 (2013), Article ID 815619, <https://doi.org/10.1155/2013/815619>.
- [13] M. Unser, A. Aldroubi, and M. Eden, "Fast B-spline Transforms for Continuous Image Representation and Interpolation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 277–285 (1991 Mar.), <https://doi.org/10.1109/34.75515>.
- [14] M. Unser, A. Aldroubi, and M. Eden, "Polynomial Spline Signal Approximations: Filter Design and Asymptotic Equivalence with Shannon's Sampling Theorem," *IEEE Trans. Inform. Theory*, vol. 38, pp. 95–103 (1992 Jan.), <https://doi.org/10.1109/18.108253>.
- [15] M. Unser, A. Aldroubi, and M. Eden, "B-Spline Signal Processing: Part I—Theory," *IEEE Trans. Signal Processing*, vol. 41, pp. 821–833 (1993 Feb.), <https://doi.org/10.1109/78.193220>.
- [16] M. Unser, A. Aldroubi, and M. Eden, "B-Spline Signal Processing: Part II—Efficient Design and Applications," *IEEE Trans. Signal Processing*, vol. 41, pp. 834–848 (1993 Feb.), <https://doi.org/10.1109/78.193221>.
- [17] M. Unser, "Splines: A Perfect Fit for Signal and Image Processing," *IEEE Signal Processing Mag.*, vol. 16, no. 6, pp. 22–38 (1999), <https://doi.org/10.1109/79.799930>.
- [18] M. Unser, "Sampling—50 Years after Shannon," *Proc. IEEE*, vol. 88, pp. 569–587 (2000 Apr.), <http://www.epfl.ch/publications/unser0001.pdf>.
- [19] M. Unser, A. Aldroubi, and M. Eden, "The L_2 Polynomial Spline Pyramid," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 364–379 (1993 Apr.), <https://doi.org/10.1109/34.206956>.
- [20] A. Aldroubi, M. Unser, and M. Eden, "Cardinal Spline Filters: Stability and Convergence to the Ideal Sinc Interpolator," *Signal Process.*, vol. 28, no. 2, pp. 127–138 (1992), [https://doi.org/10.1016/0165-1684\(92\)90030-Z](https://doi.org/10.1016/0165-1684(92)90030-Z).
- [21] "Spline Interpolation," Wikipedia, https://en.wikipedia.org/wiki/Spline_interpolation.
- [22] M. Vetterli, P. Marziliano, and T. Blu, "Sampling Signals with Finite Rate of Innovation," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428 (2002), <https://doi.org/10.1109/TSP.2002.1003065>.
- [23] D. Mackenzie, "Compressed Sensing Makes Every Pixel Count," in *What's Happening in the Mathematical Sciences*, vol. 7, pp. 114–127 (The American Mathematical Society). <https://www.ams.org/publicoutreach/math-history/hap7-pixel.pdf>.
- [24] J. R. Stuart and P. Craven, "A Hierarchical Approach to Archiving and Distribution," presented at the 137th Convention of the Audio Engineering Society (2014 Oct.), convention paper 9178.
- [25] R. Hummel, "Sampling for Spline Reconstruction," *SIAM J. Appl. Math.*, vol. 43, pp. 278–288 (1983), <https://doi.org/10.1137/0143019>.

THE AUTHOR



Jamie A. S. Angus

Jamie Andrea Shyla Angus-Whiteoak is Professor of Audio Technology at Salford University. Her interest in audio was crystallized at age 11 when she visited the WOR studios in NYC on a school trip in 1967. After this she was hooked, and spent much of her free time studying audio, radio, synthesizers, and loudspeakers, and even managed to build some! After secondary education in Scotland, at 17 she attended the University of Lethbridge in Alberta Canada. There, in addition to her studies in physics, music, computing, drama, philosophy, and English composition, she repaired their VCS3 synthesizer, and so obtained coveted access to the electronic music lab. She studied electronics at Kent (UK) doing her BSc and PhD there from 1974 to 1980. During her PhD study she became interested in A/D conversion, and worked on a sigma-delta approach, but had to give it up to concentrate on her Thesis topic of designing a general purpose Digital Signal Processor. After her PhD she joined Standard Telecommunications Laboratories, who invented optical fibres and PCM. There she worked on integrated optics, speech coding, speech synthesis, and recognition in the early 80s, and invented a novel 32kBits speech coding method. She has been ac-

tive in audio and acoustic research from then. She was appointed as the BT Lecturer at the University of York in 1983, to develop the first integrated masters (MEng) in Electronic and Communication Engineering in conjunction with British Telecom. She then co-created the UK's first Music Technology course in 1986 when it was considered a "silly idea"! She is the inventor of; modulated, wide-band, and absorbing diffusers, direct processing of Super Audio CD signals, and one of the first 4-channel digital tape recorders. She has done work on signal processing, analogue circuits, and numerous other audio technology topics.

She teaches audio and video signal processing, psychoacoustics, sound reproduction, studio design, and audio and video coding. She has co-written two textbooks and has authored or co-authored over 200 journal and conference papers and 4 patents. She is currently investigating environmentally friendly audio technology. She has been awarded an AES fellowship, the IOA Peter Barnet Memorial Award, and the AES Silver Medal Award, for her contributions to audio and acoustics. For relaxation she likes playing drums and dancing, but not at the same time.