The Sampling Theorem for Bandlimited Functions

Intuitively one feels that if one were to sample a continuous-time signal such as a speech waveform "sufficiently fast" then the sequence of sample values will form an "adequate" representation of the waveform. What constitutes a "sufficiently fast" sampling rate and, likewise, what would constitute an "adequate representation?" Let us take these questions in reverse order.

Suppose $x(t)$ is a continuous-time waveform. Sampling the waveform $x(t)$ at regular intervals at a fixed, uniform rate of $f_0 = 1/T$ samples per second\footnote{We're reusing notation here to keep the burgeoning notation within sane bounds; don't confuse the sampling interval $T$ (and the corresponding sampling frequency $f_0$) with the same notation we had used for the fundamental period $T$ of a sinusoid of frequency $f_0 = 1/T$.} yields the sequence of sample values

$$\ldots, x(-2T), x(-T), x(0), x(T), x(2T), \ldots.$$  

Write $x_n = x(nT)$ for the sequence of sample values with $n$ taking all integral values, $n = 0, \pm 1, \pm 2, \ldots$. It is trite and clear that the waveform $x(t)$ completely determines the sample sequence $\{x_n\}$—given $x(t)$ we just “read off” the values of $x(t)$ at the sampling instants $t = 0, \pm T, \pm 2T, \ldots$. Our problem, however, is the obverse—given the sample sequence $\{x_n\}$ can we unambiguously determine the entire waveform $x(t)$ (or at least a good facsimile thereof)? A mathematician would recognise this as a problem in interpolation. We would like to find a smooth interpolation through the sample points $x_n = x(nT)$ which mimics the original waveform $x(t)$. If we think of the graph of the function $x(t)$, the question is whether we can “connect the dots” corresponding to the sample values $x_n$ in a smooth fashion so as to mirror $x(t)$. Of course, there is such a way of connecting the dots: the function $x(t)$ itself! The problem is, however, that we don't know $x(t)$ ahead of time; all knowledge about $x(t)$ is encapsulated in the sequence of sampled values $x_n = x(nT)$. Thus, what we would like is some fixed way of interpolating between the sample points $x_n$ which works not only for one waveform $x(t)$ but for a whole family of waveforms.

There are really two issues here. The first deals with determining a fixed interpolation procedure that works; the second deals with determining the requisite sampling rate $f_0 = 1/T$. As we shall see shortly, Fourier theory gives us both answers in one fell and almost (at least in retrospect) transparently simple swoop.

If a fixed interpolation procedure connecting the sampled values $x_n = x(nT)$ is to work at all, intuition suggests that the originating waveform $x(t)$ cannot change drastically between successive sampling instants. It seems clear that if $x(t)$ could have an arbitrary number of wild oscillations between sampling instants then no fixed interpolation procedure can work.
This hence appears to place a restriction on not just the smoothness (or continuity) of the waveform $x(t)$, but also on its maximum allowable rate of change, or, to put it in another way, its maximum frequency content. Carrying on in this boldly intuitive vein, this suggests that the sampling frequency $f_0 = 1/T$ be at least as high as the highest frequency encountered in the waveform $x(t)$. Let us formalise this intuition.

Recall that if a signal $x(t)$ is periodic with period $T$ then it is completely represented by a discrete sequence of Fourier coefficients $a_n = a(n/T)$ corresponding to a discrete spectrum at regularly spaced intervals of width $f_0 = 1/T$ in frequency. Indeed, the representations $x(t) \leftrightarrow a_n$ are equivalent and one may move from one representation to the other at will via the Fourier formulæ:

$$x(t) = \sum_{n=-\infty}^{\infty} a_n e^{j2\pi(n/T)t} \quad \leftrightarrow \quad a_n = \frac{1}{T} \int_{-T/2}^{T/2} x(t) e^{-j2\pi(n/T)t} dt. \quad (*)$$

If $x(t)$ is aperiodic and of finite energy then its decomposition in terms of its frequency content will require all possible frequencies $f$ and not just a discrete sequence of harmonically related frequencies $nf_0$. The Fourier sum in $(*)$ is then replaced by a Fourier integral where the spectral content of the waveform $x(t)$ at each frequency $f$ is exactly the spectrum (or Fourier transform) $X(f)$. Indeed, the two representations $x(t) \leftrightarrow X(f)$ are equivalent and one may be obtained from the other via the dual relationship

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} df \quad \leftrightarrow \quad X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt. \quad (***)$$

We say that an aperiodic waveform $x(t)$ is (strictly) bandlimited if $X(f) = 0$ for $|f| \geq B$; the quantity $B$ is called the bandwidth of the waveform and may be identified with the largest frequency encountered in the waveform. The intuitive argument then suggests that the requisite sampling frequency for a waveform with bandwidth $B$ be at least as high as $2B$.\(^2\)

There are many ways to prove this result but the simplest and most direct may be to assert that we have already seen this result with the roles of time and frequency interchanged in $(*)$. Indeed, suppose $x(t)$ has bandwidth $B$ so that $X(f) = 0$ for $|f| \geq B$. Let us construct a periodic waveform $\hat{X}(f)$ in frequency whose period is $2B$ and which coincides with $X(f)$ over $|f| \leq B$. In

\(^2\) Yet another reuse of the notation!

\(^3\) As a nod to historical convention, the bandwidth $B$ refers only to the positive half of the frequency axis; recall that there is, likewise, a band of width $B$ in the negative half of the frequency axis so that the total band of frequencies over which the spectrum can take nonzero values is $2B$. \(^3\)
other words,
\[ \hat{X}(f + 2B) = \hat{X}(f) \quad \text{(for all } f), \]
\[ \hat{X}(f) = X(f) \quad (|f| \leq B). \]  

But we’ve now constructed a periodic function \( \hat{X}(f) \) of the argument \( f \) whose period is \( 2B \). This function hence has a Fourier series representation in terms of complex exponentials of the form \( e^{j2\pi n/2B} \) with Fourier coefficients, say \( \hat{x}_n \). Indeed, if we formally interchange the roles of \( f \) and \( t \) in \((\ast)\) and replace the period \( T \) by \( 2B \), we obtain\(^4\)

\[ \hat{X}(f) = \sum_{n=-\infty}^{\infty} \hat{x}_n e^{-j2\pi n/2B} f \leftrightarrow \hat{x}_n = \frac{1}{2B} \int_{-B}^{B} \hat{X}(f) e^{j2\pi n/2B} f df. \]  

Invoking duality, we may identify the Fourier coefficients in the expansion \((\ast')\) with a discrete sequence \( \hat{x}_n = \hat{x}(n/2B) \) of regularly spaced points in time where the interval between successive points is exactly \( 1/2B \) seconds.

Let’s take stock of what we’ve done. Starting with the bandlimited function \( x(t) \) we can determine its spectrum \( X(f) \) via \((\ast\ast)\) which, in turn, determines the periodic function \( \hat{X}(f) \) via \((\ast\ast')\) and which finally determines the sequence of Fourier coefficients \( \hat{x}_n \) via \((\ast')\). What is remarkable is that, starting with the discrete time sequence \( \hat{x}_n = \hat{x}(n/2B) \), we can reverse our path by first determining the periodic function \( \hat{X}(f) \) through \((\ast')\), then identifying \( X(f) \) with one period of \( \hat{X}(f) \) through \((\ast\ast')\), and finally extracting \( x(t) \) from its spectrum \( X(f) \) via the Fourier relationship \((\ast\ast)\). Thus, the representations \( x(t), X(f), \hat{X}(f), \) and \( \hat{x}_n \) are all completely equivalent from an information-theoretic point of view:

\[ \hat{x}_n \leftrightarrow \hat{X}(f) \leftrightarrow X(f) \leftrightarrow x(t). \]

We’ve thus obtained a remarkable fact.

**Slogan.** Any bandlimited function \( x(t) \) is completely specified by a discrete sequence of values \( \hat{x}_n \).

To complete the epiphany, we would like to express this symbiosis in a mathematically simple form. Let’s first examine how the sequence \( \hat{x}_n \) is related to the waveform \( x(t) \). The result follows by nothing more than meandering through \((\ast'), (\ast\ast'), (\ast\ast)\) in sequence. We commit the mathematical solemnism of writing down a long sequence of equations pleading the elegance for consistency, we’ve kept \(-j\) in the exponent for the frequency-domain expansion with \(+j\) reserved for the time-domain expansion so that the Fourier series itself has a reversed sign in the exponent. This is simply a matter of definition and nothing fundamental is affected as long as we are consistent.
and transparency of the steps in extenuation. We have

\[ \hat{x}_n \overset{(a)}{=} \frac{1}{2B} \int_{-B}^{B} \hat{X}(f)e^{j2\pi fn/2B} \, df \overset{(b)}{=} \frac{1}{2B} \int_{-B}^{B} X(f)e^{j2\pi fn/2B} \, df \]

\[ \overset{(c)}{=} \frac{1}{2B} \int_{-\infty}^{\infty} X(f)e^{j2\pi fn/2B} \, df \overset{(d)}{=} \frac{1}{2B} x(n/2B), \]

where (a) follows by definition from \((\ast)\), (b) from \((\ast\ast)\) as \(\hat{X}(f)\) coincides with \(X(f)\) over the region of integration \(-B \leq f \leq B\), again by definition, (c) follows as \(X(f)\) is identically zero for \(|f| \geq B\) so that we may extend the region of integration from \(-B \leq f \leq B\) to \(-\infty < f < \infty\), and (d) follows from \((\ast\ast)\) by identifying the integral as simply the value of \(x(t)\) evaluated at \(t = n/2B\). Thus, remarkably, the sequence values \(\hat{x}_n\) are simply the sampled values \(x_n = x(n/2B)\) scaled by the constant value \(2B\). We thus have the following truly remarkable improvement to the previous slogan.

**Slogan, refined.** Any bandlimited function \(x(t)\) with bandwidth \(B\) is completely specified by its sampled values \(x_n = x(n/2B)\).

We've thus answered one of our two questions. It suffices to sample the bandlimited waveform every \(T = 1/2B\) seconds to generate a sequence of sampled values \(x_n = x(n/2B)\) which contains all the information of the original waveform in the sense that \(x(t)\) can be exactly reconstructed from the sequence \(x_n\). This critical sampling rate of \(2B\) samples per second is called the Nyquist rate.

Having determined the requisite sampling rate, can we also explicate the underlying fixed process (independent of the choice of bandlimited waveform \(x(t)\)) which will enable us to interpolate properly through the sampled values \(x_n = x(n/2B)\) to obtain the originating function \(x(t)\)? All we have to do is meander through the relations \((\ast\ast, \ast\ast', \ast)\) (in reverse order this time). The gentle reader will find it instructive to fill in the reasoning for each of the steps below:

\[ x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft} \, df = \int_{-B}^{B} X(f)e^{j2\pi ft} \, df = \int_{-B}^{B} \hat{X}(f)e^{j2\pi ft} \, df \]

\[ = \int_{-B}^{B} \left( \sum_{n=-\infty}^{\infty} \hat{x}_n e^{-j2\pi (n/2B) f} \right) e^{j2\pi ft} \, df = \sum_{n=-\infty}^{\infty} \hat{x}_n \int_{-B}^{B} e^{j2\pi f(t-n/2B)} \, df \]

\[ = \sum_{n=-\infty}^{\infty} \frac{x(n/2B)}{2B} \int_{-B}^{B} e^{j2\pi f(t-n/2B)} \, df. \]

The integral inside the sum is elementary: when \(t = n/2B\),

\[ \int_{-B}^{B} e^{j2\pi f(t-n/2B)} \, df = \int_{-B}^{B} df = 2B \quad (t = n/2B), \]
while in all other cases

$$\int_{-B}^{B} e^{j2\pi f (t-n/B)} \, df = \left( \frac{e^{j2\pi f (t-n/2B)} - e^{-j2\pi f (t-n/2B)}}{j2\pi(t-n/2B)} \right)_{-B}^{B} = 2B \frac{\sin(2\pi B(t-n/2B))}{2\pi B(t-n/2B)} \quad (t \neq n/2B).$$

Recalling that the sinc($\cdot$) function is defined as

$$\text{sinc } \theta = \begin{cases} 
1 & \text{ if } \theta = 0, \\
\frac{\sin(\pi \theta)}{\pi \theta} & \text{ if } \theta \neq 0,
\end{cases}$$

we may combine both cases and write

$$\int_{-B}^{B} e^{j2\pi f (t-n/2B)} \, df = 2B \text{sinc}(2B(t-n/2B)).$$

Putting everything together, we have the following explicit and remarkable interpolation formula:

**Slogan, final form.** Any strictly bandlimited waveform $x(t)$ of bandwidth $B$ can be explicitly reconstructed from its values sampled uniformly every $T = 1/2B$ seconds via the interpolation formula

$$x(t) = \sum_{n=-\infty}^{\infty} x(n/2B) \text{sinc}(2B(t-n/2B)).$$

This is the famous Nyquist-Shannon sampling theorem.

The function sinc$2Bt$ whose time-shifted variants sinc$2B(t-n/2B)$ appear so ubiquitously throughout the formula is called the interpolation function. Observe that it is fixed and independent of the choice of the particular bandlimited function $x(t)$. The sampling theorem hence gives us a fixed procedure to “connect the dots” for any bandlimited function, i.e., provides a fixed interpolation formula which enables us to go from the sampled “dots” $x(n/2B)$ to the complete bandlimited waveform $x(t)$ in an exact, unambiguous way. Truly a rather high order of magic!