Abstract

This note reviews the sampling theorem and the difficulties involved in teaching it.

1 Introduction

The sampling theorem—a bandlimited continuous-time signal can be reconstructed from its sample values provided the sampling rate is greater than twice the highest frequency in the signal—is taught in every Signals and Systems course because of its fundamental nature. This note reviews two possible ways of teaching it, and discusses the pros and cons of these ways.

The notation is standard: \( x_c(t) \) is a continuous-time real-valued signal, where \( t \) ranges over the whole infinite time interval \((-\infty, \infty)\); \( X_c(j\omega) \) is the Fourier transform of \( x_c(t) \), where \( \omega \) ranges over \((-\infty, \infty)\); \( T \) is the sampling period in seconds, \( f_s = 1/T \) the sampling frequency in Hz, \( \omega_s = 2\pi f_s \) the sampling frequency in rad/s, and \( \omega_N = \omega_s/2 \) the Nyquist frequency in rad/s; \( x_c(nT) \) are the sample values, where \( n \) runs over all integers, positive, negative, and 0; \( x_c(nT) \) is denoted by \( x[n] \); and \( X(e^{j\omega}) \) is the discrete-time Fourier transform of \( x[n] \), where here \( \omega \) ranges over \((-\pi, \pi)\).

2 Aliasing

One common way to introduce the sampling theorem is to show a graph of a sinusoid and show that if you undersample, there’s a lower-frequency sinusoid that interpolates the same sample points. In the figure just below, the signal \( \sin(1.2\pi t) \) of frequency 0.6 Hz is sampled at only 1 Hz (instead of greater than 1.2 Hz), and the “aliased signal” \( \sin(-0.8\pi t) \) goes through the same sample values, shown as black dots:
This gives an intuitive idea of the concept of aliasing. There being no aliasing is a sufficient condition for being able to reconstruct the signal. However, there are other sufficient conditions, so this example doesn’t tell the whole story.

## 3 The Sampling Problem is Linear

The problem of reconstructing a signal from samples can be posed as a linear mathematical problem: Sampling is a linear projection of the continuous-time signal; the bandwidth constraint is a subspace condition. Possible lecture notes follow.

1. Suppose voltage \( v \) and current \( i \) are involved in an electric circuit, and suppose you measure the value of \( v \). Can you say what is the value of \( i \)? Not unless you know how \( v \) and \( i \) are related, such as for example \( v = R i \) where you know \( R \).

   Let’s think of this geometrically: \( x = (v, i) \) is a point in the plane. If you know that \( v = R i \) and you know \( R \), then you know that \( x \) lies on a specific line \( \mathcal{L} \) through the origin. Measuring \( v \) is a form of sampling the vector \( x \)—projecting it onto the horizontal axis. Then knowing \( x \in \mathcal{L} \) and measuring the horizontal component \( v \) of \( x \) allows you to determine \( x \) uniquely, i.e., find the vertical component \( i \).

2. Here’s a somewhat larger example. The “signal” is a vector \( x \) in \( \mathbb{R}^6 \):

   \[
x = (x_0, x_1, x_2, x_3, x_4, x_5).
\]

   Suppose you know that the signal lives in the subspace \( B \) defined by the three equations

   \[
   \begin{align*}
x_1 + 2x_2 + x_4 &= 0 \\
-2x_0 - x_1 + x_5 &= 0 \\
x_0 - 3x_4 + x_5 &= 0.
\end{align*}
   \]
You sample the signal in this way: You measure that
\[ x_0 = 1, \quad x_2 = -2, \quad x_4 = 3. \]

You want to find the missing components of \( x \). It’s easy, because you now have 6 linear equations for the 6 variables. As long as the equations are linearly independent, they uniquely determine the 6 variables. Let us write the 6 equations in the form \( Ax = b \):

\[
\begin{bmatrix}
0 & 1 & 2 & 0 & 1 & 0 \\
-2 & -1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & -3 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
x_0 \\
x_1 \\
x_2 \\
x_3 \\
x_4 \\
x_5
\end{bmatrix}
=
\begin{bmatrix}
0 \\
0 \\
0 \\
1 \\
-2 \\
3
\end{bmatrix}.
\]

You can solve by Gaussian elimination, or \( x = A^{-1}b \), or however you like.

3. Let us rephrase the problem in a slightly different way. The given subspace condition is \( Bx = 0 \), where \( B \) is the matrix

\[
B = \begin{bmatrix}
0 & 1 & 2 & 0 & 1 & 0 \\
-2 & -1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & -3 & 1
\end{bmatrix}.
\]

You measure \( y = Dx \), where

\[
y = \begin{bmatrix}
1 \\
-2 \\
3
\end{bmatrix}, \quad D = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}.
\]

The problem is this: Given \( B, D, y \) and the equations \( Bx = 0, \: Dx = y \), compute \( x \). The equation to solve is

\[
\begin{bmatrix}
B \\
D
\end{bmatrix} \begin{bmatrix}
x \\
y
\end{bmatrix} = \begin{bmatrix}
0 \\
y
\end{bmatrix},
\]

where the zero on the right-hand side is the \( 3 \times 1 \) zero vector.

4. For completeness, the definition of a \textbf{subspace \( \mathcal{B} \)} is a nonempty subset of a vector space that has the properties 1) if \( x, y \) are in \( \mathcal{B} \), so is \( x + y \), 2) if \( x \) is in \( \mathcal{B} \) and \( c \) is a real number, then \( cx \) is in \( \mathcal{B} \). The subspaces of \( \mathbb{R}^3 \) are the lines and planes through the origin, together with the origin and finally \( \mathbb{R}^3 \) itself. In the preceding example, the subspace \( \mathcal{B} \) is the nullspace of \( B \). (Nullspace is the solution space of the homogeneous equation \( Bx = 0 \).)
4 The Sampling Theorem in Discrete Time

At first glance, the task in the sampling theorem seems hopeless: Interpolating a continuum of values from only countably many? The discrete-time version seems less daunting: Given every other sample, compute the skipped ones.

The following note assumes students know the discrete-time Fourier transform and the $z$ transform.

1. The signal $x[n]$ is defined for all $-\infty < n < \infty$. It is downsampled by 2, $y[n] = x[2n]$:

```
+--------+             +--------+
|        |             |        |
|        |             |        |
|        |             |        |
|        |             |        |
|        |             |        |
|        |             |        |

\[
x[n] \downarrow 2 \quad y[n]
\]
```

The components $y[n]$ are measured and the task is to compute all the sample values of $x[n]$ that weren’t measured. The block diagram model is

```
x[n] \downarrow 2 \quad y[n]
```

where an arrow made of dashes indicates a discrete-time signal. We want to find the input given the output, i.e., we want another system with input $y$ and output $x$:

```
x[n] \downarrow 2 \quad y[n] \quad x[n]
```

So again we’re given a projection of $x[n]$, namely, all the components with even indices. For the problem to be solvable, $x[n]$ must be restricted in some way. It turns out that a sufficient restriction is that $x[n]$ lies in the subspace $B_{\pi/2}$ of signals
bandlimited to frequency \( \pi/2 \), that is, \( X(e^{j\omega}) = 0 \), \( \pi \geq |\omega| \geq \pi/2 \). This means intuitively that \( x[n] \) isn’t changing too quickly.

To repeat, the sampling problem is this: We measure \( y[n] = x[2n] \) for all \( n \) and we want to compute \( x[n] \) for all odd \( n \), with the knowledge that \( x \in B_{\pi/2} \).

2. We turn to the frequency domain. We have the \( z \)-transform

\[
X(z) = \cdots + x[-1]z + x[0] + x[1] \frac{1}{z} + \cdots
\]

and therefore

\[
X(-z) = \cdots - x[-1]z + x[0] - x[1] \frac{1}{z} + \cdots.
\]

Adding gives

\[
X(z) + X(-z) = 2Y(z^2).
\]

From this it follows that

\[
Y(z^2) = \frac{1}{2}[X(z) + X(-z)], \quad Y(e^{j2\omega}) = \frac{1}{2}[X(e^{j\omega}) + X(e^{j(\omega-\pi)})]
\]

Let us symbolically sketch the Fourier transform of \( x[n] \) as a triangle:

\[
\begin{array}{c}
X(e^{j\omega}) \\
\hline
-\pi & \pi \\
\end{array}
\]

Note that this plot isn’t meant to be realistic: \( X(e^{j\omega}) \) is complex-valued, for one thing. The plot merely shows the support of \( X(e^{j\omega}) \), i.e., the frequency range where \( X(e^{j\omega}) \) is nonzero. Remember: \( X(e^{j\omega}) \) is a periodic function of \( \omega \), of period \( 2\pi \).

Thus the graph is duplicated periodically to the left and right. Then the graph of \( X(e^{j(\omega-\pi)}) \) is that of \( X(e^{j\omega}) \) shifted to the right by \( \pi \):
Adding the two and dividing by 2 gives $Y(e^{j2\omega})$:

$$Y(e^{j2\omega})$$

Therefore the FT $Y(e^{j\omega})$ looks like

$$Y(e^{j\omega})$$

Thus in the frequency domain the effect of the downsampler is to stretch out the Fourier transform of $X(e^{j\omega})$ and scale by 1/2, but there is no aliasing. That is, the triangle of $X$ is preserved uncorrupted in $Y$.

3. The following system reconstructs $x[n]$ from $y[n]$: 

<table>
<thead>
<tr>
<th>$y[n]$</th>
<th>$\uparrow 2$</th>
<th>$w[n]$</th>
<th>$H(e^{j\omega})$</th>
<th>$x_r[n]$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The output is denoted $x_r[n]$ (r for reconstructed). It will turn out that $x_r = x$ if $x \in B_{\pi/2}$. Let’s follow the Fourier transforms from $y$ to $x_r$. We have

$$w[2n] = y[n], w[2n + 1] = 0, \quad W(z) = Y(z^2), \quad W(e^{j\omega}) = Y(e^{j2\omega}).$$

Thus the graph of $W(e^{j\omega})$ is
The ideal lowpass filter passes only the low frequency triangle, amplifying by 2:

\[ X_r(e^{j\omega}) \]

\[ \begin{array}{c}
-\pi \\
\downarrow \\
1 \\
\uparrow \\
\pi 
\end{array} \]

Thus we have shown that \( x_r = x \) if \( x \in B_{\pi/2} \).

4. Finally, we need to see how to implement the reconstruction, that is, the time-domain formula for \( x_r[n] \) as a function of \( y[n] \). Let \( h[n] \) denote the impulse response function of the filter:

\[ h[n] = \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} 2e^{j\omega n} d\omega = \frac{\sin(\frac{\pi}{2}n)}{\frac{\pi}{2}n}. \]

This is the familiar sinc function. The convolution equation is

\[ x_r[n] = h[n] * w[n] = \sum_m h[n - m]w[m]. \]

Since \( w[m] = 0 \) for \( m \) odd, thus

\[ x_r[n] = \sum_m h[n - 2m]w[2m] \]

and since \( w[2m] = y[m] \), so

\[ x_r[n] = \sum_m h[n - 2m]y[m]. \]

Equivalently,

\[ x_r[n] = \sum_m h[n - 2m]x[2m]. \]

Look at the right-hand side: If \( n \) is even, say \( n = 2k \), then the right-hand side becomes

\[ \sum_m h[2k - 2m]x[2m] = x[2k]. \]
Whereas if \( n \) is odd, all terms are required for the reconstruction of \( x[n] \).

Notice that the reconstruction operation is noncausal: To reconstruct the odd values of \( x[n] \), all past and future even values are required.

5. In summary, the following system has the property that the output equals the input for every input in \( B_{\pi/2} \):

\[
\begin{align*}
    x[n] & \quad \downarrow 2 \quad y[n] \quad \uparrow 2 \quad w[n] \quad H(e^{jo}) \quad x[n] \\
    & \quad \text{gain 2} \quad \text{cutoff \( \pi/2 \)}
\end{align*}
\]

This system has the form

\[
\begin{align*}
    x[n] & \quad \downarrow 2 \quad S \quad \uparrow 2 \quad R \quad x[n] \\
    & \quad \text{gain 2}
\end{align*}
\]

Here, \( S \) stands for (discrete-time) \textit{sampler} and \( R \) for \textit{reconstructor}. Symbolically, \( RSx = x \) for all \( x \) in \( B_{\pi/2} \). Because \( R \) goes on the left of \( S \) in the equation \( RSx = x \), we say that \( R \) is a left-inverse of \( S \) when \( S \) is restricted to the subspace \( B_{\pi/2} \).

6. Finally, we mention without proof, though it’s very similar to what was done above, that the output also equals the input in the following system for every input bandlimited to the high-frequency range \( (\pi/2, \pi) \):

\[
\begin{align*}
    x[n] & \quad \downarrow 2 \quad \uparrow 2 \quad H(e^{jo}) \quad x[n] \\
    & \quad \text{gain 2} \quad \text{passband} \ [\pi/2, \pi]
\end{align*}
\]

The filter is an ideal high-frequency filter.
5 The Sampling Theorem by Impulse Modulation

Now we turn to the “standard” sampling problem, sampling a continuous-time signal. Impulse modulation is the most common way of developing the sampling theorem in an undergraduate course.

1. The definition of $S$, the sampling operator, is

$$x[n] = x_c(nT).$$

Thus the input is a continuous-time signal and the output a discrete-time signal. The arrow convention is this:

$$x_c(t) \rightarrow S \rightarrow x[n]$$

The assumption on the input will be $x_c \in \mathcal{B}_{\omega_N}$, that is,

$$X_c(j\omega) = 0, \quad |\omega| \geq \omega_N.$$ 

In words, the input is bandlimited and the sampling frequency is greater than twice the highest frequency in the input. Under this condition, $x_c(t)$ can be recovered from the sample values $x[n]$. But note that the reconstruction is non-causal. That is, for any given $t$ not a sampling instant, to reconstruct $x_c(t)$ requires all the infinitely many sample values, past and future.

2. The first step in this development is to derive the relationship between the input and output in the frequency domain. The formula is this:

$$X(e^{j\omega T}) = \frac{1}{T} \sum_k X_c(j\omega - jk\omega_s).$$  \hspace{1cm} (1)

This formula doesn’t assume the input is bandlimited. Here’s a schematic sketch of the graph of $X_c(j\omega)$:
Equation (1) is applied, and two terms in the series are shown here, namely, $k = 0$ and $k = 1$:

$$\frac{1}{T} X_c(j\omega)$$

There is aliasing, in the sense that higher frequency components of $x_c(t)$ become baseband frequency terms under sampling. On the other hand, if $x_c \in B_{\omega_N}$, then the terms won’t overlap:

$$\frac{1}{T} X_c(j\omega)$$

And therefore

$$X(e^{j\omega T}) = \frac{1}{T} X_c(j\omega), \quad |\omega| < \omega_N$$

or equivalently

$$X(e^{j\omega}) = \frac{1}{T} X_c \left( j\frac{\omega}{T} \right), \quad |\omega| < \pi.$$  

The corresponding graph is

$$X(e^{j\omega}) = \frac{1}{T} X_c(j\omega/T)$$

3. The proof of (1) involves an artifice called impulse modulation, shown here:
The signal \( s(t) \) is a series of impulses at the sampling instants. It multiples \( x_c(t) \) to produce \( x_s(t) \). Since

\[ x_c(t)\delta(t-nT) = x_c(nT)\delta(t-nT), \]

we have

\[ x_s(t) = \sum_n x[n]\delta(t-nT). \]

The reason for introducing this setup is that all three signals are continuous-time.

There are two steps in the proof. The first connects the outputs of the two systems.

**Step 1** \( X_s(j\omega) = X(e^{j\omega T}) \)

**Proof**

\[
X_s(j\omega) = \int x_s(t)e^{-j\omega t}dt \\
= \int x_c(t)s(t)e^{-j\omega t}dt \\
= \int x_c(t)\sum_m \delta(t-mT)e^{-j\omega t}dt \\
= \sum_m \int x_c(t)\delta(t-mT)e^{-j\omega t}dt \\
= \sum_m x_c(mT)e^{-j\omega mT}dt \\
= \sum_m x[m]e^{-j\omega mT}dt \\
= X(e^{j\omega T})
\]

The second step connects the input and output of the impulse modulation system.

**Step 2** \( X_s(j\omega) = \frac{1}{T} \sum_k X_c(j\omega - jk\omega_s) \).

11
Proof The signal \( s(t) \) is periodic of period \( T \) and it can therefore be expanded in a Fourier series:

\[
s(t) = \frac{1}{T} \sum_{m} e^{jm\omega t}.
\]

Then

\[
X_s(j\omega) = \int x_s(t)e^{-j\omega t}dt
= \int x_c(t)s(t)e^{-j\omega t}dt
= \int x_c(t)\frac{1}{T} \sum_{m} e^{jm\omega t}e^{-j\omega t}dt
= \frac{1}{T} \sum_{m} \int x_c(t)e^{jm\omega t}e^{-j\omega t}dt
= \frac{1}{T} \sum_{m} \int x_c(t)e^{-j(\omega-m\omega)t}dt
= \frac{1}{T} \sum_{m} X_c(j\omega - jm\omega).
\]

Finally, putting the two steps together proves (1).

Discussion

The continuous-time proof using impulse modulation is the standard way, but I feel it can be dangerous. If you ask students how to reconstruct a signal from its sampled values many will say “lowpass filter”. By which they mean time-domain filtering of the modulated signal. They seem not to realize that the reconstruction procedure is a discrete-time to continuous-time converter, but instead to believe the modulated signal \( \sum x[n]\delta(t-nT) \) is real. This isn’t surprising since many widely used texts emphasize this way of reconstructing \( x_c(t) \). For example

“It is possible to reconstruct the function from the train of impulses ... using some process of filtration.” [1]

“If we are given a sequence of samples, \( x[n] \), we can form an impulse train ... If this impulse train is the input to an ideal lowpass continuous-time filter ..., then the output of the filter will be [the reconstructed signal].” [3]

“The original analog signal can be perfectly reconstructed by passing this impulse train through a low-pass filter. ... While this method is mathematically pure, it is difficult to generate the required narrow pulses in electronics.” [4]
Also, students seem to accept the formula
\[
\sum_n \delta(t - nT) = \frac{1}{T} \sum_m e^{jm(2\pi/T)t},
\]
which is a form of the Poisson summation formula. But there must be some lack of real understanding: The right-hand side converges? And to something that equals zero everywhere except at the sampling times, where it is infinite? This formula can be interpreted only in the framework of Schwartz distributions [6].

6 The Sampling Theorem using Fourier Series

The derivation below shows that the reconstruction formula is a Fourier series expansion.

1. Students will recall Fourier series (FS) from their Signals and Systems course. FS is normally developed in the context of periodic waveforms, such as a square wave. But it makes just as much sense to focus on signals defined on a bounded time interval. Let \( f(t) \) be a function defined for \( 0 \leq t \leq 1 \). The goal is to represent \( f(t) \) as a series of sinusoids, namely,

\[
w_n(t) = e^{-j2\pi nt}.
\]

Here \( n \) ranges over all integers, positive, negative, and 0. These sinusoids are orthonormal in the following way:

\[
\int_0^1 w_n(t)\overline{w_m(t)}dt \text{ equals 0 if } n \neq m \text{ and equals 1 if } n = m.
\]

The overbar means complex-conjugate. Denote the integral by \( \langle w_n, w_m \rangle \). This inner product yields a norm: \( \| f \| = (\langle f, f \rangle)^{1/2} \).

Notice that \( w_0 \) is a DC signal, \( w_1 \) a sinusoid of period 1 sec., \( w_3 \) a sinusoid of period 1/2 sec., and so on. The FS expansion of \( f(t) \) is

\[
f(t) = \sum_n c_n w_n(t). \tag{2}
\]

Provided \( \| f \| < \infty \), the series converges to \( f \) in the sense that \( \lim_{N \to \infty} \| f - f_N \| = 0 \), where

\[
f_N(t) = \sum_{n=-N}^N c_n w_n(t).
\]

\[1\]The minus sign is for convenience.
The Fourier coefficients can be computed by taking inner products in (2):

\[ \langle f, w_m \rangle = \sum_n \langle c_n w_n, w_m \rangle = c_m. \]

Thus the Fourier coefficients in the FS are \( c_n = \langle f, w_n \rangle \).

2. Now let's turn to the sampling theorem. For simplicity we assume \( T = 1 \). The Fourier transform, \( X_c(j\omega) \), of the signal \( x_c(t) \), is zero outside the interval \( -\pi \leq \omega \leq \pi \). Under a mild condition, \( X_c(j\omega) \) has a FS expansion in terms of sinusoids. For the time interval \([0, 1]\) the sinusoids were \( e^{-j2\pi nt} \); consequently, for the frequency interval \([-\pi, \pi]\), the appropriate sinusoids are

\[ W_n(j\omega) = e^{-j\omega n}. \]

So that these are orthonormal, the inner product is taken to be

\[ \langle W_n, W_m \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} W_n(j\omega)W_m(j\omega) d\omega. \]

Thus the FS expansion of the Fourier transform \( X_c(j\omega) \) is

\[ X_c(j\omega) = \sum_n c_n W_n(j\omega) \]

and the Fourier coefficients are

\[
\begin{align*}
c_n &= \langle X_c, W_n \rangle \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} X_c(j\omega)W_n(j\omega) d\omega \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} X_c(j\omega)e^{j\omega n} d\omega.
\end{align*}
\]

You will recognize the last integral as the inverse Fourier transform of \( X_c(j\omega) \) at \( t = n \). That is, \( c_n = x[n] \). The FS expansion is thus seen to be

\[ X_c(j\omega) = \sum_n x[n]W_n(j\omega). \] (3)
3. It remains to take inverse Fourier transforms in (3). Let \( w_n(t) \) denote the inverse Fourier transform of

\[ W_n(j\omega) = e^{-j\omega n}. \]

Namely

\[ w_n(t) = w_0(t - n), \quad w_0(n) = \frac{\sin \pi t}{\pi t}. \]

Thus the inversion formula is

\[ x_c(t) = \sum_n x[n]w_0(t - n), \quad w_0(n) = \frac{\sin \pi t}{\pi t}. \]

**Discussion**

This derivation is rigorous, though not all details are given. In particular, completeness of the sinusoidal basis in the \( L^2 \)-norm is omitted.

**7 The Sampling Theorem in Hilbert Space**

This section is for graduate students who already have had an undergraduate Signals and Systems course and who know some Hilbert space theory. A very good reference in the context of Fourier theory is [2].

1. The following Hilbert spaces are used:

   (a) The space \( \ell^2 \) of square-summable real-valued discrete-time signals, with inner product

   \[ \langle x, y \rangle = \sum_n x[n]y[n]. \]

   (b) The space \( L^2(-\infty, \infty) \) of square-integrable real-valued continuous-time signals, with inner product

   \[ \langle x, y \rangle = \int_{-\infty}^{\infty} x(t)y(t)dt. \]
(c) The Fourier transform operator \( F \) maps \( L^2(-\infty, \infty) \) onto \( L^2(-j\infty, j\infty) \), the square-integrable functions on the imaginary axis; the inner product is
\[
\langle X, Y \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\omega)Y(j\omega) d\omega.
\]

The operator \( F \) preserves inner products and is a Hilbert-space isomorphism; its adjoint is denoted by \( F^* \).

(d) The subspace \( L^2(-j\pi, j\pi) \) of \( L^2(-j\infty, j\infty) \) such that \( X(j\omega) = 0 \) for \( |\omega| > \pi \).

2. Now we can turn to the sampling problem. We have continuous-time signals in \( L^2(-\infty, \infty) \). For simplicity, we take the sampling period to be \( T = 1 \) second, and then the Nyquist frequency is \( \pi \) rad/s. Accordingly, define \( B_\pi = F^*L^2(-j\pi, j\pi) \), the subspace of \( L^2(-\infty, \infty) \) of signals \( x_c(t) \) whose Fourier transforms equal zero for \( |\omega| > \pi \). For \( x_c \in B_\pi \), the inversion formula \( x_c = F^*X_c \) is
\[
x_c(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X_c(j\omega) e^{j\omega t} d\omega.
\]

Thus the sampled signal is
\[
x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X_c(j\omega) e^{j\omega n} d\omega.
\]

This is the inner product between \( X_c(j\omega) \) and \( e^{-j\omega n} \). So let us define
\[
W_n(j\omega) = e^{-j\omega n}
\]
so that we have
\[
x[n] = \langle X_c, W_n \rangle.
\]

By Parseval’s equality, the preceding equation holds for the inverse Fourier transforms:
\[
x[n] = \langle x_c, w_n \rangle.
\]

The functions \( W_n(j\omega) = e^{-j\omega n} \) are an orthonormal basis for the frequency-domain space \( L^2(-j\pi, j\pi) \) of bandlimited transforms. Consequently, \( w_n(t) \) are an orthonormal basis for the time-domain space \( B_\pi \). These basis functions are given by
\[
w_n(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\omega n} e^{j\omega t} d\omega = \frac{\sin \pi(t - n)}{\pi(t - n)}.
\]
3. We now define the sampling operator $S$:

$$S : B_\pi \rightarrow \ell^2, \quad x = Sx_c, \quad x[n] = \langle x_c, w_n \rangle.$$ 

The adjoint operator satisfies

$$\langle y, Sx_c \rangle = \langle S^*y, x_c \rangle,$$

that is,

$$\langle y, x \rangle = \langle S^*y, x_c \rangle,$$

that is,

$$\sum_n y[n]\langle x_c, w_n \rangle = \langle S^*y, x_c \rangle,$$

Everything is real-valued in this equation, and thus

$$\sum_n y[n]w_n = S^*y.$$

Therefore,

$$S^* : \ell^2 \rightarrow B_\pi, \quad y_c = S^*y, \quad y_c(t) = \sum_n y[n]w_n(t).$$

It follows that $S^* = S^{-1}$ and therefore $S$ is an isomorphism.

A nice block diagram to describe what we’ve done is this:

The input $x_c$ lies in $B_\pi$; the sampling block with operator $S$ produces $x$ in $\ell^2$; then the adjoint operator $S^*$ reproduces $x_c$. That is,

$$S^*Sx_c = x_c, \quad x_c \in B_\pi.$$

The continuous arrow signifies continuous-time, the dashed arrow, discrete time.

Let’s summarize:

**Theorem 1** The sampling operator is a Hilbert space isomorphism when restricted to the subspace of bandlimited signals.
7.1 Sample-rate Changing

This section shows an application of the operator formalism. Consider a continuous-time signal \( x_c(t) \) that belongs to \( L^2(-\infty, \infty) \). Suppose it is sampled at 2 Hz, producing \( x[n] = x_c(n/2) \). To avoid aliasing, assume it is bandlimited to \( 2\pi \). The problem is to compute \( y[n] = x_c(n/3) \), the same signal sampled at 3 Hz.

Let \( \mathcal{B}_{2\pi} \) denote the subspace of signals bandlimited to \([-2\pi, 2\pi]\). A sinusoidal basis for \( \mathcal{B}_{2\pi} \), in the frequency domain, is

\[
W_n(j\omega) = e^{-j(\omega/2)n}.
\]

The corresponding basis functions in the time domain, the inverse Fourier transforms, are denoted \( w_n(t) \):

\[
w_n(t) = w_0 \left( t - \frac{n}{2} \right), \quad w_0(t) = \frac{2\sin 2\pi t}{2\pi t}.
\]

These are orthogonal but not unit norm: \( \|w_n\|_2 = \sqrt{2} \).

Likewise, let \( \mathcal{B}_{3\pi} \) denote the subspace of signals bandlimited to \([-3\pi, 3\pi]\). The basis in the frequency domain is

\[
V_n(j\omega) = e^{-j(\omega/3)n},
\]

and in the time domain is

\[
v_n(t) = v_0 \left( t - \frac{n}{3} \right), \quad v_0(t) = \frac{3\sin 3\pi t}{3\pi t}.
\]

These are orthogonal but \( \|v_n\|_2 = \sqrt{3} \).

Further, let \( S_2 \) denote the sampling operator \( \mathcal{B}_{2\pi} \to \ell^2 \). It maps \( x_c(t) \) to \( x[n] = x_c(n/2) \). From the Fourier transform inversion formula,

\[
x[n] = \langle x_c, w_n \rangle.
\]

The inverse operator maps \( x[n] \) to

\[
x_c(t) = \frac{1}{2} \sum_n x[n] w_n(t).
\]

The coefficient \( 1/2 \) arises because \( \langle w_n, w_n \rangle = 2 \).

And let \( S_3 \) denote the sampling operator \( \mathcal{B}_{3\pi} \to \ell^2 \). It maps \( x_c(t) \) to \( y[n] = x_c(n/3) \), where

\[
y[n] = \langle x_c, v_n \rangle.
\]

The inverse maps \( y[n] \) to

\[
x_c = \frac{1}{3} \sum_n y[n] v_n.
\]
Then the sample-rate changer is simply \( S_3 S_3^{-1} \). It acts on a sequence \( x[n] \) to produce \( y[n] \), as follows. Start with \( x[n] \) in \( \ell^2 \). Apply \( S_3^{-1} \) to get

\[
x_c = \frac{1}{2} \sum_n x[n] w_n.
\]

Now apply \( S_3 \) to get \( y[n] \):

\[
y[n] = \langle x_c, v_n \rangle = \left\langle \frac{1}{2} \sum_m x[m] w_m, v_n \right\rangle = \sum_m x[m] \frac{1}{2} \langle w_m, v_n \rangle.
\]

This is in fact the answer, but let’s provide a bit more detail:

\[
\frac{1}{2} \langle w_n, v_m \rangle = \frac{1}{2} \langle W_n, V_m \rangle = \frac{1}{4\pi} \int_{-\infty}^{\infty} W_n(j\omega)V_m(-j\omega) d\omega = \frac{1}{4\pi} \int_{-2\pi}^{2\pi} e^{-j(\omega/2)n} e^{j(\omega/3)m} d\omega.
\]

The latter integral evaluates to

\[
h[2n - 3m], \quad h[k] = \frac{\sin \frac{\pi}{3} k}{\frac{\pi}{3} k}.
\]

Thus the sample-rate changer maps \( x[n] \) to \( y[n] \) according to

\[
y[n] = \sum_m h[2n - 3m] x[m].
\]

This is equivalent to: upsample by 3, then lowpass filter with gain 3, cutoff \( \pi/3 \), and then downsample by 2.

8 Final Discussion

Digital controllers have samplers too. Normally controllers are designed in continuous time and then implemented digitally using, for example, the bilinear transformation. The digital controller should achieve close to the performance of the designed controller. So signal reconstruction from sampled data is not a direct issue. However, impulse modulation is sometimes used in the analysis of sampled-data control systems—the old star notation. It is used for the same reason as in the sampling theorem: to convert a hybrid system,
with both discrete-time and continuous-time signals, into one with only continuous-time signals. The danger here is that students then don’t fully appreciate that the system is not time-invariant, but periodically time-varying. Instead, they are told that some input-output pairs do not have transfer functions. A rigorous control theory for sampled-data systems, modeling intersample behaviour, was in fact developed [5], without using impulse modulation, but it has not yet been brought to the undergraduate level.

Given that the Nyquist rate features in the sampling theorem and the Nyquist criterion is central in feedback stability, it might be expected that confusion could arise. Here’s a startling quote from a book on power-switching voltage converters: “According to Nyquist sampling theory, the unity-gain crossover frequency, \( f_1 \), must be less than half the switching frequency to ensure system stability.”

References


