Sampling and the Discrete Fourier Transform

Chapter 7

Sampling Methods

- Sampling is most commonly done with two devices, the sample-and-hold (S/H) and the analog-to-digital-converter (ADC)
- The S/H acquires a CT signal at a point in time and holds it for later use
- The ADC converts CT signal values at discrete points in time into numerical codes which can be stored in a digital system

Sampling Methods

Sample-and-Hold

During the clock, c(t), aperture time, the response of the S/H is the same as its excitation. At the end of that time, the response holds that value until the next aperture time.

Sampling Methods

An ADC converts its input signal into a code. The code can be output serially or in parallel.

Sampling Methods

Excitation-Response Relationship for an ADC

Sampling Methods
Sampling Methods

Encoded signal samples can be converted back into a CT signal by a digital-to-analog converter (DAC).

Pulse Amplitude Modulation

Pulse amplitude modulation was introduced in Chapter 6.

Modulator

\[ p(t) = \text{rect} \left( \frac{t}{w} \right) \cdot \frac{1}{T_s} \text{comb} \left( \frac{t}{T_s} \right) \]

Pulse Amplitude Modulation

The response of the pulse modulator is

\[ y(t) = x(t)p(t) = x(t) \left[ \text{rect} \left( \frac{t}{w} \right) \cdot \frac{1}{T_s} \text{comb} \left( \frac{t}{T_s} \right) \right] \]

and its CTFT is

\[ Y(f) = Wf \sum_{k=-\infty}^{\infty} \text{sinc}(Wkf)X(f - kf_s) \]

Pulse Amplitude Modulation

If the pulse train is modified to make the pulses have a constant area instead of a constant height, the pulse train becomes

\[ p(t) = \frac{1}{w} \text{rect} \left( \frac{t}{w} \right) \cdot \frac{1}{T_s} \text{comb} \left( \frac{t}{T_s} \right) \]

and the CTFT of the modulated pulse train becomes

\[ Y(f) = f \sum_{k=-\infty}^{\infty} \text{sinc}(Wkf)X(f - kf_s) \]

Pulse Amplitude Modulation

As the aperture time, \( w \), of the pulses approaches zero the pulse train approaches an impulse train (a comb function) and the replicas of the original signal’s spectrum all approach the same size. This limit is called impulse sampling.
The fundamental consideration in sampling theory is how fast to sample a signal to be able to reconstruct the signal from the samples.

**Sampling**

The CT Signal

- High Sampling Rate
- Medium Sampling Rate
- Low Sampling Rate

Claude Elwood Shannon

**Claude Elwood Shannon**

Claude Elwood Shannon

Shannon’s Sampling Theorem

As an example, let the CT signal to be sampled be

\[ x(t) = A \text{sinc} \left( \frac{t}{w} \right) \]

Its CTFT is

\[ X_{\text{CTFT}}(f) = Aw \text{rect}(wf) \]

The DTFT of the original signal is a rectangle.

The CTFT of the original signal is

\[ X_{\text{CTFT}}(f) = Aw \text{rect}(wf) \]

a rectangle.

The DTFT of the sampled signal is

\[ X_{\text{DTFT}}(F) = Awf \text{rect}(Fw_f) \]

or

\[ X_{\text{DTFT}}(F) = Awf \sum_{k=-\infty}^{\infty} \text{rect}(F - kw_f) \]

a periodic sequence of rectangles.
Shannon’s Sampling Theorem

If the "k = 0" rectangle from the DTFT is isolated and then the transformation,

\[ F \rightarrow \frac{f}{f_s} \]

is made, the transformation is

\[ \text{Awf}(w) \rightarrow \text{Awf}(w_f) \]

If this is now multiplied by \( T_s \), the result is

\[ T_s[\text{Awf}(w_f)] = \text{Awf}(w_f) = X_{\text{CTFT}}(f) \]

which is the CTFT of the original CT signal.

Shannon’s Sampling Theorem

In this example (but not in general) the original signal can be recovered from the samples by this process:

1. Find the DTFT of the DT signal.
2. Isolate the "k = 0" function from step 1.
3. Make the change of variable, \( F \rightarrow \frac{f}{f_s} \), in the result of step 2.
4. Multiply the result of step 3 by \( T_s \), in the result of step 2.
5. Find the inverse CTFT of the result of step 4.

The recovery process works in this example because the multiple replicas of the original signal’s CTFT do not overlap in the DTFT. They do not overlap because the original signal is bandlimited and the sampling rate is high enough to separate them.

Shannon’s Sampling Theorem

If the sampling rate is high enough, in the frequency range,

\[-\frac{f_s}{2} < f < \frac{f_s}{2}\]

the CTFT of the original signal and the CTFT of the impulse-sampled signal are identical except for a scaling factor of \( f_s \). Therefore, if the impulse-sampled signal were filtered by an ideal lowpass filter with the correct corner frequency, the original signal could be recovered from the impulse-sampled signal.
Shannon’s Sampling Theorem

Suppose a signal is bandlimited with this CTFT magnitude.

If we impulse sample it at a rate, $f_s = 4f_m$, the CTFT of the impulse-sampled signal will have this magnitude.

$$f_s = 4f_m$$

Shannon’s Sampling Theorem

Suppose the same signal is now impulse sampled at a rate, $f_s = 2f_m$.

The CTFT of the impulse-sampled signal will have this magnitude.

This is the minimum sampling rate at which the original signal could be recovered.

$$f_s = 2f_m$$

Shannon’s Sampling Theorem

Now the most common form of Shannon’s sampling theorem can be stated.

*If a signal is sampled for all time at a rate more than twice the highest frequency at which its CTFT is non-zero it can be exactly reconstructed from the samples.*

This minimum sampling rate is called the Nyquist rate. A signal sampled above the Nyquist rate is oversampled and a signal sampled below the Nyquist rate is undersampled.

Harry Nyquist

2/7/1889 - 4/4/1976

Timelimited and Bandlimited Signals

- The sampling theorem says that it is possible to sample a bandlimited signal at a rate sufficient to exactly reconstruct the signal from the samples.
- But it also says that the signal must be sampled for all time. This requirement holds even for signals which are timelimited (non-zero only for a finite time).

A signal that is timelimited cannot be bandlimited. Let $x(t)$ be a timelimited signal. Then

$$x(t) = x(t) \text{rect} \left( \frac{t-t_0}{\Delta t} \right)$$

The CTFT of $x(t)$ is

$$X(f) = X(f) * \Delta t \text{sinc}(\Delta f \Delta t)$$

Since this sinc function of $f$ is not limited in $f$, anything convolved with it will also not be limited in $f$ and cannot be the CTFT of a bandlimited signal.
Sampling Bandpass Signals

There are cases in which a sampling rate below the Nyquist rate can also be sufficient to reconstruct a signal. This applies to so-called bandpass signals for which the width of the non-zero part of the CTFT is small compared with its highest frequency. In some cases, sampling below the Nyquist rate will not cause the aliases to overlap and the original signal could be recovered by using a bandpass filter instead of a lowpass filter.

\[ f_s < 2f_2 \]

Interpolation

A CT signal can be recovered (theoretically) from an impulse-sampled version by an ideal lowpass filter. If the cutoff frequency of the filter is \( f_c \) then

\[ X(f) = X(f_c) \cdot \text{rect}\left(\frac{f - f_c}{f_s} \right) \]

If the sampling is at exactly the Nyquist rate, then

\[ x(t) = \sum_{n=-\infty}^{\infty} x(nT_s) \sin\left(\frac{\pi}{T_s} t - nT_s\right) \]

Practical Interpolation

Sinc-function interpolation is theoretically perfect but it can never be done in practice because it requires samples from the signal for all time. Therefore real interpolation must make some compromises. Probably the simplest realizable interpolation technique is what a DAC does.

\[ h(t) = \begin{cases} 1 & 0 < t < T_s \\ 0 & \text{otherwise} \end{cases} = \text{rect}\left(\frac{t - T_s}{2T_s}\right) \]
Practical Interpolation

If the signal is impulse sampled and that signal excites a ZOH, the response is the same as that produced by a DAC when it is excited by a stream of encoded sample values. The transfer function of the ZOH is a sinc function.

The ZOH suppresses aliases but does not entirely eliminate them.

A "natural" idea would be to simply draw straight lines between sample values. This cannot be done in real time because doing so requires knowledge of the "next" sample value before it occurs and that would require a non-causal system. If the reconstruction is delayed by one sample time, then it can be done with a causal system.

Non-Causal First-Order Hold

Causal First-Order Hold

Sampling a Sinusoid

Cosine sampled at twice its Nyquist rate. Samples uniquely determine the signal.

Cosine sampled at exactly its Nyquist rate. Samples do not uniquely determine the signal.

A different sinusoid of the same frequency with exactly the same samples as above.

Sine sampled at its Nyquist rate. All the samples are zero.

Adding a sine at the Nyquist frequency (half the Nyquist rate) to any signal does not change the samples.

It can be shown (p. 516) that the samples from two sinusoids, $x_1(t) = A \cos(2\pi f_0 t + \theta)$ and $x_2(t) = A \cos(2\pi [f_1 + k f_s] t + \theta)$ taken at the rate $f_s$, are the same for any integer value of $k$. 

Sine sampled slightly above its Nyquist rate

Two different sinusoids sampled at the same rate with the same samples.
Sampling DT Signals

One way of representing the sampling of CT signals is by impulse sampling, multiplying the signal by an impulse train (a comb). DT signals are sampled in an analogous way. If \( x[n] \) is the signal to be sampled, the sampled signal is

\[
x_{s}[n] = x[n]\text{comb}_{N_{s}}[n]
\]

where \( N_{s} \) is the discrete time between samples and the DT sampling rate is \( F_{s} = \frac{1}{N_{s}} \).

\[x[n] = x[n]\text{comb}_{N_{s}}[n]\]

The DTFT of the sampled DT signal is

\[
X_{s}(F) = X(F)\text{comb}_{N_{s}}(F)
\]

In this example the aliases do not overlap and it would be possible to recover the original DT signal from the samples. The general rule is that

\[
X_{s}(F) = X(F)\text{rect}\left(\frac{F}{F_{s}}\right)
\]

Interpolation is accomplished by passing the impulse-sampled DT signal through a DT lowpass filter.

\[
X_{d}(F) = X_{s}(F)\text{rect}\left(\frac{F}{F_{d}}\right)
\]

The opposite of downsampling is upsampling. It is simply the reverse of downsampling. If the original signal is \( x[n] \), then the upsampled signal is

\[
x_{u}[n] = \begin{cases} 
  x[n] & \text{an integer} \\
  0 & \text{otherwise}
\end{cases}
\]

Decimation

It is common practice, after sampling a DT signal, to remove all the zero values created by the sampling process, leaving only the non-zero values. This process is decimation, first introduced in Chapter 2. The decimated DT signal is

\[
x_{d}[n] = x[n]\text{comb}_{N_{d}}[n]
\]

and its DTFT is (p. 518)

\[
X_{d}(F) = X\left(\frac{F}{N_{d}}\right)
\]

Decimation is sometimes called downsampling.

The opposite of downsampling is upsampling. If the original signal is \( x[n] \), then the upsampled signal is

\[
x_{u}[n] = \begin{cases} 
  x[n] & \text{an integer} \\
  0 & \text{otherwise}
\end{cases}
\]

where \( N_{d} - 1 \) zeros have been inserted between adjacent values of \( x[n] \). If \( X(F) \) is the DTFT of \( x[n] \), then

\[
X_{u}(F) = X(N_{d}F)
\]

is the DTFT of \( x_{u}[n] \).
Sampling DT Signals

The signal, $x[n]$, can be lowpass filtered to interpolate between the non-zero values and form $x_s(n)$.

Bandlimited Periodic Signals

• If a signal is bandlimited it can be properly sampled according to the sampling theorem.
• If that signal is also periodic its CTFT consists only of impulses.
• Since it is bandlimited, there is a finite number of (non-zero) impulses.
• Therefore the signal can be exactly represented by a finite set of numbers, the impulse strengths.

The Discrete Fourier Transform

The most widely used Fourier method in the world is the Discrete Fourier Transform (DFT). It is defined by

$$x[n] = \frac{1}{N_F} \sum_{k=0}^{N_F-1} X[k] e^{j2\pi nk/N_F} \leftrightarrow X[k] = \frac{1}{N_F} \sum_{n=0}^{N_F-1} x[n] e^{-j2\pi nk/N_F}$$

This should look familiar. It is almost identical to the DTFS.

$$x[n] = \sum_{k=0}^{N_F-1} X[k] e^{j2\pi nk/N_F} \leftrightarrow X[k] = \frac{1}{N_F} \sum_{n=0}^{N_F-1} x[n] e^{-j2\pi nk/N_F}$$

The difference is only a scaling factor. There really should not be two so similar Fourier methods with different names but, for historical reasons, there are.
The sampled signal is \( x[n] = x(nT) \) and its DTFT is

\[
X(f) = \sum_{n=-\infty}^{\infty} x(nT) \delta(f - n/T) = \sum_{n=-\infty}^{\infty} x(nT) \delta(f - n/T)
\]

The last step in the process is to periodically repeat the time-domain signal, which samples the frequency-domain signal. Then there are two periodic impulse signals which are related to each other through the DTFS. Multiplication of the DTFS harmonic function by the number of samples in one period yields the DFT.

The sampled signal is taken at time, \( t = 0 \), at a rate, \( 1/T \), then the relationship between the CTFT of \( x(t) \) and the DFT of the samples taken from it is

\[
X(f_s) = T e^{-2\pi j m/N_f} \sum_{k=-\infty}^{\infty} X[k] \delta(f - kN_f)
\]

where \( f_s = 1/T \) and \( N_f \) is the number of samples. For those harmonic numbers, \( k \), for which \( k << N_f \),

\[
X(k) = T X[k]
\]

As the sampling rate and number of samples are increased, this approximation is improved.

The sampled and windowed signal is

\[
x[n] = \sum_{m=-\infty}^{\infty} x(n - mT)
\]

The original signal and the final signal are related by

\[
X_{sws}[k] = \frac{1}{N_f} \sum_{m=-N_f/2}^{N_f/2} \delta(f - m/N_f) X(f)
\]

In words, the CTFT of the original signal is transformed by replacing \( f \) with \( f/N_f \). That result is convolved with the window function, \( w[n] = [1, 0 \leq n < N_f, 0, \text{otherwise}] \). Then that result is transformed by replacing \( f \) by \( k/N_f \). Then that result is multiplied by \( f_s \).

It can be shown (pp. 530-532) that the DFT can be used to approximate samples from the CTFT. If the signal, \( x(t) \), is an energy signal and is causal and if \( N_f \) samples are taken from it over a finite time beginning at time, \( t = 0 \), at a rate, \( f_s \), then the relationship between the CTFT of \( x(t) \) and the DFT of the samples taken from it is

\[
X(f_s) = T e^{-2\pi j m/N_f} \sum_{k=-\infty}^{\infty} X[k] \delta(f - kN_f)
\]

where \( f_s = 1/T \) and \( N_f \) is the number of samples. For those harmonic numbers, \( k \), for which \( k << N_f \),

\[
X(k) = T X[k]
\]

As the sampling rate and number of samples are increased, this approximation is improved.

If a signal, \( x(t) \), is bandlimited and periodic and is sampled above the Nyquist rate over exactly one fundamental period the relationship between the CTFS of the original signal and the DFT of the samples is (pp. 532-535)

\[
X_{sws}[k] = \sum_{m=-\infty}^{\infty} c_{0,m} X[m]
\]

That is, the DFT is a periodic-repeated version of the CTFS, scaled by the number of samples. So the impulse strengths in the base period of the DFT, divided by the number of samples, is the same set of numbers as the strengths of the CTFS impulses.
**The Fast Fourier Transform**

Probably the most used computer algorithm in signal processing is the fast Fourier transform (FFT). It is an efficient algorithm for computing the DFT. Consider a very simple case, a set of four samples from which to compute a DFT. The DFT formula is

\[ X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn/N} \]

It is convenient to use the notation, \( W = e^{-j2\pi/n} \), because then the DFT formula can be written as

\[
\begin{bmatrix}
X[0] \\
X[1] \\
X[2] \\
X[3]
\end{bmatrix} = \begin{bmatrix}
w^0 & w^0 & w^0 & w^0 \\
w^2 & w^1 & w^2 & w^3 \\
w^3 & w^2 & w^1 & w^0 \\
w^1 & w^3 & w^2 & w^1
\end{bmatrix} \begin{bmatrix}
x[0] \\
x[1] \\
x[2] \\
x[3]
\end{bmatrix}
\]

It is possible to factor the matrix into the product of two matrices:

\[
\begin{bmatrix}
w^0 & w^0 & w^0 & w^0 \\
w^2 & w^1 & w^2 & w^3 \\
w^3 & w^2 & w^1 & w^0 \\
w^1 & w^3 & w^2 & w^1
\end{bmatrix} = \begin{bmatrix}
w^0 & 0 & 0 & 1 \\
w^2 & 0 & 0 & 1 \\
w^3 & 0 & 1 & 0 \\
w^1 & 0 & 1 & 0
\end{bmatrix} \begin{bmatrix}
w^0 & w^0 & w^0 & w^0 \\
w^0 & w^0 & w^0 & w^0 \\
w^0 & w^0 & w^0 & w^0 \\
w^0 & w^0 & w^0 & w^0
\end{bmatrix}
\]

It can be shown (pp. 552-553) that 4 multiplications and 12 additions are required, compared with 16 multiplications and 12 additions using the original matrix multiplication.

The number of multiplications required for an FFT algorithm of length, \( N = 2^p \), where \( p \) is an integer is \( \frac{2N}{p} \). The speed ratio in comparison with the direct DFT algorithm is \( \frac{Np}{2} \).