THE Electrical Engineering HANDBOOK

Editor-in-Chief

RICHARD C. DORF

University of California, Davis

CRC PRESS

Boca Raton Ann Arbor London Tokyo

67.4 The Sampling Theorem

R. J. Marks II

Much of that which is ordinal is modeled as analog. Most computational engines, on the other hand, are digital. Transforming from analog to digital is straightforward: we simply sample. Regaining the original signal from these samples and assessing the information lost in the sampling process are the fundamental questions addressed by the **sampling theorem.**

The fundamental result of the sampling theorem is, remarkably, that a bandlimited signal is uniquely specified by its sufficiently close equally spaced samples. Indeed, the sampling theorem illustrates how the original signal can be regained from knowledge of the samples and the sampling rate at which they were taken.

Popularization of the sampling theorem is credited to Shannon [1948] who, in 1948, used it to show the equivalence of the information content of a bandlimited signal and a sequence of discrete numbers. Shannon was aware of the pioneering work of Whittaker [1915] and Whittaker's son [1929] in formulating the sampling theorem. Kotel'nikov's [1933] independent discovery in the then Soviet Union deserves mention. Higgins [1985] credits Borel [1897] with first recognizing that a signal could be recovered from its samples.

Surveys of sampling theory are in the widely cited paper of Jerri [1977] and in two books by the author [1991, 1993]. Marvasti [1987] has written a book devoted to nonuniform sampling.

The Cardinal Series

If a signal has finite energy, the minimum **sampling rate** is equal to two samples per period of the highest frequency component of the signal. Specifically, if the highest frequency component of the signal is B Hz, then the signal, $x(t)$, can be recovered from the samples by

$$x(t) = \frac{1}{\pi} \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2B}\right) \frac{\sin[\pi(2Bt - n)]}{2Bt - n} \qquad (67.66)$$

The frequency B is also referred to as the signal's bandwidth and, if B is finite, $x(t)$ is said to be bandlimited. The signal, $x(t)$, is here being sampled at a rate of $2B$ samples per second. If sampling were done at a lower rate, the replications would overlap and the information about $X(\omega)$ [and thus $x(t)$] is irretrievably lost. Undersampling results in *aliased* data. The minimum sampling rate at which **aliasing** does not occur is referred to as the **Nyquist rate** which, in our example, is $2B$. Eq. (67.66) was dubbed the **cardinal series** by the junior Whittaker [1929].

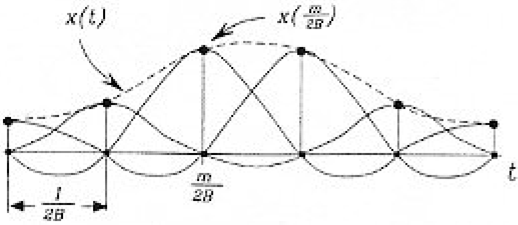

FIGURE 67.11 Illustration of the interpolation that results from the cardinal series. A sinc function, weighted by the sample, is placed at each sample bottom. The sum of the sincs exactly generates the original bandlimited function from which the samples were taken.

A signal is bandlimited in the low-pass sense if there is a $B > 0$ such that

$$X(\omega) = X(\omega)\Pi\left(\frac{\omega}{4\pi B}\right) \tag{67.67}$$

where the gate function $\Pi(\xi)$ is one for $\xi \leq 1/2$ and is otherwise zero, and

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\,dt \tag{67.68}$$

is the **Fourier transform** of $x(t)$. That is, the spectrum is identically zero for $\omega > 2\pi B$. The B parameter is referred to as the **signal's bandwidth**. The inverse Fourier transform is

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)\,e^{j\omega t}\,d\omega \tag{67.69}$$

The sampling theorem reduces the normally continuum infinity of ordered pairs required to specify a function to a countable—although still infinite—set. Remarkably, these elements are obtained directly by sampling.

How can the cardinal series interpolate uniquely the bandlimited signal from which the samples were taken? Could not the same samples be generated from another bandlimited signal? The answer is no. Bandlimited functions are smooth. Any behavior deviating from smooth would result in high-frequency components which in turn invalidates the required property of being bandlimited. The smoothness of the signal between samples precludes arbitrary variation of the signal there.

Let's examine the cardinal series more closely. Evaluate Eq. (67.74) at $t = m/2B$. Since sinc(n) is one for $n = 0$ and is otherwise zero, only the sample at $t = m/2B$ contributes to the interpolation at that point. This is illustrated in Fig. 67.11, where the reconstruction of a signal from its samples using the cardinal series is shown. The value of $x(t)$ at a point other than a sample location [e.g., $t = (m + \frac{1}{2})/2B$] is determined by all of the sample values.

Proof of the Sampling Theorem

Borel [1897] and Shannon [1948] both discussed the sampling theorem as the Fourier transform dual of the Fourier series. Let $x(t)$ have a bandwidth of B. Consider the periodic signal

$$Y(\omega) = \sum_{n=-\infty}^{\infty} X(\omega - 4\pi nB) \tag{67.70}$$

The function $Y(\omega)$ is a periodic function with period $4\pi B$. From Eq. (67.67) $X(\omega)$ is zero for

$\omega > 2\pi B$ and is thus finite in extent. The terms in Eq. (67.70) therefore do not overlap. Periodic functions can be expressed as a Fourier series.

$$Y(\omega) = \sum_{n=-\infty}^{\infty} \alpha_n \exp\left(\frac{-jn\omega}{2B}\right) \tag{67.71}$$

where the Fourier series coefficients are

$$\alpha_n = \frac{1}{4\pi B}\int_{-2\pi B}^{2\pi B} Y(\omega)\exp\left(\frac{jn\omega}{2B}\right) d\omega$$

or

$$\alpha_n = \frac{1}{2B} x\left(\frac{n}{2B}\right) \tag{67.72}$$

where we have used the inverse Fourier transform in Eq. (67.69). Substituting into the Fourier series in Eq. (67.71) gives

$$Y(\omega) = \frac{1}{2B}\sum_{n=-\infty}^{\infty} x\left(\frac{n}{2B}\right)\exp\left(\frac{-jn\omega}{2B}\right) \tag{67.73}$$

Since a period of $Y(\omega)$ is $X(\omega)$, we can get back the original spectrum by

$$X(\omega) = Y(\omega)\prod\left(\frac{\omega}{4\pi B}\right)$$

Substitute Eq. (67.73) and inverse transforming gives, using Eq. (67.69),

$$x(t) = \frac{1}{4\pi B}\int_{-2\pi B}^{2\pi B}\sum_{n=-\infty}^{\infty} x\left(\frac{n}{2B}\right)\exp\left(\frac{-jn\omega}{2B}\right)e^{j\omega t} d\omega$$

or

$$x(t) = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2B}\right)\text{sinc}\,(2Bt - n) \tag{67.74}$$

where

$$\text{sinc}(t) = \frac{\sin \pi t}{\pi t}$$

is the inverse Fourier transform of $\Pi(\omega/2\pi)$. Eq. (67.74) is, of course, the cardinal series.

The sampling theorem generally converges uniformly, in the sense that

$$\lim_{N\to\infty} | x(t) - x_N(t) |^2 = 0$$

where the truncated cardinal series is

$$x_N(t) = \sum_{n=-N}^{N} x\left(\frac{n}{2B}\right)\text{sinc}\,(2Bt - n) \tag{67.75}$$

Sufficient conditions for uniform convergence are [Marks, 1991]

1. the signal, $x(t)$, has finite energy, E,

$$E = \int_{-\infty}^{\infty} |x(t)|^2 \, dt < \infty$$

2. or $X(\omega)$ has finite area,

$$A = \int_{-\infty}^{\infty} |X(\omega)| \, d\omega < \infty$$

Care must be taken in the second case, though, when singularities exist at $\omega = \pm 2\pi B$. Here, sampling may be required to be strictly greater than $2B$. Such is the case, for example, for the signal, $x(t) = \sin(2\pi Bt)$. Although the signal is bandlimited, and although its Fourier transform has finite area, all of the samples of $x(t)$ taken at $t = n/2B$ are zero. The cardinal series in Eq. (67.74) will thus interpolate to zero everywhere. If the sampling rate is a bit greater than $2B$, however, the samples are not zero and the cardinal series will uniformly converge to the proper answer.

The Time-Bandwidth Product

The cardinal series requires knowledge of an infinite number of samples. In practice, only a finite number of samples are required. If most of the energy of a signal exists in the interval $0 \le t \le T$, and we sample at the Nyquist rate of $2B$ samples per second, then a total of $S = \langle 2BT \rangle$ samples are taken. ($\langle \theta \rangle$ denotes the largest number not exceeding θ.) The number S is a measure of the degrees of freedom of the signal and is referred to as its **time-bandwidth product.** A 5-min single-track audio recording requiring fidelity up to 20,000 Hz, for example, requires a minimum of $S = 2 \times 20{,}000 \times 5 \times 60 = 12$ million samples. In practice, audio sampling is performed well above the Nyquist rate.

Sources of Error

Exact interpolation using the cardinal series assumes that (1) the values of the samples are known exactly, (2) the sample locations are known exactly, and (3) an infinite number of terms are used in the series. Deviation from these requirements results in interpolation error due to (1) data noise, (2) jitter, and (3) truncation, respectively. The effect of data error on the restoration can be significant. Some innocently appearing sampling theorem generalizations, when subjected to performance analysis in the presence of data error, are revealed as ill-posed. In other words, a bounded error on the data can result in unbounded error on the restoration [Marks, 1991].

Data Noise

The source of data noise can be the signal from which samples are taken, or from round-off error due to finite sampling precision. If the noise is additive and random, instead of the samples

$$x\left(\frac{n}{2B}\right)$$

we must deal with the samples

$$x\left(\frac{n}{2B}\right) + \xi\left(\frac{n}{2B}\right)$$

(a) **Interpolation.** The tails of a signal are known and we wish to restore the middle.

(b) **Extrapolation.** We wish to generate the tails of a function with knowledge of the middle.

(c) **Prediction.** A signal for $t > 0$ is to be estimated from knowledge of the signal for $t < 0$.

Final Remarks

Since its popularization in the late 1940s, the sampling theorem has been studied in depth. More than 1000 papers have been generated on the topic [Marks, 1993]. Its understanding is fundamental in matching the largely continuous world to digital computation engines.

Defining Terms

Aliasing: A phenomenon that occurs when a signal is undersampled. High-frequency information about the signal is lost.

Cardinal series: The formula by which samples of a bandlimited signal are interpolated to form a continuous time signal.

Fourier transform: The mathematical operation that converts a time-domain signal into the frequency domain.

Jitter: A sample is temporally displaced by an unknown, usually small, interval.

Kramer's generalization: A sampling theory based on other than Fourier transforms and frequency.

Lagrangian interpolation: A classic interpolation procedure used in numerical analysis. The sampling theorem is a special case.

Nyquist rate: The minimum sampling rate that does not result in aliasing.

Papoulis' generalization: A sampling theory applicable to many cases wherein signal samples are obtained either nonuniformly and/or indirectly.

Sampling rate: The number of samples per second.

Sampling theorem: Samples of a bandlimited signal, if taken close enough together, exactly specify the continuous time signal from which the samples were taken.

Signal bandwidth: The maximum frequency component of a signal.

Time bandwidth product: The product of a signal's duration and bandwidth approximates the number of samples required to characterize the signal.

Truncation error: The error that occurs when a finite number of samples are used to interpolate a continuous time signal.

References

E. Borel, "Sur l'interpolation," *C.R. Acad. Sci. Paris*, vol. 124, pp. 673–676, 1897.

J. R. Higgins, "Five short stories about the cardinal series," *Bull. Am. Math. Soc.*, vol. 12, pp. 45–89, 1985.

A. J. Jerri, "The Shannon sampling theorem—its various extension and applications: a tutorial review," *Proc. IEEE*, vol. 65, pp. 1565–1596, 1977.

V. A. Kotel'nikov, "On the transmission capacity of 'ether' and wire in electrocommunications," *Izd. Red. Upr. Svyazi RKKA* (Moscow), 1933.

R. J. Marks II, *Introduction to Shannon Sampling and Interpolation Theory*, New York: Springer-Verlag, 1991.

R. J. Marks II, Ed., *Advanced Topics in Shannon Sampling and Interpolation Theory*, New York: Springer-Verlag, 1993.

F. A. Marvasti, *A Unified Approach to Zero-Crossing and Nonuniform Sampling*, Oak Park, Ill.: Nonuniform, 1987.

Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379, 623, 1948.

T. Whittaker, "On the functions which are represented by the expansions of the interpolation theory," *Proc. Royal Society of Edinburgh*, vol. 35, pp. 181–194, 1915.

M. Whittaker, "The Fourier theory of the cardinal functions," *Proc. Math. Soc. Edinburgh*, vol. 1, pp. 169–176, 1929.

Information

An in-depth study of the sample theorem and its numerous variations is provided in R. J. Marks II, Ed., *Introduction to Shannon Sampling and Interpolation Theory*, New York: Springer-Verlag, 1991.

In-depth studies of modern sampling theory with over 1000 references are available in R. J. Marks II, Ed., *Advanced Topics in Shannon Sampling and Interpolation Theory*, New York: Springer-Verlag, 1993.

The specific case of nonuniform sampling is treated in the monograph by F. A. Marvasti, *A Unified Approach to Zero-Crossing and Nonuniform Sampling*, Oak Park, Ill: Nonuniform, 1987.

The sampling theorem is treated generically in the *IEEE Transactions on Signal Processing*. For applications, topical journals are the best source of current literature.