Mathematical Foundations of Data Sciences

Gabriel Peyré
CNRS & DMA, École Normale Supérieure

CHAPTER 1

Shannon Sampling Theory

September 9, 2026

Digital measurements retain only samples of a continuous signal, so reconstruction depends on what information sampling preserves. Understanding these limits guides the choice of sampling rate and the filtering applied before acquisition. Starting from continuous and discrete signal models, we use Fourier analysis to prove a sampling theorem, explain aliasing, and distinguish exact interpolation from the additional error introduced by quantization.

Shannon's foundational work of 1948–1949 addresses three closely related topics:

1. Sampling determines when a continuous signal can be recovered from discrete measurements. Quantization then replaces the measured values by symbols from a finite alphabet.
2. Source coding uses symbol probabilities to represent a sequence of symbols compactly as a binary string.
3. Channel coding adds redundancy to protect the coded sequence against transmission errors, such as flipped bits. It is also known as error-correcting coding.

Chapter 3 develops source coding theory. We do not cover channel coding. The main reference for this chapter is [1].

1.1 Continuous and Discrete Signals

To analyze signal-processing methods, we often begin with a continuous model of the physical input to a sensor. Examples include microphones, digital cameras, and medical imaging devices. For example, a one-dimensional signal can be modeled by $f_0 \in L^2([0,1])$, where $[0,1]$ is the acquisition interval, often representing time. A grayscale image can be modeled by $f_0 \in L^2([0,1]^2)$ on the unit square.

Although sound and natural images provide many of our examples, most of the methods extend to multidimensional data. After normalizing the signal values, these data can be represented by mappings

$$f_0 : [0,1]^d \longrightarrow [0,1]^s$$

Here d is the dimension of the acquisition domain and s is the number of channels. Grayscale images have $(d,s) = (2,1)$, grayscale videos have $(d,s) = (3,1)$, and RGB color images have $(d,s) = (2,3)$. Multispectral images have $d = 2$ and many channels, each recording a different wavelength range. Figures 1.1 and 1.2 illustrate these types of data.

1.1.1 Acquisition and Sampling

Signal acquisition maps a continuous signal to a finite-dimensional vector of measurements. A microphone records samples along a time axis, whereas a digital camera records samples on a two-dimensional pixel grid. Abstractly, an acquisition operator maps

$$f_0 \in L^2([0,1]^d) \mapsto f \in \mathbb{C}^N$$

Figure 1.3 shows examples of discretized signals.

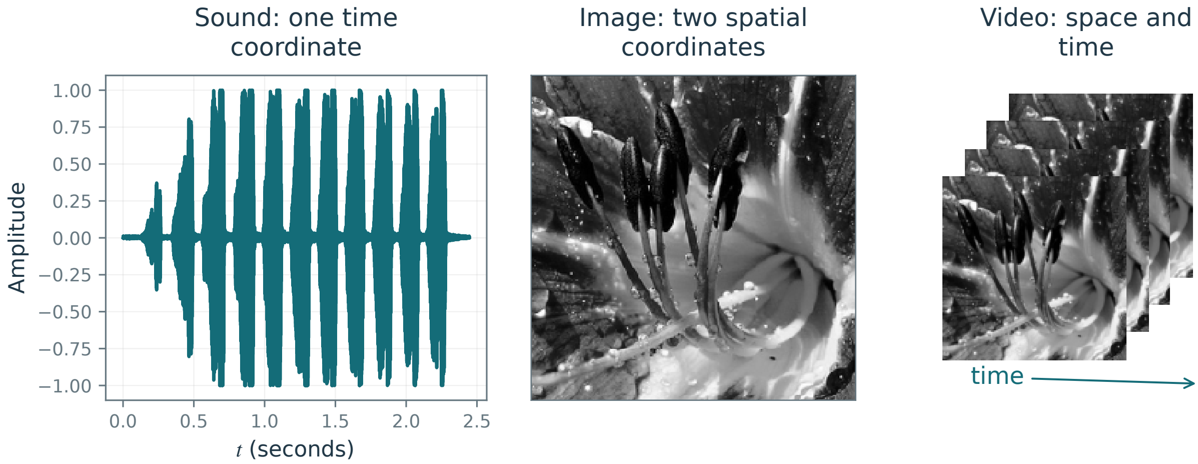


Figure 1.1. Examples of sound ($d = 1$), an image ($d = 2$), and a video ($d = 3$).

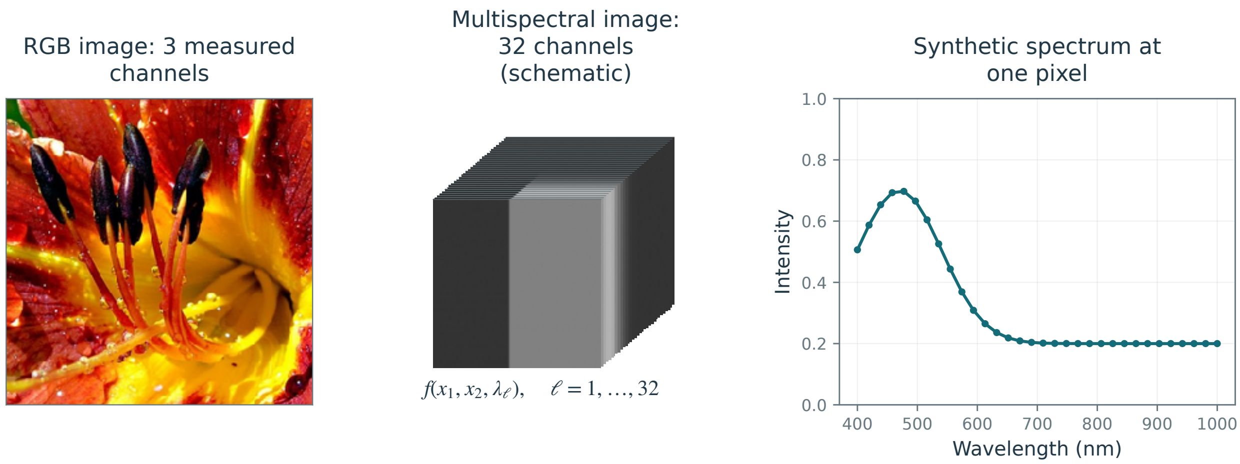


Figure 1.2. A color flower image and a synthetic multispectral image with 32 channels, with an example spectrum.

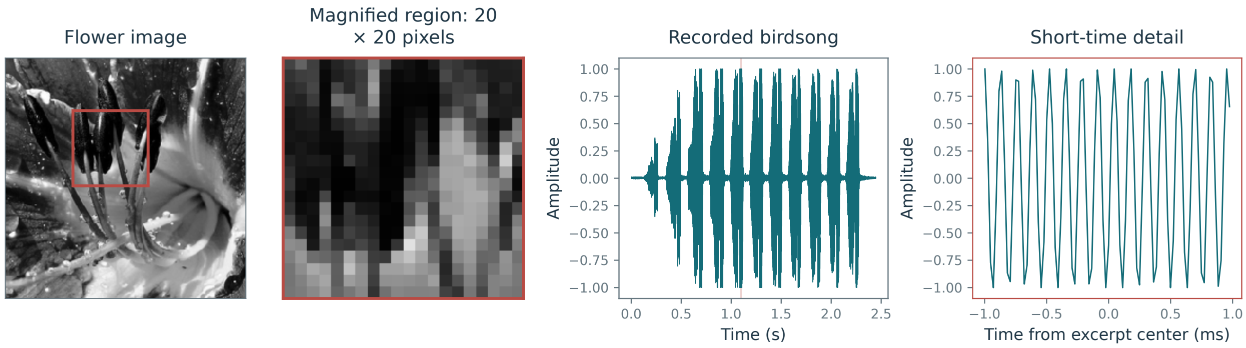


Figure 1.3. Spatial and temporal detail: a flower image with a pixelated crop, and a recorded birdsong waveform with a short-time enlargement. Matching red boxes identify the spatial crop.

1.1.2 Linear Translation-Invariant Sampling

To model translation-invariant sampling, extend the signal outside its acquisition domain according to the chosen boundary convention, convolve it with a fixed impulse response h, and sample the result:

$$f_n = \int_{\mathbb{R}} f_0(x)h(n/N - x)\,\mathrm{d}x = (f_0 \star h)(n/N). \tag{1.1}$$

The impulse response $h(x)$ depends on the device. It is typically a smooth low-pass kernel concentrated near $x = 0$. Its effective width S controls spatial resolution: a wide kernel blurs the signal, whereas a narrow kernel may fail to suppress frequencies that cause aliasing. A width of order $1/N$ balances these effects at the sampling scale.

Section 1.2 explains when a bandlimited signal can be reconstructed from its samples. Smoothness alone does not guarantee exact recovery.

1.2 Shannon Sampling Theorem

The Fourier transform. For $f \in L^1(\mathbb{R})$, its Fourier transform is defined as

$$\forall\, \omega \in \mathbb{R}, \quad \hat{f}(\omega) \overset{\text{def.}}{=} \int_{\mathbb{R}} f(x)e^{-\mathrm{i}x\omega}\mathrm{d}x. \tag{1.2}$$

For $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, Plancherel's identity gives $\|\hat{f}\|_2^2 = 2\pi\|f\|_2^2$. The map $f \mapsto \hat{f}$ therefore extends continuously to $L^2(\mathbb{R})$. This extension is the L^2 limit, as $T \to +\infty$, of the truncated integrals $\int_{-T}^{T} f(x)e^{-\mathrm{i}x\omega}\mathrm{d}x$. When $\hat{f} \in L^1(\mathbb{R})$, Fourier inversion gives

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\omega)e^{\mathrm{i}x\omega}\mathrm{d}\omega, \tag{1.3}$$

where the equality holds almost everywhere. The right-hand side defines the continuous representative of f, which vanishes at $\pm\infty$. We use this representative whenever pointwise values are needed.

The Fourier transform $\mathcal{F} : f \mapsto \hat{f}$ exchanges regularity and decay. For an integer $p \geqslant 0$, if $f \in W^{p,1}(\mathbb{R})$, then $\mathcal{F}(f^{(p)})(\omega) = (\mathrm{i}\omega)^p\hat{f}(\omega)$ so that $|\hat{f}(\omega)| = O(1/|\omega|^p)$. Conversely,

$$\int_{\mathbb{R}} (1 + |\omega|)^p|\hat{f}(\omega)|\mathrm{d}\omega < +\infty \quad \Longrightarrow \quad f \in C^p(\mathbb{R}). \tag{1.4}$$

For example, the decay estimate $\hat{f}(\omega) = O(1/|\omega|^{p+1+\varepsilon})$ for $\varepsilon > 0$ implies $f \in C^p(\mathbb{R})$.

A related measure of smoothness is the squared Sobolev norm $\|f\|_{H^1}^2 = (2\pi)^{-1}\int_{\mathbb{R}}(1 + |\omega|^2)|\hat{f}(\omega)|^2\,\mathrm{d}\omega = \|f\|_2^2 + \|f'\|_2^2$, where f' is the weak derivative. Sobolev spaces provide a natural setting for weak solutions of PDEs. Later chapters use Sobolev regularity to control approximation and denoising errors.

Fourier series. Let $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ denote the circle of length 2π. A function $f \in L^2(\mathbb{T})$ is identified with a 2π-periodic function on $\mathbb{R}$. Its Fourier coefficients are

$$\forall\, k \in \mathbb{Z}, \quad \hat{f}_k \overset{\text{def.}}{=} \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-\mathrm{i}xk}\mathrm{d}x.$$

Equivalently, $\hat{f}_k = \langle f, e_k\rangle$, where $\langle f, g\rangle \overset{\text{def.}}{=} \frac{1}{2\pi}\int_{\mathbb{T}} f(x)\bar{g}(x)\mathrm{d}x$ and $e_k(x) \overset{\text{def.}}{=} e^{\mathrm{i}xk}$. The functions $(e_k)_{k\in\mathbb{Z}}$ form an orthonormal basis for this inner product, giving the expansion

$$f = \sum_{k\in\mathbb{Z}} \langle f, e_k\rangle e_k \tag{1.5}$$

which means $\|f - \sum_{k=-N}^{N}\langle f, e_k\rangle e_k\|_{L^2(\mathbb{T})} \to 0$ for $N \to +\infty$. For a differentiable function, the series (1.5) converges pointwise at every $x \in \mathbb{T}$. If f is of class C^2 on $\mathbb{T}$, then $\hat{f}_k = O(1/k^2)$, which gives normal convergence and hence uniform convergence. For a piecewise C^1 function, the Fourier series converges at a jump to the average of the one-sided limits. Gibbs oscillations prevent uniform convergence across the jump.

Poisson formula. The Poisson formula relates sampling in the signal domain to periodization in the Fourier domain. For a function $h(\omega)$ on $\mathbb{R}$, define its periodization by

$$h_P(\omega) \stackrel{\text{def.}}{=} \sum_n h(\omega - 2\pi n). \tag{1.6}$$

If $h \in L^1(\mathbb{R})$, the series converges absolutely almost everywhere and $\|h_P\|_{L^1(\mathbb{T})} \leqslant \|h\|_{L^1(\mathbb{R})}$, with equality for nonnegative functions. Here the norm on $\mathbb{T}$ uses unnormalized Lebesgue measure over one period. Indeed,

$$|h_P(x)| \leqslant (|h|_P)(x), \quad \Longrightarrow \quad \|h_P\|_{L^1} \leqslant \||h|_P\|_{L^1} = \|h\|_{L^1}.$$

For $h = \hat{f}$, Proposition 1.1 states that the following diagram

$$\begin{array}{cccccc} & f(x) & \xrightarrow{\mathcal{F}} & \hat{f}(\omega) & \\ \text{sampling} & \downarrow & & \downarrow & \text{periodization} \\ & (f(n))_n & \xrightarrow{\text{Fourier series}} & \sum_n f(n)e^{-\mathrm{i}\omega n} & \end{array}$$

is commutative. The minus sign in $\sum_n f(n)e^{-\mathrm{i}\omega n}$ is opposite to the sign in Fourier series synthesis; this accounts for the index reversal in the proof below.

Proposition 1.1 (Poisson formula). *Assume that $f \in C(\mathbb{R})$, that $\hat{f}$ has compact support, and that $|f(x)| \leqslant C(1+|x|)^{-3}$ for some C. Then*

$$\forall\, \omega \in \mathbb{R}, \quad \sum_n f(n)e^{-\mathrm{i}\omega n} = \hat{f}_P(\omega). \tag{1.7}$$

Proof. The decay of f makes $\hat{f}$ continuously differentiable. Its compact support makes the periodization locally a finite sum, so $(\hat{f})_P$ is also C^1. It therefore equals its Fourier series:

$$(\hat{f})_P(\omega) = \sum_k c_k e^{\mathrm{i}k\omega}, \tag{1.8}$$

where

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} (\hat{f})_P(\omega)e^{-\mathrm{i}k\omega}\mathrm{d}\omega = \frac{1}{2\pi}\int_0^{2\pi} \sum_n \hat{f}(\omega - 2\pi n)e^{-\mathrm{i}k\omega}\mathrm{d}\omega.$$

Absolute integrability follows from

$$\int_0^{2\pi} \sum_n |\hat{f}(\omega - 2\pi n)e^{-\mathrm{i}k\omega}|\mathrm{d}\omega = \int_{\mathbb{R}} |\hat{f}|$$

which is finite because $\hat{f}$ is continuous and compactly supported. We may therefore exchange the sum and integral:

$$c_k = \sum_n \frac{1}{2\pi}\int_0^{2\pi} \hat{f}(\omega - 2\pi n)e^{-\mathrm{i}k\omega}\mathrm{d}\omega = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(\omega)e^{-\mathrm{i}k\omega}\mathrm{d}\omega = f(-k)$$

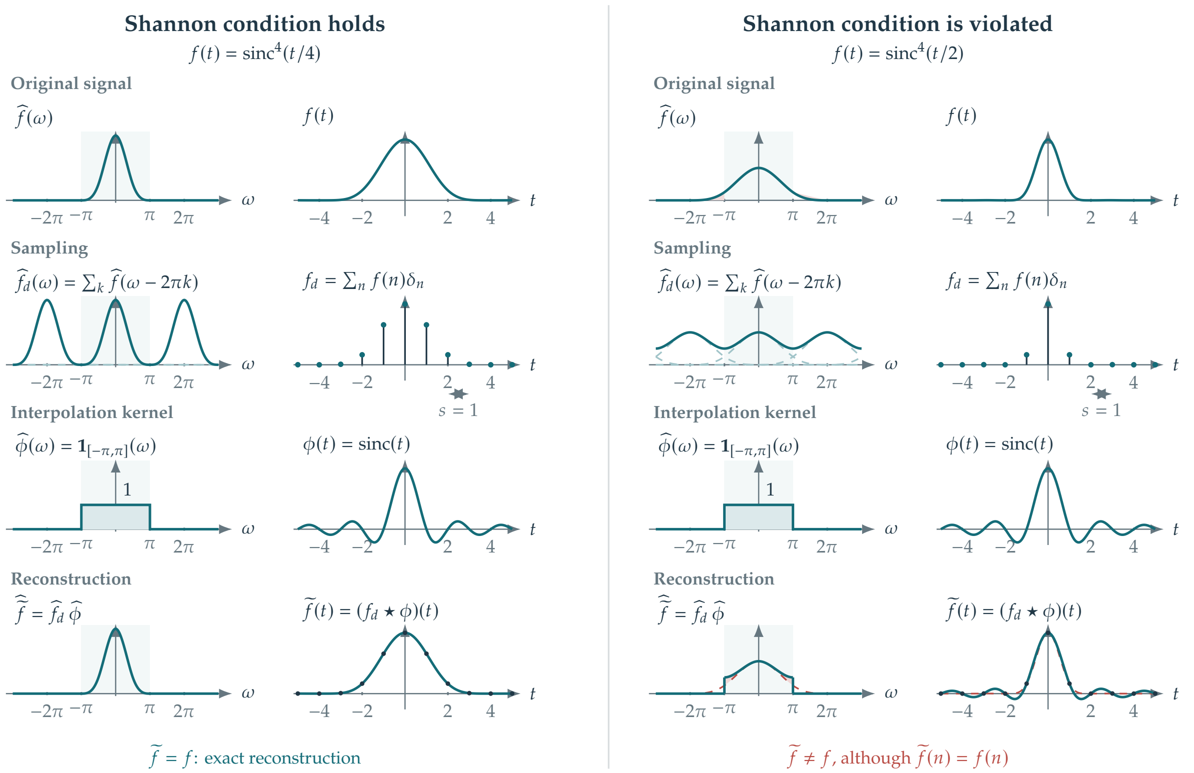
where we used the inverse Fourier transform formula (1.3), which is valid because $\hat{f} \in L^1(\mathbb{R})$. □

Shannon's sampling theorem. Shannon's sampling theorem gives a sufficient condition for reconstructing a signal from the samples $f \mapsto (f(ns))_n$ at spacing $s > 0$. The bandlimiting condition is $\mathrm{supp}(\hat{f}) \subset [-\pi/s, \pi/s]$. By Fourier inversion (1.3), such a signal is C^∞ and has an entire extension to the complex plane. Shannon's 1949 paper placed the sampling theorem at the heart of communication theory, while acknowledging earlier interpolation results and Nyquist's work on transmission rates. Figure 1.4 illustrates the reconstruction argument and the aliasing that occurs when spectral replicas overlap; see also Figure 1.7.

Theorem 1.2. *If $f \in C(\mathbb{R})$, $|f(x)| \leqslant C(1+|x|)^{-3}$ for some C and $\mathrm{supp}(\hat{f}) \subset [-\pi/s, \pi/s]$, then one has*

$$\forall\, x \in \mathbb{R}, \quad f(x) = \sum_n f(ns)\,\mathrm{sinc}(x/s - n) \quad \textit{where} \quad \mathrm{sinc}(u) = \frac{\sin(\pi u)}{\pi u}, \quad \mathrm{sinc}(0) = 1 \tag{1.9}$$

with uniform convergence.



$$\mathrm{sinc}(t) = \frac{\sin(\pi t)}{\pi t}, \qquad \widetilde{f}(t) = \sum_{n\in\mathbb{Z}} f(n)\,\mathrm{sinc}(t-n)$$

Pale band: $[-\pi, \pi]$. Dashed spectral curves: replicas. Red shading: aliased mass.
Dashed red in the reconstruction row: the original signal and spectrum. Dots: the shared samples.

Figure 1.4. Sampling and Shannon interpolation in the frequency and time domains, with unit sample spacing. Left: the bandlimiting condition of Theorem 1.2 holds. Right: overlapping spectral replicas cause aliasing; the reconstructed signal still passes through every sample. The rows show the original signal, sampling, the interpolation kernel, and reconstruction.

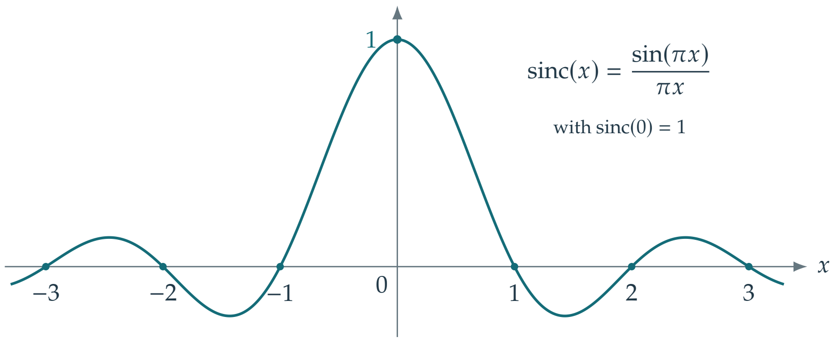


Figure 1.5. The sinc interpolation kernel.

Proof. Set $g \stackrel{\text{def.}}{=} f(s\cdot)$. Then $\hat{g} = 1/s\hat{f}(\cdot/s)$, as follows from the substitution $z = sx$:

$$\hat{g}(\omega) = \int f(sx)e^{-\mathrm{i}\omega x}\mathrm{d}x = \frac{1}{s}\int f(z)e^{-\mathrm{i}(\omega/s)z}\mathrm{d}z = \hat{f}(\omega/s)/s,$$

It is therefore enough to prove the result for unit sample spacing:

$$g(x) = \sum_n g(n)\,\mathrm{sinc}(x-n),$$

For the rest of the proof, write f in place of g. The compact support hypothesis implies $\hat{f}(\omega) = 1_{[-\pi,\pi]}(\omega)\hat{f}_P(\omega)$. Combining the inversion formula (1.3) with the Poisson formula (1.7) gives

$$f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\hat{f}_P(\omega)e^{\mathrm{i}\omega x}\mathrm{d}\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_n f(n)e^{\mathrm{i}\omega(x-n)}\mathrm{d}\omega.$$

The decay of f gives $\int_{-\pi}^{\pi}\sum_n |f(n)e^{\mathrm{i}\omega(x-n)}|\mathrm{d}\omega = 2\pi\sum_n|f(n)| < +\infty$. We may therefore exchange summation and integration:

$$f(x) = \sum_n f(n)\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{\mathrm{i}\omega(x-n)}\mathrm{d}\omega = \sum_n f(n)\,\mathrm{sinc}(x-n).$$

The series converges uniformly by the Weierstrass test, since $|\mathrm{sinc}(x-n)| \leqslant 1$ and $\sum_n |f(n)| < \infty$. □

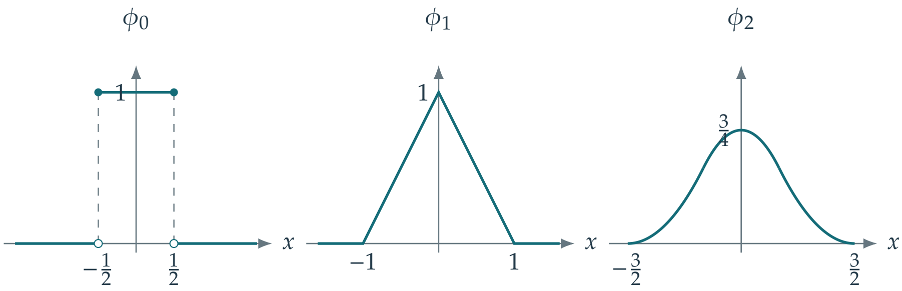
The sinc kernel decays slowly and oscillates, making it inconvenient for local interpolation. Oversampling gives more flexibility: if $\mathrm{supp}(\hat{f}) \subset [-\pi/s', \pi/s']$ with $s' > s$, the spectrum occupies a strict subinterval of the Nyquist interval. One can then choose a reconstruction kernel with a smooth, compactly supported Fourier transform. The resulting kernel decays faster than every inverse power of time. Smooth compact spectral support does not, in general, imply exponential decay.

For local interpolation, cardinal B-splines offer another useful family of kernels. Define $\varphi_0 = 1_{[-1/2,1/2]}$ and $\varphi_k = \varphi_{k-1} \star \varphi_0$. The function φ_k is piecewise polynomial of degree k, with support $[-(k+1)/2, (k+1)/2]$. For $k \geqslant 1$, it belongs to C^{k-1} and has a bounded weak derivative of order k. The reconstruction takes the form $f \approx \tilde{f} \stackrel{\text{def.}}{=} \sum_n a_n\varphi(\cdot - n)$, with $\varphi = \varphi_k$. The interpolation conditions $\tilde{f}(n) = f(n)$ define a linear system for the coefficients $(a_n)_n$ in terms of the samples $(f(n))_n$. For general sample data, the identity $a_n = f(n)$ holds only for $k \in \{0,1\}$, corresponding to piecewise constant and piecewise affine interpolation. A common choice is cubic spline interpolation, which corresponds to $k = 3$.

`Associated code: test_sampling.m`

When the spectral replicas overlap, sampling creates aliasing: high-frequency components become indistinguishable from lower-frequency components. A low-pass filter applied before sampling attenuates frequencies beyond the Nyquist interval $[-\pi/s, \pi/s]$. This reduces aliasing at the cost of losing high-frequency information; see Figure 1.7.

Quantization. To store or transmit a sampled signal digitally, one quantizes its values to finite precision. Section 6.1 develops transform coding, in which an invertible linear transform, usually orthogonal, is applied



$$\phi_0 = \mathbf{1}_{[-1/2,1/2]}, \qquad \phi_k = \phi_{k-1} * \phi_0$$

Figure 1.6. Cardinal splines as basis functions for interpolation.

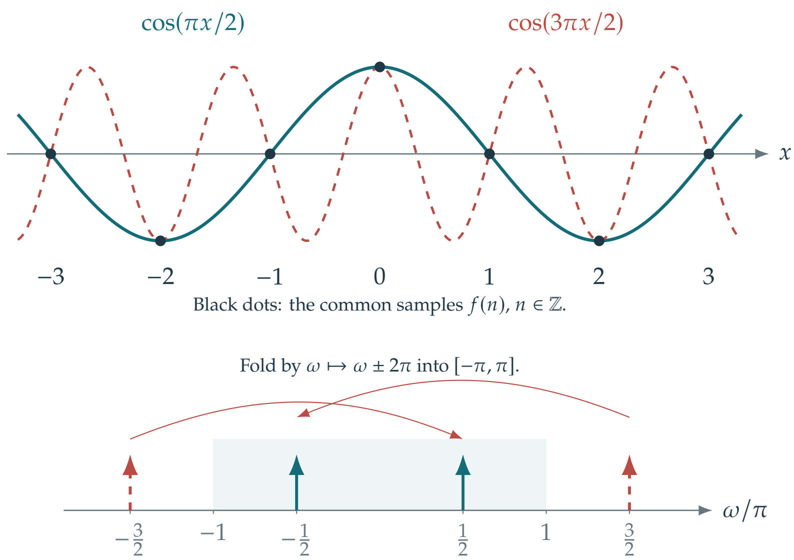


Figure 1.7. Aliasing of a sine wave. This periodic example is interpreted through its discrete spectral lines rather than as an $L^2(\mathbb{R})$ signal.

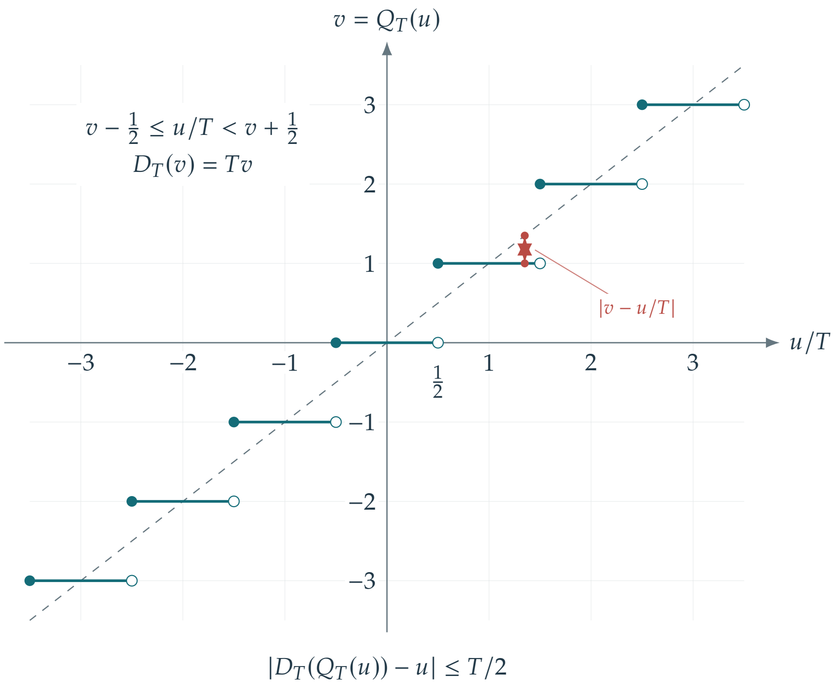


Figure 1.8. A uniform scalar quantizer.

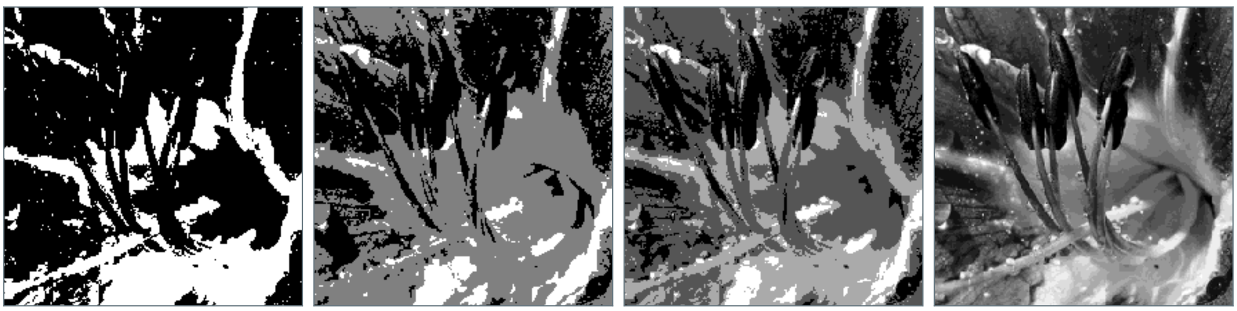


Figure 1.9. Quantizing an image with $K \in \{2, 3, 4, 16\}$ gray levels. The normalized intensities lie in $[0, 1)$; finite-level quantizers also specify clipping or endpoint conventions.

before quantization. Transforming the data can make the coefficients easier to compress. Directly quantizing the sampled values is a simpler alternative.

For example, a step size $s = 1/N$ gives the finite vector $(u_n \overset{\text{def.}}{=} f(n/N))_{n=1}^{N} \in \mathbb{R}^N$. This records only finitely many samples, which do not determine a general bandlimited signal. Exact recovery from these data requires additional assumptions, such as a specified finite-dimensional signal model.

For a quantization step $T > 0$, the quantizer $v_n = Q_T(u_n) \in \mathbb{Z}$ records the index of the nearest multiple of T, i.e.

$$v = Q_T(u) \quad \Leftrightarrow \quad v - \frac{1}{2} \leqslant u/T < v + \frac{1}{2},$$

as illustrated by the quantizer diagram. Dequantization replaces each index by the cell midpoint $D_T(v) \overset{\text{def.}}{=} Tv$, which minimizes the worst-case error within that cell. The resulting error is bounded by $T/2$: $|D_T(Q_T(u)) - u| \leqslant T/2$. Quantization introduces the loss relative to the sampled vector. Source coding and decoding then preserve the quantized symbols exactly.

References

[1] Stephane Mallat. *A wavelet tour of signal processing: the sparse way*. Academic press, 2008.