UJEP

Department of Mathematics

Mathematics for Data Engineers

KMA/MDI

Part IV

Function series and
orthogonal signal expansions

Lecture 10

Yaroslav Bazaikin

October 2026

Preface

This part of *Mathematics for Data Engineers*, KMA/MDI, contains Lecture 10: function series, Fourier coefficients, and Haar expansions. It continues Parts I–III. The finite projection problem is unchanged; completeness determines whether increasingly rich models recover the whole signal in integral norm.

We assume the integral inner product and projection arguments of Part III. Six review questions and eight exercises follow the lecture; the last two exercises, marked with an asterisk, extend the core.

The task throughout this lecture

Represent a known signal by finitely many coefficients. Specify which basis functions are allowed, minimize integral squared error, and quantify the information lost by truncation. A frequency cutoff, a cell resolution, and a number of retained coefficients impose different constraints.

Starting point and notation

For real signals on $I = [a, b]$, with $a < b$, use

$$\langle f, g\rangle = \int_I f(t)g(t)\,\mathrm{d}t, \qquad \|f\|_2^2 = \int_I |f(t)|^2\,\mathrm{d}t, \qquad \mathrm{RMSE}(f, p) = \frac{\|f - p\|_2}{\sqrt{|I|}}.$$

The residual is $f - p$; squared error has no factor $1/2$. Signal energy means squared integral norm. For $f \neq 0$, relative RMS error is $\|f - p\|_2 \,/\, \|f\|_2$; it is dimensionless. RMSE has the same units as the signal.

Elements of $L^2(I)$ are classes of square-integrable measurable functions, identified by equality almost everywhere. Integral coefficients and errors depend only on the class. A pointwise statement additionally refers to a chosen representative. The inner product exists for L^2 functions by Cauchy–Schwarz.

An *orthonormal system* (q_m) satisfies $\langle q_m, q_k\rangle = \delta_{mk}$, where δ_{mk} is one for $m = k$ and zero otherwise. Basis functions are fixed; their coefficients are the unknowns in an approximation problem. Once a target f is given, projection computes those coefficients from f.

Fourier calculations use dimensionless phase $t \in [-\pi, \pi]$ and

$$S_N f(t) = \frac{a_0}{2} + \sum_{n=1}^{N} (a_n \cos nt + b_n \sin nt).$$

The *cutoff* N is the highest retained frequency index. The model space has dimension $2N + 1$; a particular target can have zero coefficients. Haar calculations use normalized time $t \in [0, 1]$: $H_J f$ is constant on 2^J equal cells. A level j and a location k specify a Haar detail.

Contents

Preface **i**

Starting point and notation **ii**

10 Fourier and Haar series **1**

10.1 Partial sums and two meanings of convergence 1
10.2 Orthogonal coordinates, optimality, and energy 1
10.3 Fourier coefficients separate frequencies 2
10.4 A complete Fourier calculation and its convergence 3
10.5 An error budget chooses the truncation 5
10.6 Haar coordinates describe local averages and details 6
10.7 Refinement, convergence, and a jump inside a cell 8
10.8 Further perspective: coefficient selection and physical frequency 10
10.9 Questions and exercises 12

From series to integral transforms **13**

Lecture 10

Fourier and Haar series

A finite expansion solves an approximation problem in a prescribed space. A series describes the limit as these spaces grow. The limiting claim must specify its meaning of convergence.

10.1 Partial sums and two meanings of convergence

For fixed functions $q_0, q_1, \ldots$ and scalars $c_0, c_1, \ldots$, the *partial sums* of $\sum_{m=0}^{\infty} c_m q_m$ are

$$P_M(t) = \sum_{m=0}^{M} c_m q_m(t).$$

Writing an infinite expansion asserts a limit of these finite functions.

Pointwise and mean-square convergence

The sequence P_M converges *pointwise* to f on I if $P_M(t) \to f(t)$ for every fixed $t \in I$. For $f, P_M \in L^2(I)$, convergence *in* L^2 means

$$\|f - P_M\|_2^2 = \int_I |f(t) - P_M(t)|^2 \, \mathrm{d}t \longrightarrow 0.$$

On a finite interval this is equivalent to mean squared error, and hence RMSE, tending to zero. It does not prescribe values on a set of measure zero.

Lecture 9's narrowing triangular pulses converge to zero in L^2 while their central value remains one. Conversely, on $[0, 1]$ define $h_n(t) = \sqrt{n}$ for $0 < t < 1/n$ and zero elsewhere. Then $h_n(t) \to 0$ at every fixed point, including zero, but

$$\|h_n\|_2^2 = n \cdot \frac{1}{n} = 1.$$

The width shrinks while the amplitude grows. Thus pointwise convergence alone does not control the integrated squared error, and L^2 convergence alone does not imply convergence at every point.

For bounded representatives, *uniform convergence* means $\sup_{t \in I} |f(t) - P_M(t)| \to 0$. It implies L^2 convergence through

$$\|f - P_M\|_2^2 \leq |I| \sup_{t \in I} |f(t) - P_M(t)|^2.$$

A requirement to preserve a local peak or a transition therefore needs more than an RMSE specification.

10.2 Orthogonal coordinates, optimality, and energy

Let $(q_m)_{m \geq 0}$ be orthonormal in real $L^2(I)$, let $f \in L^2(I)$, and set $W_M = \operatorname{span}\{q_0, \ldots, q_M\}$. The problem is

$$\min_{z_0, \ldots, z_M \in \mathbb{R}} \left\| f - \sum_{m=0}^{M} z_m q_m \right\|_2^2.$$

Its coefficients and reconstructed function are

$$c_m = \langle f, q_m \rangle , \qquad P_M f = \sum_{m=0}^{M} c_m q_m .$$

Indeed, $\langle f - P_M f, q_k \rangle = c_k - c_k = 0$ for every $k \le M$.

Reasoning: orthogonality certifies the minimum. For any $v \in W_M$, the two terms in $f - v = (f - P_M f) + (P_M f - v)$ are orthogonal. Therefore

$$\|f - v\|_2^2 = \|f - P_M f\|_2^2 + \|P_M f - v\|_2^2 . \tag{10.1}$$

The last term is nonnegative and vanishes exactly when $v = P_M f$ in L^2. Linear independence makes its coefficients unique as well. □

Taking $v = 0$ and using orthonormality gives

$$\|f - P_M f\|_2^2 = \|f\|_2^2 - \sum_{m=0}^{M} |c_m|^2 . \tag{10.2}$$

Adding a coordinate removes exactly $|c_m|^2$ of squared error. The partial sums of these nonnegative energies are bounded by $\|f\|_2^2$, so $\sum_{m=0}^{\infty} |c_m|^2 \le \|f\|_2^2$: *Bessel's inequality.*

An orthonormal system is *complete* if its finite linear combinations are dense in $L^2(I)$: for each f and $\varepsilon > 0$, some finite combination v satisfies $\|f - v\|_2 < \varepsilon$. Once W_M contains this v, optimality gives $\|f - P_M f\|_2 < \varepsilon$. Thus completeness implies $P_M f \to f$ in L^2. Letting $M \to \infty$ in (10.2) then yields *Parseval's identity*:

$$\|f\|_2^2 = \sum_{m=0}^{\infty} |c_m|^2 . \tag{10.3}$$

Orthogonality alone is insufficient: the normalized constant is orthonormal by itself, but cannot approximate a nonzero mean-zero signal arbitrarily well.

10.3 Fourier coefficients separate frequencies

On $[-\pi, \pi]$ the functions

$$q_0 = \frac{1}{\sqrt{2\pi}}, \qquad q_n^c = \frac{\cos nt}{\sqrt{\pi}}, \qquad q_n^s = \frac{\sin nt}{\sqrt{\pi}}, \qquad n \ge 1,$$

are orthonormal. For instance, $2 \cos nt \cos mt = \cos((n-m)t) + \cos((n+m)t)$. Every nonconstant integer cosine integrates to zero over this interval, giving zero when $n \ne m$ and squared norm π when $n = m$. The sine identities give the same result; sine–cosine products are odd.

Fix N and minimize $\|f - p\|_2^2$ over $p \in W_N^{\mathrm{F}} = \operatorname{span}\{1, \cos t, \sin t, \dots, \cos Nt, \sin Nt\}$. Projection gives $p = S_N f$, with

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt \, \mathrm{d}t \quad (n \ge 0), \qquad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt \, \mathrm{d}t \quad (n \ge 1). \tag{10.4}$$

The multiplier of 1 is its projection coefficient $(2\pi)^{-1} \int f = a_0/2$, the mean. Normalized coordinates are $c_0 = \sqrt{\pi/2}\, a_0$, $c_n^c = \sqrt{\pi}\, a_n$, and $c_n^s = \sqrt{\pi}\, b_n$. Consequently

$$\|f - S_N f\|_2^2 = \|f\|_2^2 - \left[\frac{\pi a_0^2}{2} + \pi \sum_{n=1}^{N} (a_n^2 + b_n^2) \right] . \tag{10.5}$$

A *harmonic* of index n makes n cycles per period. Its sine and cosine coefficients specify its phase and amplitude; its contribution to energy is $\pi(a_n^2 + b_n^2)$. A frequency cutoff preserves the mean and restricts the model to these harmonics.

Example 10.1. A compact model of cyclic latency

During a repeatable processing cycle, a modeled latency deviation from a fixed reference is $f(t) = 2 + 3\cos t - \sin 2t$ milliseconds, with phase $t \in [-\pi, \pi]$. This means that $L(t) = L_{\text{ref}} + f(t)$, negative deviations mean latency below the reference. Store only a mean and a first harmonic, minimizing squared error over the cycle.

Orthogonality in (10.4) gives

$$a_0 = \frac{1}{\pi}(2 \cdot 2\pi) = 4, \qquad a_1 = \frac{1}{\pi}(3 \cdot \pi) = 3, \qquad b_1 = 0, \qquad b_2 = -1.$$

All other coefficients vanish. Hence the optimal model is $S_1 f = 2 + 3\cos t$; its residual is $-\sin 2t$. Directly,

$$\|f\|_2^2 = 8\pi + 9\pi + \pi = 18\pi, \qquad \|f - S_1 f\|_2^2 = \pi, \qquad \text{RMSE}(f, S_1 f) = \frac{1}{\sqrt{2}} \text{ ms.}$$

The mean deviation stays 2 ms. Retained energy is $17/18$ of the total, while relative RMS error is $1/\sqrt{18}$. Allowing the second harmonic reproduces the model exactly. These statements concern the specified trace; they do not certify an error relative to unknown measurements.

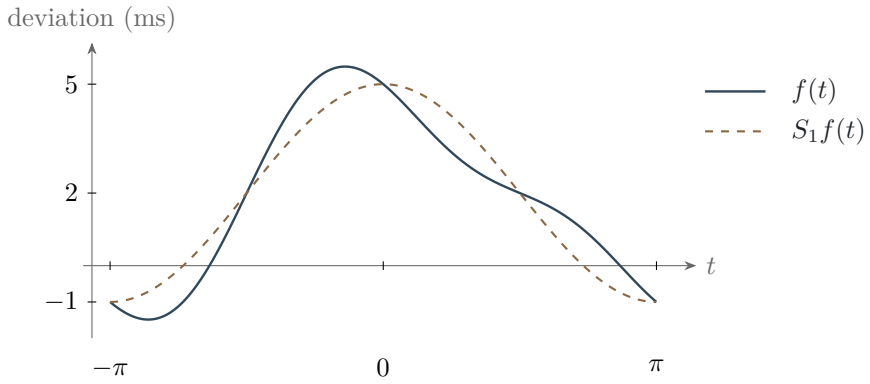


FIGURE 10.1. The latency-deviation model and its best first-harmonic approximation. The mean stays 2 ms; the discarded second harmonic contributes RMSE $1/\sqrt{2}$ ms.

10.4 A complete Fourier calculation and its convergence

Example 10.2. Encoding a repeated ramp

A synthetic calibration trace rises linearly during a device cycle and resets at its boundary. In the normalized phase and output units of Lecture 9, it is $s(t) = t$ on $(-\pi, \pi)$, repeated with period 2π. We seek its optimal model for any frequency cutoff N. Assign value zero at each reset; this fixes a representative without changing integral coefficients.

Oddness gives $a_0 = a_n = 0$. For $n \geq 1$, integration by parts gives

$$\begin{aligned} b_n &= \frac{2}{\pi} \int_0^\pi t \sin nt \, \mathrm{d}t \\ &= \frac{2}{\pi} \left[-\frac{t \cos nt}{n} + \frac{\sin nt}{n^2} \right]_0^\pi = \frac{2(-1)^{n+1}}{n}. \end{aligned}$$

Thus

$$S_N s(t) = 2 \sum_{n=1}^{N} \frac{(-1)^{n+1}}{n} \sin nt. \tag{10.6}$$

At $N = 1$ this is $2\sin t$, exactly the projection in Lecture 9. Each added harmonic reduces the integral error, but the reset requires a separate convergence statement.

Fourier convergence: two precise assertions

The trigonometric system is complete in $L^2([-\pi, \pi])$. Therefore, for every $f \in L^2$,

$$\|f - S_N f\|_2 \to 0, \qquad \|f\|_2^2 = \frac{\pi a_0^2}{2} + \pi \sum_{n=1}^{\infty} (a_n^2 + b_n^2).$$

Completeness is used without proof.

For a pointwise assertion, suppose the periodic representative is *piecewise* C^1: a period has finitely many pieces on each of which the function and its derivative extend continuously to the two ends. Then

$$S_N f(t) \longrightarrow \frac{f(t-) + f(t+)}{2} \quad \text{at every } t.$$

Both limits belong to the periodic extension. This sufficient theorem is stated without its technical proof. At a continuity point the limit is $f(t)$; at a jump it is the midpoint of the one-sided limits.

For the ramp, these limits equal t inside $(-\pi, \pi)$. At the reset $t = \pi$ they are π and $-\pi$, so the series converges to zero. In fact, $S_N s(\pm\pi) = 0$ for every N. Hence

$$t = 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin nt, \qquad -\pi < t < \pi,$$

is a justified pointwise identity. Assigning different values at the endpoints leaves the coefficients and L^2 limit unchanged, but need not give pointwise recovery of those assigned values.

Uniform recovery fails even with value zero assigned at the reset: for fixed N, $s(t) \to \pi$ and $S_N s(t) \to 0$ as $t \to \pi^-$. Thus the supremum error is at least π. Nearby oscillatory overshoot is the *Gibbs phenomenon*. Increasing N narrows the affected region while the maximum error stays nonzero; an RMSE tolerance alone cannot certify accurate reset values.

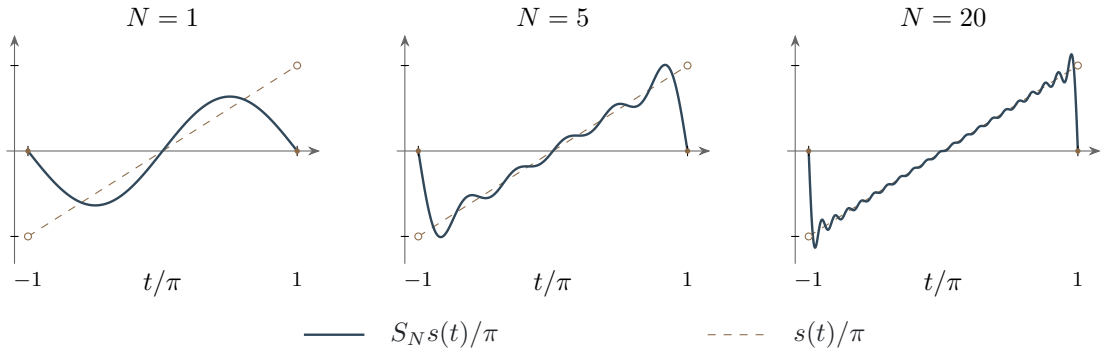


FIGURE 10.2. The normalized partial sums $S_N s(t)/\pi$ (solid) and the ramp $s(t)/\pi$ (dashed), against t/π. All three panels use the same axes, with vertical ticks at ± 1. Endpoint circles distinguish the ramp's one-sided limits from its chosen value zero at the periodic jump. The oscillations near the boundary are computed from (10.6).

10.5 An error budget chooses the truncation

For the ramp, $\|s\|_2^2 = \int_{-\pi}^{\pi} t^2 \,\mathrm{d}t = 2\pi^3/3$. Substitution of its coefficients in (10.5) gives

$$E_N := \|s - S_N s\|_2^2 = \frac{2\pi^3}{3} - 4\pi \sum_{n=1}^{N} \frac{1}{n^2}. \tag{10.7}$$

Completeness gives $E_N \to 0$, so Parseval yields $\sum_{n=1}^{\infty} n^{-2} = \pi^2/6$. The same error can therefore be written as the discarded tail

$$E_N = 4\pi \sum_{n=N+1}^{\infty} \frac{1}{n^2}.$$

Reasoning: coefficient decay gives a usable bound. Since x^{-2} decreases, comparison with neighbouring rectangles gives, for $N \geq 1$,

$$\int_{N+1}^{\infty} x^{-2} \,\mathrm{d}x \leq \sum_{n=N+1}^{\infty} \frac{1}{n^2} \leq \int_{N}^{\infty} x^{-2} \,\mathrm{d}x.$$

Multiply by 4π and divide by the period length 2π:

$$\frac{2}{N+1} \leq \mathrm{RMSE}(s, S_N s)^2 \leq \frac{2}{N}. \tag{10.8}$$

□

Example 10.3. Choosing a representation from an RMSE specification

An encoder must reproduce the ramp with RMSE at most 0.25 in its normalized output units. It uses the fixed models $S_N s$, so the design variable is the integer cutoff N. The bound guarantees the requirement whenever

$$\sqrt{2/N} \leq 1/4 \quad \Longleftrightarrow \quad N \geq 32.$$

For this signal, the exact error is

$$\mathrm{RMSE}(s, S_N s) = \left(\frac{\pi^2}{3} - 2 \sum_{n=1}^{N} \frac{1}{n^2} \right)^{1/2}.$$

It is about 1.136 at $N = 1$, 0.25197 at $N = 31$, and 0.24806 at $N = 32$. Since each added coefficient decreases the error, 32 is also the smallest feasible cutoff in this family. The mean and cosine coefficients are known to be zero, leaving 32 nonzero sine coefficients in the retained model.

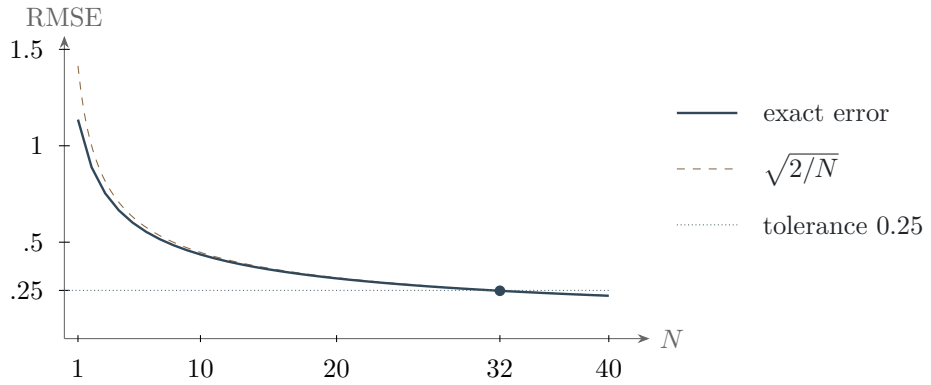


FIGURE 10.3. The exact RMSE at integer cutoffs $N = 1, \dots, 40$, joined to guide the eye, and its upper bound. The marked cutoff $N = 32$ meets the stated tolerance. Neither curve describes maximum pointwise error.

10.6 Haar coordinates describe local averages and details

To resolve changes by location and scale, set $\phi(t) = 1$ on $[0, 1]$ and define the Haar function on $\mathbb{R}$ by

$$\psi(t) = \begin{cases} 1, & 0 \le t < 1/2, \\ -1, & 1/2 \le t < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Its integral is zero and its squared integral norm is one. For integers $j \ge 0$ and $0 \le k < 2^j$, put

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k). \tag{10.9}$$

Its nonzero interval is $I_{j,k} = [k2^{-j}, (k+1)2^{-j})$, of width $h = 2^{-j}$. It is positive on the left half and negative on the right. The level j determines width; the location k chooses the cell. Its squared norm is $2^j h = 1$. Values at cell endpoints are irrelevant in L^2; use the left limit for reconstructions at $t = 1$.

Reasoning: local cancellation gives orthogonality. At one level, different nonzero intervals are disjoint. At different levels, they are either disjoint or the finer cell lies in a half where the coarser function is constant. The product integral is then that constant times the finer function's zero integral. Each detail is also orthogonal to ϕ. Hence ϕ and the $\psi_{j,k}$ form an orthonormal system. □

The coordinates of a target $f \in L^2([0, 1])$ are

$$\alpha = \int_0^1 f(t)\,\mathrm{d}t, \qquad d_{j,k} = \int_0^1 f(t)\psi_{j,k}(t)\,\mathrm{d}t.$$

Let m_L, m_R be the means of f on the two halves of $I_{j,k}$. Each half has length $h/2$, so

$$d_{j,k} = 2^{j/2}\left(\frac{h}{2}m_L - \frac{h}{2}m_R\right) = 2^{-j/2-1}(m_L - m_R). \tag{10.10}$$

A *detail coefficient* is this normalized local contrast. Its sign says which half has the larger mean; its square gives the energy lost if this coordinate is omitted.

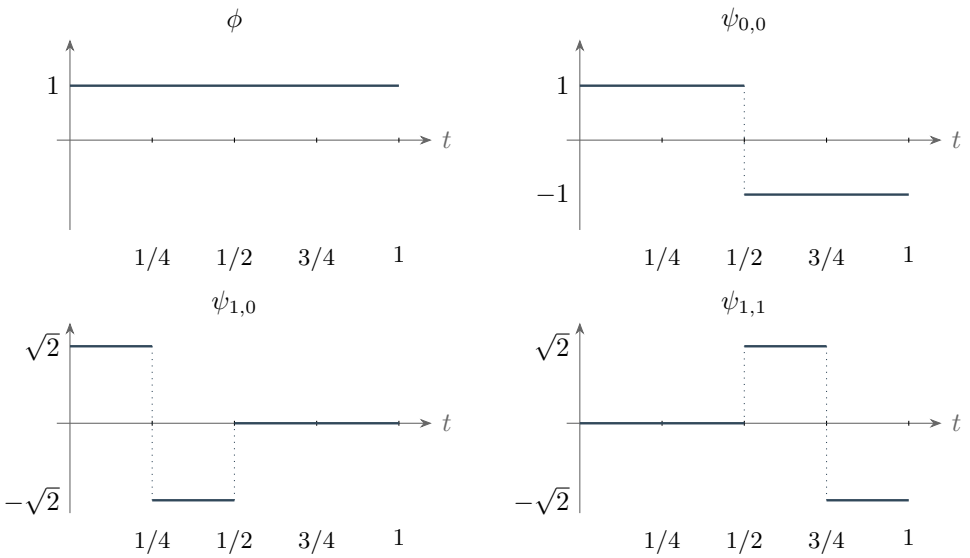


FIGURE 10.4. The first four Haar functions, with identical axis scales. Each has integral squared norm one; the last two need height $\sqrt{2}$ because their supports have length $1/2$. Dotted vertical connectors mark jumps, not intermediate signal values. Endpoint conventions do not affect the integrals.

A resolution level is a space of cell averages

Let V_J consist of functions constant on each of the 2^J cells of width 2^{-J}, interpreted in L^2. The functions ϕ and $\psi_{j,k}$ with $0 \le j < J$ belong to V_J and are orthonormal. Their number is

$$1 + \sum_{j=0}^{J-1} 2^j = 2^J = \dim V_J,$$

so they are a basis. Their projection is

$$H_J f = \alpha\phi + \sum_{j=0}^{J-1} \sum_{k=0}^{2^j-1} d_{j,k}\psi_{j,k}, \qquad H_0 f = \alpha\phi. \tag{10.11}$$

It minimizes $\|f - p\|_2^2$ over $p \in V_J$. To identify its values directly, let I be a cell of length h and $m = h^{-1}\int_I f$. Since $\int_I (f - m) = 0$,

$$\int_I (f - c)^2 \,\mathrm{d}t = \int_I (f - m)^2 \,\mathrm{d}t + h(c - m)^2.$$

Thus the unique best cell value is its mean. This both explains the reconstruction and proves that it preserves the integral on each cell.

Example 10.4. Compressing an ingestion-rate record

During a four-second window, an ingestion rate is held at $(1, 3, 2, 6)$ MB/s on successive one-second intervals. These reuse Lecture 4's values with a stated model between samples. With physical time τ and normalized time $t = \tau/(4\,\mathrm{s})$, let $g(t)$ be this four-cell step signal. Store two coefficients so that reconstruction is constant on each half-window and minimizes squared error.

The admissible space is $V_1 = \operatorname{span}\{\phi, \psi_{0,0}\}$: a function $A\phi + B\psi_{0,0}$ has levels $A + B$ and $A - B$ on the two halves. To see what is retained and lost, compute all four coordinates in V_2:

$$\begin{aligned} \alpha &= \tfrac{1}{4}(1+3+2+6) = 3, \\ d_{0,0} &= \tfrac{1}{4}(1+3-2-6) = -1, \\ d_{1,0} &= \tfrac{\sqrt{2}}{4}(1-3) = -\tfrac{1}{\sqrt{2}}, \qquad\qquad d_{1,1} = \tfrac{\sqrt{2}}{4}(2-6) = -\sqrt{2}. \end{aligned}$$

Each factor $1/4$ is a cell width; $\sqrt{2}$ is the height of a level-one basis function. These give

$$g = 3\phi - \psi_{0,0} - \tfrac{1}{\sqrt{2}}\psi_{1,0} - \sqrt{2}\psi_{1,1} \quad \text{almost everywhere.}$$

On the first quarter the sum is $3 - 1 - 1 = 1$; on the last it is $3 + 1 + 2 = 6$. Higher details vanish because the rate is constant on every quarter.

The two stored coordinates are $\alpha = 3$ and $d_{0,0} = -1$. Hence $H_1 g$ has levels $3 - 1 = 2$ and $3 + 1 = 4$ MB/s. These are the half-window averages: reconstruction preserves 4 MB in the first two seconds and 8 MB in the last two, but loses the fluctuations within each pair.

$$\|g\|_2^2 = \tfrac{1}{4}(1+9+4+36) = \frac{25}{2} = 9 + 1 + \frac{1}{2} + 2, \qquad \|g - H_1 g\|_2^2 = \frac{1}{2} + 2 = \frac{5}{2}.$$

Normalized time has interval length one, so RMSE is $\sqrt{5/2}$ MB/s. Retained energy is $10/(25/2) = 80\%$; relative RMS error is $1/\sqrt{5}$.

The relation to Lecture 4's discrete unit vectors $\boldsymbol{h}_\ell$ is exact. Each continuous basis function has cell values $2\boldsymbol{h}_\ell$, so its coefficient is

$$c_\ell = \frac{1}{4}\sum_{i=1}^{4} g_i(2h_{\ell,i}) = \frac{1}{2}\boldsymbol{h}_\ell^{\mathsf{T}}\boldsymbol{g}.$$

Thus continuous coefficients are half the discrete ones and integrated energy is one quarter of 50. Physical-time squared integrals gain the factor 4 s; division by the window length leaves the same RMSE.

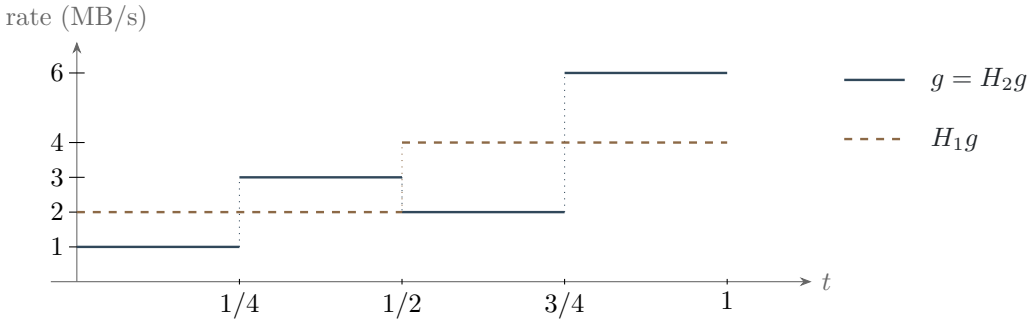


FIGURE 10.5. The ingestion rate against normalized time and its best two-cell approximation. Half-window data totals are preserved. Restoring both local details removes squared error $5/2$ in normalized time. Dotted connectors mark jumps.

10.7 Refinement, convergence, and a jump inside a cell

The Haar system is complete in $L^2([0,1])$. With partial sums taken by whole levels,

$$f = \alpha\phi + \sum_{j=0}^{\infty}\sum_{k=0}^{2^j-1} d_{j,k}\psi_{j,k} \quad \text{in } L^2,$$

and Parseval gives

$$\|f - H_J f\|_2^2 = \sum_{j=J}^{\infty} \sum_{k=0}^{2^j-1} |d_{j,k}|^2. \tag{10.12}$$

Refining from level J to $J+1$ removes exactly $\sum_k |d_{J,k}|^2$ of squared error.

For a continuous signal, its value and its cell average differ by at most its oscillation on that cell. Uniform continuity makes this bound tend to zero as the width shrinks. For a general $f \in L^2$, choose a continuous v close to f; such functions are dense, a fact used without proof. Optimality gives

$$\|f - H_J f\|_2 \le \|f - H_J v\|_2 \le \|f - v\|_2 + \|v - H_J v\|_2 .$$

Both terms can be made small, proving convergence from this density fact. It does not assert convergence at every point of a chosen representative.

Example 10.5. A service transition inside a reporting cell

A service is inactive before one third of a reporting window and active afterwards. In normalized time its state is $w(t) = 0$ for $0 \le t < 1/3$ and $w(t) = 1$ for $1/3 \le t \le 1$. A dashboard stores mean state on 2^J equal time cells. Each stored value is the fraction of its cell spent active. We seek the error and the behaviour at the transition.

Only the cell $[a, a+h)$ containing $1/3$ is inexact, with $h = 2^{-J}$. Put

$$\theta = \frac{1/3 - a}{h}.$$

The inactive part has length $h\theta$, and the active part length $h(1-\theta)$. The reconstructed cell value is therefore $1 - \theta$. Adding the two contributions to squared error gives

$$\begin{aligned} \|w - H_J w\|_2^2 &= h\theta\,[0 - (1-\theta)]^2 + h(1-\theta)\,[1 - (1-\theta)]^2 \\ &= h\theta(1-\theta) \le h/4. \end{aligned}$$

All other cells are reproduced exactly.

The fractional part of $2^J/3$ is θ: it is $1/3$ for even J and $2/3$ for odd J. Hence

$$\|w - H_J w\|_2^2 = \frac{2}{9\,2^J} \longrightarrow 0. \tag{10.13}$$

Yet the displayed value at the exact transition alternates between $2/3$ and $1/3$. There is no pointwise limit there. At any other point, sufficiently fine cells stay on one side of the transition and recover the state exactly. Finer reporting reduces integrated distortion, but a cell occupancy fraction does not specify the instantaneous binary state at the transition.

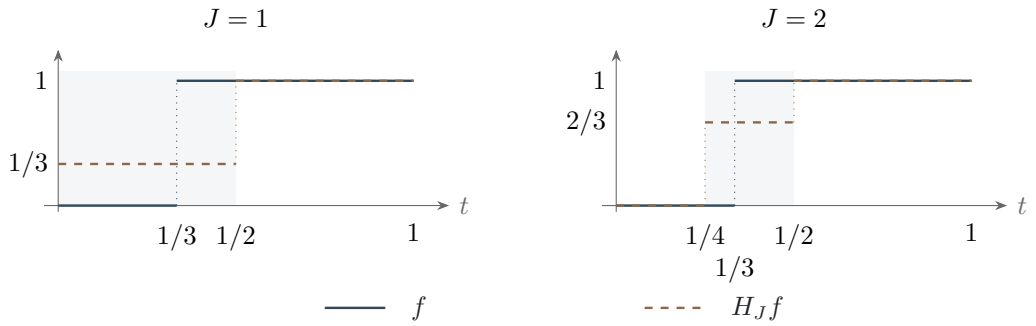


FIGURE 10.6. A fixed jump (solid) and its cell averages (dashed). Only the shaded cell contributes to the error. From $J = 1$ to $J = 2$, its width halves and the squared error falls from $1/9$ to $1/18$, even though the approximation at the jump changes from $1/3$ to $2/3$.

10.8 Further perspective: coefficient selection and physical frequency

Choosing a set of coordinates

For a fixed orthonormal system and a finite index set A, let $P_A f = \sum_{m \in A} c_m q_m$, where $c_m = \langle f, q_m \rangle$. Then

$$\|f - P_A f\|_2^2 = \|f\|_2^2 - \sum_{m \in A} |c_m|^2.$$

Among sets of K indices, a set of K largest coefficient magnitudes minimizes error: exchanging a smaller retained magnitude for a larger discarded one reduces it. Ties can give several optimal sets. If the mean must be retained, fix that coordinate first and apply the rule to the remaining budget. The count concerns coefficients; selecting arbitrary indices also requires recording those indices.

Example 10.6. Spend one detail on resolution or on a local contrast?

Compress the ingestion record from Example 10.4 with its mean and one detail. This time only the global mean must be preserved; half-window averages are not prescribed. The available detail energies are

$$|d_{0,0}|^2 = 1, \qquad |d_{1,0}|^2 = \frac{1}{2}, \qquad |d_{1,1}|^2 = 2.$$

Keeping the largest gives

$$\widetilde{g} = 3\phi - \sqrt{2}\psi_{1,1}, \qquad (\widetilde{g}_1, \widetilde{g}_2, \widetilde{g}_3, \widetilde{g}_4) = (3, 3, 1, 5),$$

and discarded energy $\|g - \widetilde{g}\|_2^2 = 1 + 1/2 = 3/2$. Its RMSE is $\sqrt{3/2}$ MB/s, below $\sqrt{5/2}$ for $H_1 g$. The model spends its detail on the larger change in the second half. It preserves the global mean 3, but both half-window means now equal 3, rather than 2 and 4.

Both models retain two coefficients. Their different errors reflect different admissible spaces: $\widetilde{g} \notin V_1$. A smaller distortion here does not meet the earlier requirement to preserve each half-window total.

Complex coefficients use the same geometry

For complex functions use $\langle f, g\rangle = \int_{-\pi}^{\pi} f\overline{g}$, linear in its first argument as in Parts I and III. The orthonormal modes are $e_n(t) = e^{int}/\sqrt{2\pi}$, $n \in \mathbb{Z}$. For unnormalized modes write

$$\widehat{f}_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,\mathrm{d}t, \qquad S_N f(t) = \sum_{n=-N}^{N} \widehat{f}_n e^{int}. \tag{10.14}$$

Conjugation supplies the minus sign. Normalized coordinates are $\gamma_n = \sqrt{2\pi}\,\widehat{f}_n$, so

$$\|f\|_2^2 = \sum_{n\in\mathbb{Z}} |\gamma_n|^2 = 2\pi \sum_{n\in\mathbb{Z}} |\widehat{f}_n|^2.$$

For real f, Euler's formula gives

$$\widehat{f}_0 = \frac{a_0}{2}, \qquad \widehat{f}_n = \frac{a_n - ib_n}{2}, \qquad \widehat{f}_{-n} = \frac{a_n + ib_n}{2} = \overline{\widehat{f}_n} \quad (n \geq 1).$$

The negative indices provide the conjugate terms that make reconstruction real. For the latency trace, $\widehat{f}_0 = 2$, $\widehat{f}_{\pm 1} = 3/2$, and $\widehat{f}_2 = i/2$, $\widehat{f}_{-2} = -i/2$.

A physical period specifies the frequency grid

For physical time τ, a period $T > 0$, and $x \in L^2([0, T])$, put $\omega_0 = 2\pi/T$. Then

$$\widehat{x}_n = \frac{1}{T}\int_0^T x(\tau)e^{-in\omega_0\tau}\,\mathrm{d}\tau, \qquad x = \sum_{n\in\mathbb{Z}} \widehat{x}_n e^{in\omega_0\tau} \quad \text{in } L^2 \text{ over one period.}$$

The nth positive harmonic has frequency n/T cycles per time unit and angular frequency $2\pi n/T$. For a four-second cycle, the first and third are 0.25 and 0.75 Hz. Division by T gives the coefficients the signal's amplitude units, while

$$\int_0^T |x(\tau)|^2\,\mathrm{d}\tau = T\sum_n |\widehat{x}_n|^2, \qquad x_{\mathrm{RMS}}^2 = \sum_n |\widehat{x}_n|^2.$$

A Fourier series represents a periodic extension on a discrete frequency grid. Part V instead defines the Fourier transform of a nonperiodic signal on the whole real line, with a continuous frequency parameter and different normalization.

10.9 Questions and exercises

Questions for review

T10.1. Define a partial sum, pointwise convergence, and convergence in L^2. Use narrowing pulses to explain why neither of the last two implies the other in general.

T10.2. State the finite projection problem and prove its optimality and energy identity. Define completeness and explain why it changes Bessel's inequality into Parseval's identity.

T10.3. Derive the real Fourier coefficients from projection. Explain $a_0/2$, the factors π in energy, and the dimension $2N + 1$ of the cutoff space.

T10.4. State a sufficient pointwise Fourier convergence theorem. Apply it at a ramp's reset, and derive an RMSE bound from its coefficient decay.

T10.5. Define the Haar indices and normalization. Derive the difference-of-means formula and explain why a complete resolution level gives cell averages.

T10.6. Compare a frequency cutoff, a Haar resolution, and selection of a fixed number of coordinates. Which constraints and error guarantees does each impose?

Exercises

10.1. A modeled periodic sensor deviation is $f(t) = 1 + 2\cos 2t - 3\sin t$ on $[-\pi, \pi]$. Find its nonzero Fourier coefficients, $S_1 f$, $S_2 f$, and total energy. Calculate the squared error and RMSE of the first-harmonic model.

10.2. For the repeated ramp $s(t) = t$ on $(-\pi, \pi)$, find $S_2 s$ and its squared integral error. Give the series limit inside the interval and at both endpoints. Explain the effect of reassigning the endpoint values.

10.3. Choose a cutoff guaranteeing ramp RMSE at most 0.1 using (10.8). How many nonzero sine coefficients are retained? Does this ensure maximum error at most 0.1 for the representative $s(t) = t$ on the closed interval?

10.4. Hold the values $(2, 0, 4, 2)$ on the quarters of $[0, 1]$. Compute all four Haar coefficients and reconstruct the signal. Find $H_1 f$ and its squared error; verify the energy identity with the cell width included.

10.5. For $f(t) = t$ on $[0, 1]$, calculate $\alpha, d_{0,0}, d_{1,0}, d_{1,1}$. Find $H_1 f$ from cell means and integrate its squared residual. Compare with the error of $H_0 f$.

10.6. A binary state changes from zero to one at $t = 3/8$. Find the cell values and squared errors of $H_1 w$ and $H_2 w$. Determine the first level giving exact recovery almost everywhere and explain why.

10.7. (*) Compute the Fourier coefficients of $f(t) = |t|$ on $[-\pi, \pi]$. State the pointwise limit, including at periodic endpoints. Use coefficient decay to show $|f(t) - S_N f(t)| \le 4/(\pi N)$ for all t and $N \ge 1$. Explain the difference from the ramp's boundary behaviour.

10.8. (*) Convert Exercise 10.1 to complex coefficients and check Parseval. For $x(\tau) = 1 + 2\cos(\pi\tau) - 3\sin(\pi\tau/2)$, with τ in seconds, give a common period and the positive frequencies in Hz.

From series to integral transforms

Both expansions minimize squared error by orthogonal projection. Their basis functions encode different information.

	Fourier series	Haar series
Basis function	A global sine, cosine, or complex oscillation	A local positive–negative contrast
Index	Frequency; phase is encoded in coefficients	Scale and position
Finite model	Frequencies up to N	Constants on 2^J cells
Optimality	Least squared error in W_N^{F}	Least squared error in V_J
Lost energy	Discarded frequency coordinates	Discarded local details
At a jump	Midpoint limit under the stated smoothness condition	Cell averages; the value at the jump can fail to converge

The coefficients are determined by the target and the chosen basis; discarding them imposes an approximation. A low integrated error may coexist with a large local discrepancy. Selecting the largest coefficients optimizes a coefficient budget, while fixed frequency or resolution limits impose additional structure.

Part V uses the same Fourier sign convention: the forward kernel is $e^{-i\omega t}$ and the inverse has factor $1/(2\pi)$. Its transform has units of amplitude times time, whereas Fourier-series coefficients have amplitude units. The distinction follows from integration over the whole line versus averaging over one period.

Haar indices correspond to scale $a = 2^{-j}$ and left endpoint $b = k2^{-j}$ in the normalized function $a^{-1/2}\psi((t-b)/a)$. This prepares wavelet analysis by scale and position. Allowing arbitrary scales and shifts in Part V gives a different family; orthogonality of the dyadic system is not an assertion about all those functions.