

Review Session 2

Sangmyung Ha

Fall 2026

1 Basic Theory of Hilbert Spaces

1

We first introduce some terminology. Throughout, the inner product on $\mathcal{H}$ is denoted by $\langle \cdot, \cdot \rangle$, and the induced norm is $\|x\|_{\mathcal{H}} = \langle x, x \rangle^{1/2}$. For complex inner product spaces, we use the convention that the inner product is linear in its first argument and conjugate-linear in its second argument.

1. An inner product space $\mathcal{H}$ is a *Hilbert space* if it is complete under the induced norm: every Cauchy sequence in $\mathcal{H}$ converges in norm to an element of $\mathcal{H}$.
2. A family $\{e_{\alpha} : \alpha \in I\} \subset \mathcal{H}$ is *orthonormal* if $\langle e_{\alpha}, e_{\beta} \rangle = 0$ for $\alpha \neq \beta$ and $\|e_{\alpha}\|_{\mathcal{H}} = 1$ for every $\alpha \in I$. An orthonormal family is an *orthonormal basis* if

$$\overline{\text{span}\{e_{\alpha} : \alpha \in I\}} = \mathcal{H}.$$

Here, span denotes the set of finite linear combinations, and the closure is taken in the norm of $\mathcal{H}$. Thus, the basis condition means that every element of $\mathcal{H}$ can be approximated arbitrarily well by finite linear combinations of the basis vectors.

3. A Hilbert space is *separable* if it contains a countable dense subset. Equivalently, it has an orthonormal basis that is finite or countably infinite.

The following theorem gives the expansion of an arbitrary element in terms of an orthonormal basis.

Theorem 1 (Orthonormal expansion and Parseval's identity). *Let $\mathcal{H}$ be a separable Hilbert space with an orthonormal basis $\{e_n : n \in J\}$, where $J = \{1, \dots, d\}$ or $J = \mathbb{N}$. Then every $x \in \mathcal{H}$ has the unique expansion*

$$x = \sum_{n \in J} \langle x, e_n \rangle e_n.$$

When $J = \mathbb{N}$, the series converges in the norm of $\mathcal{H}$, meaning that

$$\left\| x - \sum_{n=1}^N \langle x, e_n \rangle e_n \right\|_{\mathcal{H}} \longrightarrow 0 \quad \text{as } N \rightarrow \infty.$$

¹For a more detailed treatment, see the slides *Basic Theory of Hilbert Spaces* in the Canvas file *Supplementary Materials*.

Moreover, Parseval's identity holds:

$$\|x\|_{\mathcal{H}}^2 = \sum_{n \in J} |\langle x, e_n \rangle|^2.$$

When J is finite, both expansions are finite sums.

Proof. For $N \in J$, define

$$\mathcal{H}_N = \text{span}\{e_1, \dots, e_N\}, \quad S_N x = \sum_{n=1}^N \langle x, e_n \rangle e_n.$$

Orthonormality implies that, for every $j \leq N$,

$$\langle x - S_N x, e_j \rangle = \langle x, e_j \rangle - \sum_{n=1}^N \langle x, e_n \rangle \langle e_n, e_j \rangle = 0.$$

Hence $x - S_N x$ is orthogonal to $\mathcal{H}_N$. For any $y \in \mathcal{H}_N$, the decomposition

$$x - y = (x - S_N x) + (S_N x - y)$$

is therefore orthogonal. By the Pythagorean identity,

$$\|x - y\|_{\mathcal{H}}^2 = \|x - S_N x\|_{\mathcal{H}}^2 + \|S_N x - y\|_{\mathcal{H}}^2.$$

In particular,

$$\|x - S_N x\|_{\mathcal{H}} \leq \|x - y\|_{\mathcal{H}}.$$

Thus, $S_N x$ is the unique best approximation to x in $\mathcal{H}_N$, or equivalently, the orthogonal projection of x onto $\mathcal{H}_N$.

Suppose first that $J = \mathbb{N}$. Fix $\varepsilon > 0$. Since the linear span of the basis vectors is dense in $\mathcal{H}$, there exist an integer M and an element $y \in \mathcal{H}_M$ such that $\|x - y\|_{\mathcal{H}} < \varepsilon$. For every $N \geq M$, we have $y \in \mathcal{H}_N$, and therefore

$$\|x - S_N x\|_{\mathcal{H}} \leq \|x - y\|_{\mathcal{H}} < \varepsilon.$$

This proves that $S_N x \rightarrow x$ in norm. If $J = \{1, \dots, d\}$, then $\mathcal{H}_d = \mathcal{H}$, so $x - S_d x$ is orthogonal to all of $\mathcal{H}$ and must be zero.

To establish uniqueness, suppose that $x = \sum_{n \in J} c_n e_n$, with norm convergence when $J = \mathbb{N}$. Taking the inner product with e_m and using continuity of the inner product gives

$$\langle x, e_m \rangle = c_m.$$

Finally, since $S_N x$ and $x - S_N x$ are orthogonal,

$$\|x\|_{\mathcal{H}}^2 = \|x - S_N x\|_{\mathcal{H}}^2 + \sum_{n=1}^N |\langle x, e_n \rangle|^2.$$

Letting $N \rightarrow \infty$ in the infinite-dimensional case, or setting $N = d$ in the finite-dimensional case, proves Parseval's identity. $\square$

1.1 Examples

1. Eigenbases and population PCA.

By the spectral theorem, every real symmetric matrix $A \in \mathbb{R}^{n \times n}$ admits an orthonormal basis of eigenvectors $v_1, \dots, v_n$. When an eigenvalue is repeated, we choose an orthonormal basis within its eigenspace. Applying Theorem 1 to $\mathbb{R}^n$ with its Euclidean inner product gives

$$x = \sum_{i=1}^n \langle x, v_i \rangle_{\mathbb{R}^n} v_i = \sum_{i=1}^n (v_i' x) v_i \quad \text{for every } x \in \mathbb{R}^n.$$

Now let X_t be a centered random vector in $\mathbb{R}^n$ satisfying $\mathbb{E}\|X_t\|_{\mathbb{R}^n}^2 < \infty$, and take

$$A = \text{Cov}(X_t) = \mathbb{E}[X_t X_t'].$$

Order its eigenvalues and corresponding orthonormal eigenvectors so that

$$A v_i = \lambda_i v_i, \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0.$$

Applying the Euclidean expansion to each realization gives

$$X_t(\omega) = \sum_{i=1}^n (v_i' X_t(\omega)) v_i.$$

The vectors v_i are deterministic elements of $\mathbb{R}^n$, while the coefficients $\xi_{it} = v_i' X_t$ are scalar random variables. This is a change of basis in $\mathbb{R}^n$ for each realization of X_t , not a change of basis in the space $L^2(\Omega, \mathcal{F}, \mathbb{P})$. The random coefficients satisfy

$$\mathbb{E}[\xi_{it}] = 0, \quad \mathbb{E}[\xi_{it} \xi_{jt}] = v_i' A v_j = \begin{cases} \lambda_i, & i = j, \\ 0, & i \neq j. \end{cases}$$

In particular, the random coefficients are uncorrelated, with variances given by the eigenvalues.

For $1 \leq r < n$, let $V_r = (v_1, \dots, v_r)$ and define the truncated expansion

$$\tilde{X}_t = V_r V_r' X_t = \sum_{i=1}^r (v_i' X_t) v_i.$$

This is the projection of X_t onto the subspace generated by the eigenvectors associated with the r largest eigenvalues. It minimizes expected squared Euclidean reconstruction error among projections onto deterministic r -dimensional linear subspaces:

$$\begin{aligned} \mathbb{E}\|X_t - \tilde{X}_t\|_{\mathbb{R}^n}^2 &= \min_{\substack{Q \in \mathbb{R}^{n \times r} \\ Q'Q = I_r}} \mathbb{E}\|X_t - QQ'X_t\|_{\mathbb{R}^n}^2 \\ &= \sum_{i=r+1}^n \lambda_i. \end{aligned}$$

Indeed, for any $Q \in \mathbb{R}^{n \times r}$ with $Q'Q = I_r$,

$$\mathbb{E}\|X_t - QQ'X_t\|_{\mathbb{R}^n}^2 = \text{tr}(A) - \text{tr}(Q'AQ).$$

Writing $a_i = \|Q'v_i\|_{\mathbb{R}^r}^2$, we have $0 \leq a_i \leq 1$ and $\sum_{i=1}^n a_i = r$, so

$$\text{tr}(Q'AQ) = \sum_{i=1}^n \lambda_i a_i \leq \sum_{i=1}^r \lambda_i.$$

Equality is attained by $Q = V_r$. Thus, minimizing reconstruction error is equivalent to maximizing the variance captured by the projection.

The reconstruction criterion can also be written using scalar L^2 norms:

$$\mathbb{E}\|X_t - \tilde{X}_t\|_{\mathbb{R}^n}^2 = \sum_{j=1}^n \|X_{jt} - \tilde{X}_{jt}\|_{L^2(\Omega, \mathcal{F}, \mathbb{P})}^2.$$

This is the population version of principal component analysis, or PCA. The minimizing subspace need not be unique when $\lambda_r = \lambda_{r+1}$.

2. The Fourier basis.

Let $L^2([-\pi, \pi])$ denote the space of complex-valued square-integrable functions on $[-\pi, \pi]$, with functions that agree almost everywhere are identified. We use the normalized inner product

$$\langle f, g \rangle_{L^2} = \int_{-\pi}^{\pi} f(\lambda) \overline{g(\lambda)} \frac{d\lambda}{2\pi}, \quad \|f\|_{L^2}^2 = \int_{-\pi}^{\pi} |f(\lambda)|^2 \frac{d\lambda}{2\pi}.$$

For each $n \in \mathbb{Z}$, define

$$\varphi_n(\lambda) = e^{in\lambda}.$$

These functions form an orthonormal family. To see this,

$$\|\varphi_n\|_{L^2}^2 = \int_{-\pi}^{\pi} e^{in\lambda} e^{-in\lambda} \frac{d\lambda}{2\pi} = \int_{-\pi}^{\pi} 1 \frac{d\lambda}{2\pi} = 1.$$

For $n \neq m$,

$$\begin{aligned} \langle \varphi_n, \varphi_m \rangle_{L^2} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)\lambda} d\lambda \\ &= \frac{e^{i(n-m)\pi} - e^{-i(n-m)\pi}}{2\pi i(n-m)} \\ &= \frac{\sin((n-m)\pi)}{\pi(n-m)} \\ &= 0. \end{aligned}$$

The third equality follows from Euler's formula $e^{ix} = \cos x + i \sin x$ and because cosine is an even function. The last equality follows because $n - m$ is a nonzero integer.

A separate completeness result states that the linear span of $\{\varphi_n : n \in \mathbb{Z}\}$, consisting of trigonometric polynomials, is dense in $L^2([-\pi, \pi])$. We omit its proof. Together with orthonormality, this establishes that $\{\varphi_n : n \in \mathbb{Z}\}$ is an orthonormal basis.

For any $f \in L^2([-\pi, \pi])$, define its Fourier coefficients by

$$c_n = \langle f, \varphi_n \rangle_{L^2} = \int_{-\pi}^{\pi} f(\lambda) e^{-in\lambda} \frac{d\lambda}{2\pi}, \quad n \in \mathbb{Z}.$$

Theorem 1 gives the Fourier expansion

$$f(\lambda) = \sum_{n \in \mathbb{Z}} c_n e^{in\lambda},$$

where equality is understood in L^2 . More explicitly, the symmetric partial sums satisfy

$$\lim_{K \rightarrow \infty} \int_{-\pi}^{\pi} \left| f(\lambda) - \sum_{n=-K}^K c_n e^{in\lambda} \right|^2 \frac{d\lambda}{2\pi} = 0.$$

This is convergence in mean square over the frequency variable λ ; the Hilbert-space theorem does not assert convergence at every λ or uniform convergence. Parseval's identity gives

$$\|f\|_{L^2}^2 = \int_{-\pi}^{\pi} |f(\lambda)|^2 \frac{d\lambda}{2\pi} = \sum_{n \in \mathbb{Z}} |c_n|^2.$$

Connection with spectral densities.

Let $\{X_t : t \in \mathbb{Z}\}$ be a real-valued, centered, covariance-stationary scalar process with autocovariance function

$$\gamma_X(h) = \mathbb{E}[X_t X_{t-h}], \quad h \in \mathbb{Z}.$$

Suppose the process has a spectral density. To match the normalized inner product used above, we define f_X as the spectral density relative to the measure $d\lambda/(2\pi)$, so that

$$\gamma_X(h) = \int_{-\pi}^{\pi} f_X(\lambda) e^{ih\lambda} \frac{d\lambda}{2\pi}.$$

If, in addition, $f_X \in L^2([-\pi, \pi])$, its Fourier coefficients satisfy

$$\langle f_X, \varphi_n \rangle_{L^2} = \int_{-\pi}^{\pi} f_X(\lambda) e^{-in\lambda} \frac{d\lambda}{2\pi} = \gamma_X(-n) = \gamma_X(n).$$

The last equality uses the symmetry of the autocovariance function of a real-valued scalar process. Thus, with this normalization, the autocovariances are the Fourier coefficients of the spectral density, and

$$f_X(\lambda) = \sum_{n \in \mathbb{Z}} \gamma_X(n) e^{in\lambda} \quad \text{in } L^2([-\pi, \pi]).$$

Parseval's identity then yields

$$\int_{-\pi}^{\pi} |f_X(\lambda)|^2 \frac{d\lambda}{2\pi} = \sum_{n \in \mathbb{Z}} |\gamma_X(n)|^2.$$

A spectral density is integrable but need not be square-integrable. The covariance–spectral-density integral relation remains valid without square integrability, but the L^2 expansion and Parseval identity above require $f_X \in L^2([-\pi, \pi])$. One convenient sufficient condition is absolute summability of the autocovariances:

$$\sum_{h \in \mathbb{Z}} |\gamma_X(h)| < \infty.$$

Under this stronger condition, the spectral-density series converges absolutely and uniformly, giving a bounded continuous version of f_X and hence ensuring that $f_X \in L^2([-\pi, \pi])$.

2 Spectral Representation of a Covariance-Stationary Process

Throughout this section, time is indexed by $t \in \mathbb{Z}$. We work in the scalar complex Hilbert space $\mathcal{L}_{\mathbb{C}}^2(\Omega, \mathcal{F}, \mathbb{P})$, with inner product and norm

$$\langle U, V \rangle_{\mathcal{L}^2(\Omega)} = \mathbb{E}[U \overline{V}], \quad \|U\|_{\mathcal{L}^2(\Omega)}^2 = \mathbb{E}|U|^2.$$

Even when the process of interest is real-valued, complex-valued random variables are useful because the spectral representation involves complex exponentials. Frequencies are understood modulo 2π and represented by the half-open interval $\mathbb{T} = [-\pi, \pi)$. Thus, the equivalent frequencies $-\pi$ and π are counted only once. We use the negative-exponential convention $e^{-it\lambda}$ throughout the process representation.

A finite-frequency example. Let $\lambda_1, \dots, \lambda_n \in \mathbb{T}$ be distinct deterministic frequencies, and let $Z_1, \dots, Z_n \in \mathcal{L}_{\mathbb{C}}^2(\Omega, \mathcal{F}, \mathbb{P})$ be centered, pairwise orthogonal random variables:

$$\mathbb{E}[Z_j] = 0, \quad \mathbb{E}[Z_j \overline{Z_\ell}] = 0 \quad \text{for } j \neq \ell.$$

Define

$$X_t^{(n)} = \sum_{j=1}^n e^{-it\lambda_j} Z_j, \quad t \in \mathbb{Z}. \tag{1}$$

Since $e^{-it\lambda_j} = \cos(t\lambda_j) - i \sin(t\lambda_j)$, this process is a finite combination of oscillatory components with random coefficients.

We can show directly that $X_t^{(n)}$ is covariance stationary. First,

$$\mathbb{E}[X_t^{(n)}] = \sum_{j=1}^n e^{-it\lambda_j} \mathbb{E}[Z_j] = 0.$$

Next, writing $\sigma_j^2 = \mathbb{E}|Z_j|^2$, we obtain

$$\begin{aligned} \gamma_n(h) &:= \mathbb{E} \left[X_{t+h}^{(n)} \overline{X_t^{(n)}} \right] \\ &= \sum_{j=1}^n \sum_{\ell=1}^n e^{-i(t+h)\lambda_j} e^{it\lambda_\ell} \mathbb{E}[Z_j \overline{Z_\ell}] \\ &= \sum_{j=1}^n e^{-ih\lambda_j} \sigma_j^2. \end{aligned}$$

The last equality follows from orthogonality. The covariance depends only on the lag h , and

$$\mathbb{E}|X_t^{(n)}|^2 = \gamma_n(0) = \sum_{j=1}^n \sigma_j^2 < \infty.$$

Define a finite nonnegative measure on the frequency domain by

$$F_n(B) = \sum_{j=1}^n \sigma_j^2 \mathbf{1}_B(\lambda_j), \quad B \in \mathcal{B}(\mathbb{T}).$$

The measure F_n places mass σ_j^2 at frequency λ_j , and

$$\gamma_n(h) = \int_{\mathbb{T}} e^{-ih\lambda} F_n(d\lambda).$$

This is the spectral measure of the finite-frequency process.

Orthogonal random measures. For a general covariance-stationary process, the finite collection of random coefficients is replaced by an orthogonal random measure. Let F be a finite nonnegative Borel measure on $\mathbb{T}$. A *centered orthogonal random measure with control measure F* is a map

$$Z : \mathcal{B}(\mathbb{T}) \longrightarrow \mathcal{L}_{\mathbb{C}}^2(\Omega, \mathcal{F}, \mathbb{P})$$

such that

$$\mathbb{E}[Z(B)] = 0, \quad \mathbb{E}[Z(B)\overline{Z(C)}] = F(B \cap C), \quad B, C \in \mathcal{B}(\mathbb{T}), \quad (2)$$

and Z is countably additive in the $\mathcal{L}^2$ norm. More explicitly, for pairwise disjoint sets $B_1, B_2, \dots$,

$$\left\| Z\left(\bigcup_{j=1}^{\infty} B_j\right) - \sum_{j=1}^N Z(B_j) \right\|_{\mathcal{L}^2(\Omega)} \longrightarrow 0 \quad \text{as } N \rightarrow \infty.$$

In particular,

$$F(B) = \mathbb{E}|Z(B)|^2,$$

and disjoint frequency sets give orthogonal random variables:

$$B \cap C = \emptyset \implies \langle Z(B), Z(C) \rangle_{\mathcal{L}^2(\Omega)} = 0.$$

Thus, F is an ordinary deterministic measure, whereas Z takes values in a Hilbert space of random variables. The additivity of Z is understood in mean square, not as a statement that every realization defines an ordinary finite complex measure.

For a simple function $g(\lambda) = \sum_{j=1}^m c_j \mathbf{1}_{B_j}(\lambda)$ on disjoint Borel sets, define

$$\int_{\mathbb{T}} g(\lambda) Z(d\lambda) = \sum_{j=1}^m c_j Z(B_j).$$

The integral for a general $g \in \mathcal{L}^2(\mathbb{T}, F)$ is obtained by approximation with simple functions and convergence in mean square. We give this construction and its isometry property in the next section.

The spectral representation theorem.

Theorem 2 (Spectral representation). *Let $\{X_t : t \in \mathbb{Z}\}$ be a centered covariance-stationary process with $X_t \in \mathcal{L}_{\mathbb{C}}^2(\Omega, \mathcal{F}, \mathbb{P})$, and define*

$$\gamma(h) = \mathbb{E}[X_{t+h} \overline{X_t}], \quad h \in \mathbb{Z}.$$

There exist a unique finite nonnegative Borel measure F on $\mathbb{T}$ and a centered orthogonal random measure Z with control measure F such that

$$X_t = \int_{\mathbb{T}} e^{-it\lambda} Z(d\lambda) \quad \text{in } \mathcal{L}^2, \quad t \in \mathbb{Z}. \quad (3)$$

The same random measure Z is used for every t . Moreover,

$$\gamma(h) = \int_{\mathbb{T}} e^{-ih\lambda} F(d\lambda), \quad h \in \mathbb{Z}, \quad (4)$$

and

$$F(\mathbb{T}) = \gamma(0) = \mathbb{E}|X_t|^2.$$

The measure F is called the spectral measure of the process.

The covariance representation is the Herglotz representation, and the process representation is the Cramér spectral representation. We omit the existence proofs; see Yu. A. Rozanov, *Stationary Random Processes*, for a detailed treatment. The covariance formula will also follow directly from the isometry established in the next section.

For a covariance-stationary process with mean μ , apply the theorem to $X_t - \mu$. The representation then becomes

$$X_t = \mu + \int_{\mathbb{T}} e^{-it\lambda} Z(d\lambda).$$

In what follows, we continue to work with centered processes.

Approximation by finite sums. The stochastic integral in (3) can be approximated by sums resembling Riemann sums. For each n , choose a partition

$$-\pi = \lambda_{n,0} < \lambda_{n,1} < \cdots < \lambda_{n,n} = \pi, \quad \Lambda_{n,j} = [\lambda_{n,j-1}, \lambda_{n,j}),$$

and a point $\xi_{n,j} \in \Lambda_{n,j}$ for each j . Assume that the mesh of the partition satisfies

$$\delta_n := \max_{1 \leq j \leq n} (\lambda_{n,j} - \lambda_{n,j-1}) \longrightarrow 0.$$

Define

$$X_t^{(n)} = \sum_{j=1}^n e^{-it\xi_{n,j}} Z(\Lambda_{n,j}).$$

For each fixed n , the increments $Z(\Lambda_{n,1}), \dots, Z(\Lambda_{n,n})$ are centered and pairwise orthogonal. Thus, $X_t^{(n)}$ is a covariance-stationary process of the finite-sum form (1). The increments associated with different partitions all come from the same random measure Z .

Using the isometry (6) proved below, for every fixed $t \in \mathbb{Z}$,

$$\begin{aligned}\mathbb{E}|X_t - X_t^{(n)}|^2 &= \sum_{j=1}^n \int_{\Lambda_{n,j}} |e^{-it\lambda} - e^{-it\xi_{n,j}}|^2 F(d\lambda) \\ &\leq t^2 \delta_n^2 F(\mathbb{T}) \\ &\longrightarrow 0,\end{aligned}$$

where we used $|e^{-ita} - e^{-itb}| \leq |t| |a - b|$. Hence

$$X_t = \lim_{n \rightarrow \infty} \sum_{j=1}^n e^{-it\xi_{n,j}} Z(\Lambda_{n,j}) \quad \text{in } \mathcal{L}^2.$$

This is convergence in mean square for each fixed t . Merely increasing the number of intervals is not enough; their maximum width must tend to zero.

Similarly,

$$\mathbb{E} \left[X_{t+h}^{(n)} \overline{X_t^{(n)}} \right] = \sum_{j=1}^n e^{-ih\xi_{n,j}} F(\Lambda_{n,j}),$$

and

$$\left| \sum_{j=1}^n e^{-ih\xi_{n,j}} F(\Lambda_{n,j}) - \int_{\mathbb{T}} e^{-ih\lambda} F(d\lambda) \right| \leq |h| \delta_n F(\mathbb{T}) \longrightarrow 0.$$

The last limit is ordinary convergence of complex numbers. Unlike the integral with respect to Z , the integral with respect to F is an ordinary deterministic measure integral.

Fourier–Stieltjes interpretation. Define the cumulative spectral process

$$Z(\lambda) = Z([-\pi, \lambda)), \quad -\pi \leq \lambda \leq \pi.$$

Then

$$Z([a, b)) = Z(b) - Z(a), \quad -\pi \leq a < b \leq \pi,$$

and the representation is also written as

$$X_t = \int_{-\pi}^{\pi} e^{-it\lambda} dZ(\lambda).$$

This is a *stochastic Fourier–Stieltjes representation*, with the integral defined in mean square. It should not generally be interpreted as an ordinary Riemann–Stieltjes integral computed separately for every outcome ω . In particular, the construction does not require $\lambda \mapsto Z(\lambda, \omega)$ to have bounded variation, or require sample paths of X_t to be periodic, nonperiodic, or absolutely integrable.

For a real-valued process, the spectral measure and random measure have the conjugate symmetry

$$F(-B) = F(B), \quad Z(-B) = \overline{Z(B)},$$

where reflection is understood on the frequency circle and the second equality holds in $\mathcal{L}^2$. Contributions at opposite frequencies therefore combine to give real-valued random variables.

The spectral representation is a Hilbert-space decomposition into frequency components. The increment $Z(B)$ is the random coefficient associated with a frequency band B , and $F(B) = \mathbb{E}|Z(B)|^2$ measures the variance associated with that band.

3 Isometry between the Time and Frequency Domains

Let $\{X_t : t \in \mathbb{Z}\}$ be the centered covariance-stationary process considered above, with spectral representation

$$X_t = \int_{\mathbb{T}} e^{-it\lambda} Z(d\lambda).$$

Define its time-domain Hilbert space by

$$\mathcal{M} = \overline{\text{span}_{\mathbb{C}}\{X_t : t \in \mathbb{Z}\}}, \quad (5)$$

where the closure is taken in $\mathcal{L}_{\mathbb{C}}^2(\Omega, \mathcal{F}, \mathbb{P})$. Thus, $\mathcal{M}$ consists of finite complex linear combinations of the X_t and their mean-square limits. The closure is essential: an arbitrary element of $\mathcal{M}$ need not be a finite linear combination of the process.

The corresponding frequency-domain Hilbert space is $\mathcal{L}^2(\mathbb{T}, F)$, consisting of complex-valued measurable functions g satisfying

$$\int_{\mathbb{T}} |g(\lambda)|^2 F(d\lambda) < \infty,$$

with functions that agree F -almost everywhere identified. Its inner product is

$$\langle g, q \rangle_{\mathcal{L}^2(F)} = \int_{\mathbb{T}} g(\lambda) \overline{q(\lambda)} F(d\lambda).$$

This space depends on the spectral measure F of the process.

We construct the map

$$\tau : \mathcal{L}^2(\mathbb{T}, F) \longrightarrow \mathcal{L}_{\mathbb{C}}^2(\Omega, \mathcal{F}, \mathbb{P}), \quad \tau(g) = \int_{\mathbb{T}} g(\lambda) Z(d\lambda),$$

and then show that it is a linear isometry whose range is exactly $\mathcal{M}$.

Step 1: Define the map for simple functions. Let

$$g(\lambda) = \sum_{j=1}^m c_j \mathbf{1}_{B_j}(\lambda),$$

where $B_1, \dots, B_m$ are pairwise disjoint Borel sets. Define

$$\tau(g) = \sum_{j=1}^m c_j Z(B_j).$$

Finite additivity of Z implies that this definition does not depend on the particular disjoint-set representation of the simple function. Moreover, a set with zero F -measure has a zero increment in $\mathcal{L}^2$, since $\mathbb{E}|Z(B)|^2 = F(B)$.

Using orthogonality,

$$\begin{aligned}
\|\tau(g)\|_{\mathcal{L}^2(\Omega)}^2 &= \mathbb{E} \left| \sum_{j=1}^m c_j Z(B_j) \right|^2 \\
&= \sum_{j=1}^m \sum_{\ell=1}^m c_j \bar{c}_\ell \mathbb{E}[Z(B_j) \overline{Z(B_\ell)}] \\
&= \sum_{j=1}^m |c_j|^2 F(B_j) \\
&= \int_{\mathbb{T}} |g(\lambda)|^2 F(d\lambda) \\
&= \|g\|_{\mathcal{L}^2(F)}^2.
\end{aligned}$$

Thus, τ preserves norms on simple functions. By expressing two simple functions on a common finite partition, we also obtain

$$\|\tau(g) - \tau(q)\|_{\mathcal{L}^2(\Omega)} = \|g - q\|_{\mathcal{L}^2(F)}.$$

Step 2: Extend the map by mean-square approximation. Let $g \in \mathcal{L}^2(\mathbb{T}, F)$. Since simple functions are dense in this space, there exists a sequence of simple functions g_n such that

$$\|g_n - g\|_{\mathcal{L}^2(F)} \longrightarrow 0.$$

The approximating functions may be complex-valued; no monotonicity assumption is required.

The identity for simple functions gives

$$\|\tau(g_n) - \tau(g_m)\|_{\mathcal{L}^2(\Omega)} = \|g_n - g_m\|_{\mathcal{L}^2(F)} \longrightarrow 0 \quad \text{as } n, m \rightarrow \infty.$$

Therefore, $\{\tau(g_n)\}$ is a Cauchy sequence. Completeness of $\mathcal{L}_{\mathbb{C}}^2(\Omega, \mathcal{F}, \mathbb{P})$ implies that it has a mean-square limit. Define

$$\tau(g) := \lim_{n \rightarrow \infty} \tau(g_n) \quad \text{in } \mathcal{L}^2(\Omega).$$

This definition is independent of the approximating sequence: if $q_n \rightarrow g$ in $\mathcal{L}^2(F)$, then

$$\|\tau(g_n) - \tau(q_n)\|_{\mathcal{L}^2(\Omega)} = \|g_n - q_n\|_{\mathcal{L}^2(F)} \leq \|g_n - g\|_{\mathcal{L}^2(F)} + \|q_n - g\|_{\mathcal{L}^2(F)} \longrightarrow 0.$$

The extended map is linear, and continuity of the norm gives

$$\|\tau(g)\|_{\mathcal{L}^2(\Omega)}^2 = \int_{\mathbb{T}} |g(\lambda)|^2 F(d\lambda) = \|g\|_{\mathcal{L}^2(F)}^2. \quad (6)$$

Equivalently,

$$\mathbb{E} \left| \int_{\mathbb{T}} g(\lambda) Z(d\lambda) \right|^2 = \int_{\mathbb{T}} |g(\lambda)|^2 F(d\lambda).$$

This is the *isometry property* of stochastic integration with respect to an orthogonal random measure. The argument uses approximation in the $\mathcal{L}^2$ norm and completeness, rather than monotone convergence.

Step 3: Show that inner products are preserved. For simple functions $g = \sum_j c_j \mathbf{1}_{B_j}$ and $q = \sum_j d_j \mathbf{1}_{B_j}$ expressed on a common disjoint partition,

$$\mathbb{E}[\tau(g)\overline{\tau(q)}] = \sum_j c_j \overline{d_j} F(B_j) = \int_{\mathbb{T}} g(\lambda) \overline{q(\lambda)} F(d\lambda).$$

Approximate general $g, q \in \mathcal{L}^2(\mathbb{T}, F)$ by simple functions. Continuity of the inner products, which follows from the Cauchy–Schwarz inequality, then yields

$$\langle \tau(g), \tau(q) \rangle_{\mathcal{L}^2(\Omega)} = \langle g, q \rangle_{\mathcal{L}^2(F)}. \quad (7)$$

In integral notation,

$$\mathbb{E} \left[\left(\int_{\mathbb{T}} g(\lambda) Z(d\lambda) \right) \overline{\left(\int_{\mathbb{T}} q(\lambda) Z(d\lambda) \right)} \right] = \int_{\mathbb{T}} g(\lambda) \overline{q(\lambda)} F(d\lambda).$$

Thus, the time-domain and frequency-domain representations have the same norms and inner products.

Step 4: Identify the range of the map. For $t \in \mathbb{Z}$, define

$$\varphi_t(\lambda) = e^{-it\lambda}.$$

The spectral representation gives

$$\tau(\varphi_t) = X_t.$$

The linear span of the exponentials is dense in $\mathcal{L}^2(\mathbb{T}, F)$:

$$\overline{\text{span}_{\mathbb{C}}\{\varphi_t : t \in \mathbb{Z}\}} = \mathcal{L}^2(\mathbb{T}, F). \quad (8)$$

This density result holds for any finite Borel measure on the frequency circle. It follows from the uniform density of trigonometric polynomials in continuous functions on the circle, together with the density of continuous functions in $\mathcal{L}^2(\mathbb{T}, F)$.

It is important that density does not imply orthonormality. In fact,

$$\langle \varphi_t, \varphi_s \rangle_{\mathcal{L}^2(F)} = \int_{\mathbb{T}} e^{-i(t-s)\lambda} F(d\lambda) = \gamma(t-s),$$

which need not be zero when $t \neq s$. Thus, the exponentials are generally not an orthonormal basis under the spectral measure F .

We now show that the range of τ is exactly $\mathcal{M}$. First, given $g \in \mathcal{L}^2(\mathbb{T}, F)$, choose trigonometric polynomials

$$p_n(\lambda) = \sum_{|j| \leq N_n} a_{n,j} e^{-ij\lambda}$$

such that $\|p_n - g\|_{\mathcal{L}^2(F)} \rightarrow 0$. By linearity,

$$\tau(p_n) = \sum_{|j| \leq N_n} a_{n,j} X_j.$$

By the isometry,

$$\|\tau(p_n) - \tau(g)\|_{\mathcal{L}^2(\Omega)} = \|p_n - g\|_{\mathcal{L}^2(F)} \rightarrow 0.$$

Since $\mathcal{M}$ is closed, $\tau(g) \in \mathcal{M}$.

Conversely, let $Y \in \mathcal{M}$. There exist finite linear combinations $Y_n = \sum_{|j| \leq N_n} a_{n,j} X_j$ such that $Y_n \rightarrow Y$ in $\mathcal{L}^2(\Omega)$. Set

$$p_n(\lambda) = \sum_{|j| \leq N_n} a_{n,j} e^{-ij\lambda}.$$

Then

$$\|p_n - p_m\|_{\mathcal{L}^2(F)} = \|Y_n - Y_m\|_{\mathcal{L}^2(\Omega)} \rightarrow 0.$$

Completeness of $\mathcal{L}^2(\mathbb{T}, F)$ gives a limit g with $p_n \rightarrow g$ in $\mathcal{L}^2(F)$. Continuity of τ implies that $\tau(g) = Y$.

Finally, the isometry implies injectivity: if $\tau(g) = \tau(q)$ in $\mathcal{L}^2(\Omega)$, then $\|g - q\|_{\mathcal{L}^2(F)} = 0$, so $g = q$ as elements of $\mathcal{L}^2(\mathbb{T}, F)$. We have therefore established the linear isometric bijection

$$\tau : \mathcal{L}^2(\mathbb{T}, F) \rightarrow \mathcal{M}.$$

This is the time–frequency isometry. Its range is the closed linear span of the process, not generally the entire space $\mathcal{L}_{\mathbb{C}}^2(\Omega, \mathcal{F}, \mathbb{P})$.

3.1 Isometry in Use

Let $\alpha, \beta \in \mathcal{L}^2(\mathbb{T}, F)$, and define

$$U = \int_{\mathbb{T}} \alpha(\lambda) Z(d\lambda), \quad V = \int_{\mathbb{T}} \beta(\lambda) Z(d\lambda).$$

These are elements of $\mathcal{M}$. In particular, if

$$p_n(\lambda) = \sum_{|j| \leq N_n} a_{n,j} e^{-ij\lambda}, \quad \|p_n - \alpha\|_{\mathcal{L}^2(F)} \rightarrow 0,$$

then

$$\left\| U - \sum_{|j| \leq N_n} a_{n,j} X_j \right\|_{\mathcal{L}^2(\Omega)} = \|\alpha - p_n\|_{\mathcal{L}^2(F)} \rightarrow 0.$$

The same argument applies to V . The coefficients of the approximating polynomials may depend on n ; density alone does not supply an orthonormal-basis expansion with a fixed coefficient sequence in $\mathcal{L}^2(\mathbb{T}, F)$.

By (7),

$$\begin{aligned} \mathbb{E}[U\overline{V}] &= \langle U, V \rangle_{\mathcal{L}^2(\Omega)} \\ &= \langle \alpha, \beta \rangle_{\mathcal{L}^2(F)} \\ &= \int_{\mathbb{T}} \alpha(\lambda) \overline{\beta(\lambda)} F(d\lambda). \end{aligned}$$

In particular,

$$\mathbb{E}|U|^2 = \int_{\mathbb{T}} |\alpha(\lambda)|^2 F(d\lambda).$$

These identities follow from finite sums and passage to mean-square limits. There is no need to introduce a double stochastic integral or manipulate an expression such as $\mathbb{E}[dZ(\lambda)d\overline{Z(\lambda)}]$ as though the increments were ordinary differentials.

Recovering the autocovariance function. Apply the inner-product identity to

$$X_t = \int_{\mathbb{T}} e^{-it\lambda} Z(d\lambda), \quad X_{t-k} = \int_{\mathbb{T}} e^{-i(t-k)\lambda} Z(d\lambda).$$

Then

$$\begin{aligned} \gamma(k) &= \mathbb{E}[X_t \overline{X_{t-k}}] \\ &= \langle X_t, X_{t-k} \rangle_{\mathcal{L}^2(\Omega)} \\ &= \left\langle e^{-it\lambda}, e^{-i(t-k)\lambda} \right\rangle_{\mathcal{L}^2(F)} \\ &= \int_{\mathbb{T}} e^{-it\lambda} e^{i(t-k)\lambda} F(d\lambda) \\ &= \int_{\mathbb{T}} e^{-ik\lambda} F(d\lambda). \end{aligned}$$

This recovers the covariance representation directly from the process representation.

When a spectral density exists. Let ν denote normalized Lebesgue measure on $\mathbb{T}$:

$$\nu(d\lambda) = \frac{d\lambda}{2\pi}.$$

If F is absolutely continuous with respect to ν , the Radon–Nikodym theorem gives a non-negative integrable function

$$f = \frac{dF}{d\nu}, \quad F(d\lambda) = f(\lambda) \frac{d\lambda}{2\pi}.$$

We call f the spectral density under the normalized-Lebesgue convention. In particular,

$$\int_{-\pi}^{\pi} f(\lambda) \frac{d\lambda}{2\pi} = \gamma(0),$$

and

$$\gamma(k) = \int_{-\pi}^{\pi} e^{-ik\lambda} f(\lambda) \frac{d\lambda}{2\pi}. \quad (9)$$

Under the alternative convention in which the density is defined relative to $d\lambda$, the density is $f(\lambda)/(2\pi)$.

A spectral measure need not have a density. The representation in terms of F applies whether the measure is absolutely continuous, atomic, or has a singular continuous component. In particular, covariance stationarity alone does not justify replacing $F(d\lambda)$ by a density times $d\lambda$.

Distinguishing the two frequency-domain spaces. When a density exists, the space used for stochastic integration has norm

$$\|g\|_{\mathcal{L}^2(F)}^2 = \int_{-\pi}^{\pi} |g(\lambda)|^2 f(\lambda) \frac{d\lambda}{2\pi}.$$

This is different from the normalized-Lebesgue space used for ordinary Fourier expansions:

$$\|g\|_{\mathcal{L}^2(\nu)}^2 = \int_{-\pi}^{\pi} |g(\lambda)|^2 \frac{d\lambda}{2\pi}.$$

The exponentials form an orthonormal basis in $\mathcal{L}^2(\mathbb{T}, \nu)$, but generally not in $\mathcal{L}^2(\mathbb{T}, F)$. Consequently, membership of α and β in $\mathcal{L}^2(\mathbb{T}, F)$ does not by itself justify applying the usual Fourier orthonormal-expansion theorem to them.

Recovering the spectral density from the autocovariances. The existence of a spectral density guarantees that $f \in \mathcal{L}^1(\mathbb{T}, \nu)$, but not necessarily that $f \in \mathcal{L}^2(\mathbb{T}, \nu)$. Suppose, in addition, that

$$\int_{-\pi}^{\pi} |f(\lambda)|^2 \frac{d\lambda}{2\pi} < \infty.$$

The functions $\psi_k(\lambda) = e^{ik\lambda}$, $k \in \mathbb{Z}$, form an orthonormal basis in $\mathcal{L}^2(\mathbb{T}, \nu)$. By (9), the Fourier coefficients of f are

$$\langle f, \psi_k \rangle_{\mathcal{L}^2(\nu)} = \int_{-\pi}^{\pi} f(\lambda) e^{-ik\lambda} \frac{d\lambda}{2\pi} = \gamma(k).$$

Therefore,

$$f(\lambda) = \sum_{k \in \mathbb{Z}} \gamma(k) e^{ik\lambda} \quad \text{in } \mathcal{L}^2(\mathbb{T}, \nu).$$

More explicitly,

$$\lim_{N \rightarrow \infty} \int_{-\pi}^{\pi} \left| f(\lambda) - \sum_{k=-N}^N \gamma(k) e^{ik\lambda} \right|^2 \frac{d\lambda}{2\pi} = 0.$$

Parseval's identity also gives

$$\int_{-\pi}^{\pi} |f(\lambda)|^2 \frac{d\lambda}{2\pi} = \sum_{k \in \mathbb{Z}} |\gamma(k)|^2.$$

A convenient stronger sufficient condition is absolute summability of the autocovariances:

$$\sum_{k \in \mathbb{Z}} |\gamma(k)| < \infty.$$

Under this condition, the spectral measure has a bounded continuous density given by

$$f(\lambda) = \sum_{k \in \mathbb{Z}} \gamma(k) e^{ik\lambda},$$

with absolute and uniform convergence. In particular, this density belongs to $\mathcal{L}^2(\mathbb{T}, \nu)$.

In general, the full autocovariance sequence uniquely determines the spectral measure F , and F determines the full autocovariance sequence. When a spectral density exists, it is determined up to equality almost everywhere, rather than at every frequency. Thus, the covariance and spectral representations describe the same second-order information, while the form and mode of convergence of a Fourier-series reconstruction require the additional conditions stated above.