Appendix A

Mathematical Complements

Important mathematical concepts are reviewed without proof. Sections A.1–A.5 present results of real and complex analysis, including properties of Hilbert spaces, bases and linear operators [58]. Random vectors and Dirac distributions are covered in the last two sections.

A.1 Functions and Integration

Analog signals are modeled by measurable functions. We first give the main theorems of Lebesgue integration. A function f is said to be *integrable* if $\int_{-\infty}^{+\infty} |f(t)|\,dt < +\infty$. The space of integrable functions is written $\mathbf{L^1}(\mathbb{R})$. Two functions f_1 and f_2 are equal in $\mathbf{L^1}(\mathbb{R})$ if

$$\int_{-\infty}^{+\infty} |f_1(t) - f_2(t)|\,dt = 0.$$

This means that $f_1(t)$ and $f_2(t)$ can differ only on a set of points of measure 0. We say that they are *almost everywhere* equal.

The Fatou lemma gives an inequality when taking a limit under the Lebesgue integral of positive functions.

Lemma A.1 (Fatou). *Let $\{f_n\}_{n\in\mathbb{N}}$ be a family of positive functions $f_n(t) \geqslant 0$. If $\lim_{n\to+\infty} f_n(t) = f(t)$ almost everywhere then*

$$\int_{-\infty}^{+\infty} f(t)\,dt \leqslant \lim_{n\to+\infty} \int_{-\infty}^{+\infty} f_n(t)\,dt.$$

The dominated convergence theorem supposes the existence of an integrable upper bound to obtain an equality when taking a limit under a Lebesgue integral.

Theorem A.1 (Dominated Convergence). *Let $\{f_n\}_{n\in\mathbb{N}}$ be a family such that $\lim_{n\to+\infty} f_n(t) = f(t)$ almost everywhere. If*

$$\forall n \in \mathbb{N} \quad |f_n(t)| \leqslant g(t) \quad \textit{and} \quad \int_{-\infty}^{+\infty} g(t)\,dt < +\infty \tag{A.1}$$

then f is integrable and

$$\int_{-\infty}^{+\infty} f(t)\,dt = \lim_{n\to+\infty} \int_{-\infty}^{+\infty} f_n(t)\,dt.$$

The Fubini theorem gives a sufficient condition for inverting the order of integrals in multidimensional integrations.

Theorem A.2 (Fubini)**.** *If* $\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty}|f(x_1,x_2)|dx_1\right)dx_2 < +\infty$ *then*

$$\begin{aligned}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x_1,x_2)\,dx_1\,dx_2 &= \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} f(x_1,x_2)\,dx_1\right)dx_2 \\ &= \int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} f(x_1,x_2)\,dx_2\right)dx_1\ .\end{aligned}$$

Convexity A function $f(t)$ is said to be *convex* if for all $p_1, p_2 > 0$ with $p_1 + p_2 = 1$ and all $(t_1, t_2) \in \mathbb{R}^2$,

$$f(p_1 t_1 + p_2 t_2) \leqslant p_1\, f(t_1) + p_2\, f(t_2)\ .$$

The function $-f$ satisfies the reverse inequality and is said to be *concave*. If f is convex then the Jensen inequality generalizes this property for any $p_k \geqslant 0$ with $\sum_{k=1}^{K} p_k = 1$ and any $t_k \in \mathbb{R}$:

$$f\left(\sum_{k=1}^{K} p_k t_k\right) \leqslant \sum_{k=1}^{K} p_k\, f(t_k)\ . \tag{A.2}$$

The following theorem relates the convexity to the sign of the second order derivative.

Theorem A.3. *If f is twice differentiable, then f is convex if and only if $f''(t) \geqslant 0$ for all $t \in \mathbb{R}$.*

The notion of convexity also applies to sets $\Omega \subset \mathbb{R}^n$. This set is convex if for all $p_1, p_2 > 0$ with $p_1 + p_2 = 1$ and all $(x_1, x_2) \in \Omega^2$, then $p_1 x_1 + p_2 x_2 \in \Omega$. If Ω is not convex then its convex hull is defined as the smallest convex set that includes Ω.

A.2 Banach and Hilbert Spaces

Banach Space Signals are often considered as vectors. To define a distance, we work within a vector space $\mathbf{H}$ that admits a norm. A norm satisfies the following properties:

$$\forall f \in \mathbf{H}\ ,\quad \|f\| \geqslant 0 \ \text{ and }\ \|f\| = 0 \quad\Leftrightarrow\quad f = 0, \tag{A.3}$$

$$\forall \lambda \in \mathbb{C}\quad \|\lambda f\| = |\lambda|\,\|f\|, \tag{A.4}$$

$$\forall f, g \in \mathbf{H}\ ,\quad \|f + g\| \leqslant \|f\| + \|g\|. \tag{A.5}$$

With such a norm, the convergence of $\{f_n\}_{n\in\mathbb{N}}$ to f in $\mathbf{H}$ means that

$$\lim_{n\to+\infty} f_n = f \quad\Leftrightarrow\quad \lim_{n\to+\infty} \|f_n - f\| = 0.$$

To guarantee that we remain in $\mathbf{H}$ when taking such limits, we impose a completeness property, using the notion of *Cauchy sequences*. A sequence $\{f_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence if for any $\varepsilon > 0$, if n and p are large enough, then $\|f_n - f_p\| < \varepsilon$. The space $\mathbf{H}$ is said to be *complete* if every Cauchy sequence in $\mathbf{H}$ converges to an element of $\mathbf{H}$.

Example A.1. *For any integer $p \geqslant 1$ we define over discrete sequences $f[n]$*

$$\|f\|_p = \left(\sum_{n=-\infty}^{+\infty} |f[n]|^p\right)^{1/p}\ .$$

The space $\boldsymbol{\ell}^p = \{f\ :\ \|f\|_p < +\infty\}$ is a Banach space with the norm $\|f\|_p$.

Example A.2. *The space $\mathbf{L^p}(\mathbb{R})$ is composed of the measurable functions f on $\mathbb{R}$ for which*

$$\|f\|_p = \left(\int_{-\infty}^{+\infty} |f(t)|^p\,dt\right)^{1/p} < +\infty.$$

This integral defines a norm for $p \geqslant 1$ and $\mathbf{L^p}(\mathbb{R})$ is a Banach space, provided one identifies functions that are equal almost everywhere.

Hilbert Space Whenever possible, we work in a space that has an inner product to define angles and orthogonality. A *Hilbert space* **H** is a Banach space with an inner product. The inner product of two vectors $\langle f, g\rangle$ is linear with respect to its first argument:

$$\forall \lambda_1, \lambda_2 \in \mathbb{C} \quad , \quad \langle \lambda_1 f_1 + \lambda_2 f_2, g\rangle = \lambda_1 \langle f_1, g\rangle + \lambda_2 \langle f_2, g\rangle. \tag{A.6}$$

It has an Hermitian symmetry:

$$\langle f, g\rangle = \langle g, f\rangle^*.$$

Moreover

$$\langle f, f\rangle \geqslant 0 \;\text{ and }\; \langle f, f\rangle = 0 \quad\Leftrightarrow\quad f = 0.$$

One can verify that $\|f\| = \langle f, f\rangle^{1/2}$ is a norm. The positivity (A.3) implies the Cauchy-Schwarz inequality:

$$|\langle f, g\rangle| \leqslant \|f\|\,\|g\|, \tag{A.7}$$

which is an equality if and only if f and g are linearly dependent.

We write $\mathbf{V}^\perp$ the orthogonal complement of a subspace $\mathbf{V}$ of $\mathbf{H}$. All vectors of $\mathbf{V}$ are orthogonal to all vectors of $\mathbf{V}^\perp$ and $\mathbf{V} \oplus \mathbf{V}^\perp = \mathbf{H}$.

Example A.3. *An inner product between discrete signals $f[n]$ and $g[n]$ can be defined by*

$$\langle f, g\rangle = \sum_{n=-\infty}^{+\infty} f[n]\, g^*[n].$$

It corresponds to an $\boldsymbol{\ell}^2(\mathbb{Z})$ norm:

$$\|f\|^2 = \langle f, f\rangle = \sum_{n=-\infty}^{+\infty} |f[n]|^2.$$

The space $\boldsymbol{\ell}^2(\mathbb{Z})$ of finite energy sequences is therefore a Hilbert space. The Cauchy-Schwarz inequality (A.7) proves that

$$\left|\sum_{n=-\infty}^{+\infty} f[n]\, g^*[n]\right| \leqslant \left(\sum_{n=-\infty}^{+\infty} |f[n]|^2\right)^{1/2} \left(\sum_{n=-\infty}^{+\infty} |g[n]|^2\right)^{1/2}.$$

Example A.4. *Over analog signals $f(t)$ and $g(t)$, an inner product can be defined by*

$$\langle f, g\rangle = \int_{-\infty}^{+\infty} f(t)\, g^*(t)\, dt.$$

The resulting norm is

$$\|f\| = \left(\int_{-\infty}^{+\infty} |f(t)|^2\, dt\right)^{1/2}.$$

The space $\mathbf{L^2}(\mathbb{R})$ of finite energy functions is thus also a Hilbert space. In $\mathbf{L^2}(\mathbb{R})$, the Cauchy-Schwarz inequality (A.7) is

$$\left|\int_{-\infty}^{+\infty} f(t)\, g^*(t)\, dt\right| \leqslant \left(\int_{-\infty}^{+\infty} |f(t)|^2\, dt\right)^{1/2} \left(\int_{-\infty}^{+\infty} |g(t)|^2\, dt\right)^{1/2}.$$

Two functions f_1 and f_2 are equal in $\mathbf{L^2}(\mathbb{R})$ if

$$\|f_1 - f_1\|^2 = \int_{-\infty}^{+\infty} |f_1(t) - f_2(t)|^2\, dt = 0,$$

which means that $f_1(t) = f_2(t)$ for almost all $t \in \mathbb{R}$.

A.3 Bases of Hilbert Spaces

Orthonormal Basis A family $\{e_n\}_{n\in\mathbb{N}}$ of a Hilbert space $\mathbf{H}$ is orthogonal if for $n \neq p$

$$\langle e_n, e_p \rangle = 0.$$

If for $f \in \mathbf{H}$ there exists a sequence $a[n]$ such that

$$\lim_{N\to+\infty} \|f - \sum_{n=0}^{N} a[n]\, e_n\| = 0,$$

then $\{e_n\}_{n\in\mathbb{N}}$ is said to be an *orthogonal basis* of $\mathbf{H}$. The orthogonality implies that necessarily $a[n] = \langle f, e_n\rangle / \|e_n\|^2$ and we write

$$f = \sum_{n=0}^{+\infty} \frac{\langle f, e_n\rangle}{\|e_n\|^2}\, e_n. \tag{A.8}$$

A Hilbert space that admits an orthogonal basis is said to be *separable*.

The basis is *orthonormal* if $\|e_n\| = 1$ for all $n \in \mathbb{N}$. Computing the inner product of $g \in \mathbf{H}$ with each side of (A.8) yields a Parseval equation for orthonormal bases:

$$\langle f, g\rangle = \sum_{n=0}^{+\infty} \langle f, e_n\rangle\, \langle g, e_n\rangle^*. \tag{A.9}$$

When $g = f$, we get an energy conservation called the *Plancherel formula*:

$$\|f\|^2 = \sum_{n=0}^{+\infty} |\langle f, e_n\rangle|^2. \tag{A.10}$$

The Hilbert spaces $\boldsymbol{\ell}^2(\mathbb{Z})$ and $\mathbf{L^2}(\mathbb{R})$ are separable. For example, the family of translated Diracs $\{e_n[k] = \delta[k-n]\}_{n\in\mathbb{Z}}$ is an orthonormal basis of $\boldsymbol{\ell}^2(\mathbb{Z})$. Chapter 7 and Chapter 8 construct orthonormal bases of $\mathbf{L^2}(\mathbb{R})$ with wavelets, wavelet packets and local cosine functions.

Riesz Bases In an infinite dimensional space, if we loosen up the orthogonality requirement, we must still impose a partial energy equivalence to guarantee the stability of the basis. A family of vectors $\{e_n\}_{n\in\mathbb{N}}$ is said to be a Riesz basis of $\mathbf{H}$ if it is linearly independent and if there exist $B \geqslant A > 0$ such that

$$\forall f \in \mathbf{H} \quad , \quad A\,\|f\|^2 \leqslant \sum_{n=0}^{+\infty} |\langle f, e_n\rangle|^2 \leqslant B\,\|f\|^2. \tag{A.11}$$

Section 5.1.2 proves that there exists a unique dual basis $\{\tilde{e}_n\}_{n\in\mathbb{N}}$ characterized by biorthogonality relations

$$\forall (n,p) \in \mathbb{N}^2 \quad , \quad \langle e_n, \tilde{e}_p\rangle = \delta[n-p], \tag{A.12}$$

and which satisfies

$$\forall f \in \mathbf{H} \quad , \quad f = \sum_{n=0}^{+\infty} \langle f, \tilde{e}_n\rangle\, e_n = \sum_{n=0}^{+\infty} \langle f, e_n\rangle\, \tilde{e}_n.$$

A.4 Linear Operators

Classical signal processing algorithms are mostly based on linear operators. An operator U from a Hilbert space $\mathbf{H}_1$ to another Hilbert space $\mathbf{H}_2$ is linear if

$$\forall \lambda_1, \lambda_2 \in \mathbb{C}\ ,\ \forall f_1, f_2 \in \mathbf{H}\ ,\ U(\lambda_1\, f_1 + \lambda_2\, f_2) = \lambda_1\, U(f_1) + \lambda_2\, U(f_2).$$

The null space and image spaces of U are defined by

$$\mathbf{NullU} = \{h \in \mathbf{H}_1 \ :\ Uh = 0\} \quad \text{and} \quad \mathbf{ImU} = \{g \in \mathbf{H}_2 \ :\ \exists h \in H_1\, ,\, g = Uh\}.$$

Sup Norm The sup operator norm of U is defined by

$$\|U\|_S = \sup_{f \in \mathbf{H}_1} \frac{\|Uf\|}{\|f\|}. \tag{A.13}$$

If this norm is finite, then U is continuous. Indeed, $\|Uf - Ug\|$ becomes arbitrarily small if $\|f-g\|$ is sufficiently small.

Adjoint The *adjoint* of U is the operator U^* from $\mathbf{H}_2$ to $\mathbf{H}_1$ such that for any $f \in \mathbf{H}_1$ and $g \in \mathbf{H}_2$

$$\langle Uf, g\rangle = \langle f, U^* g\rangle.$$

The null and image spaces of adjoint operators are orthogonal complement:

$$\mathbf{NullU} = (\mathbf{ImU}^*)^{\perp} \quad \text{and} \quad \mathbf{ImU} = (\mathbf{NullU}^*)^{\perp}.$$

When U is defined from $\mathbf{H}$ into itself, it is *self-adjoint* if $U = U^*$. It is also said to be *symmetric*.

A non-zero vector $f \in \mathbf{H}$ is a called an *eigenvector* if there exists an *eigenvalue* $\lambda \in \mathbb{C}$ such that

$$Uf = \lambda\, f.$$

In a finite dimensional Hilbert space (Euclidean space), a self-adjoint operator is always diagonalized by an orthogonal basis $\{e_n\}_{0 \leqslant n < N}$ of eigenvectors

$$Ue_n = \lambda_n\, e_n.$$

When U is self-adjoint the eigenvalues λ_n are real. For any $f \in \mathbf{H}$,

$$Uf = \sum_{n=0}^{N-1} \langle Uf, e_n\rangle\, e_n = \sum_{n=0}^{N-1} \lambda_n \langle f, e_n\rangle\, e_n.$$

For any U, the operators U^*U and UU^* are self-adjoint and have the same eigenvalues. These eigenvalues are called *singular values* of U.

In an infinite dimensional Hilbert space, the eigenvalues of symmetric operators are generalized by introducing the spectrum of the operator.

Orthogonal Projector Let $\mathbf{V}$ be a subspace of $\mathbf{H}$. A *projector* $P_{\mathbf{V}}$ on $\mathbf{V}$ is a linear operator that satisfies

$$\forall f \in \mathbf{H} \;\; , \;\; P_{\mathbf{V}} f \in \mathbf{V} \;\; \text{and} \;\; \forall f \in \mathbf{V} \;\; , \;\; P_{\mathbf{V}} f = f.$$

The projector $P_{\mathbf{V}}$ is *orthogonal* if

$$\forall f \in \mathbf{H} \;\; , \;\; \forall g \in \mathbf{V} \;\; , \;\; \langle f - P_{\mathbf{V}} f,\, g\rangle = 0.$$

The following properties are often used.

Theorem A.4. *If $P_{\mathbf{V}}$ is a projector on $\mathbf{V}$ then the following statements are equivalent:*

(i) *$P_{\mathbf{V}}$ is orthogonal.*
(ii) *$P_{\mathbf{V}}$ is self-adjoint.*
(iii) $\|P_{\mathbf{V}}\|_S = 1$.
(iv) $\forall f \in \mathbf{H}$, $\|f - P_{\mathbf{V}} f\| = \min_{g \in \mathbf{V}} \|f - g\|$.
(v) *If $\{e_n\}_{n \in \mathbb{N}}$ is an orthogonal basis of $\mathbf{V}$ then*

$$P_{\mathbf{V}} f = \sum_{n=0}^{+\infty} \frac{\langle f, e_n\rangle}{\|e_n\|^2}\, e_n. \tag{A.14}$$

(vi) *If $\{e_n\}_{n \in \mathbb{N}}$ and $\{\tilde{e}_n\}_{n \in \mathbb{N}}$ are biorthogonal Riesz bases of $\mathbf{V}$ then*

$$P_{\mathbf{V}} f = \sum_{n=0}^{+\infty} \langle f, e_n\rangle\, \tilde{e}_n = \sum_{n=0}^{+\infty} \langle f, \tilde{e}_n\rangle\, e_n \;. \tag{A.15}$$

Limit and Density Argument Let $\{U_n\}_{n\in\mathbb{N}}$ be a sequence of linear operators from $\mathbf{H}$ to $\mathbf{H}$. Such a sequence *converges weakly* to a linear operator U_∞ if

$$\forall f \in \mathbf{H} \quad , \quad \lim_{n\to+\infty} \|U_n f - U_\infty f\| = 0.$$

To find the limit of operators it is often preferable to work in a well chosen subspace $\mathbf{V} \subset \mathbf{H}$ which is dense. A space $\mathbf{V}$ is *dense* in $\mathbf{H}$ if for any $f \in \mathbf{H}$ there exist $\{f_m\}_{m\in\mathbb{N}}$ with $f_m \in \mathbf{V}$ such that

$$\lim_{m\to+\infty} \|f - f_m\| = 0.$$

The following theorem justifies this approach.

Theorem A.5 (Density)**.** *Let $\mathbf{V}$ be a dense subspace of $\mathbf{H}$. Suppose that there exists C such that $\|U_n\|_S \leqslant C$ for all $n \in \mathbb{N}$. If*

$$\forall f \in \mathbf{V} \quad , \quad \lim_{n\to+\infty} \|U_n f - U_\infty f\| = 0 \ ,$$

then

$$\forall f \in \mathbf{H} \quad , \quad \lim_{n\to+\infty} \|U_n f - U_\infty f\| = 0.$$

A.5 Separable Spaces and Bases

Tensor Product Tensor products are used to extend spaces of one-dimensional signals into spaces of multiple dimensional signals. A tensor product $f_1 \otimes f_2$ between vectors of two Hilbert spaces $\mathbf{H}_1$ and $\mathbf{H}_2$ satisfies the following properties:

Linearity

$$\forall \lambda \in \mathbb{C} \ , \quad \lambda\,(f_1 \otimes f_2) = (\lambda\, f_1) \otimes f_2 = f_1 \otimes (\lambda\, f_2). \tag{A.16}$$

Distributivity

$$(f_1 + g_1) \otimes (f_2 + g_2) = (f_1 \otimes f_2) + (f_1 \otimes g_2) + (g_1 \otimes f_2) + (g_1 \otimes g_2). \tag{A.17}$$

This tensor product yields a new Hilbert space $\mathbf{H} = \mathbf{H}_1 \otimes \mathbf{H}_2$ that includes all vectors of the form $f_1 \otimes f_2$ where $f_1 \in \mathbf{H}_1$ and $f_2 \in \mathbf{H}_2$, as well as linear combinations of such vectors. An inner product in $\mathbf{H}$ is derived from inner products in $\mathbf{H}_1$ and $\mathbf{H}_2$ by

$$\langle f_1 \otimes f_2 \, , \, g_1 \otimes g_2 \rangle_{\mathbf{H}} \ = \ \langle f_1, g_1 \rangle_{\mathbf{H}_1} \, \langle f_2, g_2 \rangle_{\mathbf{H}_2}. \tag{A.18}$$

Separable Bases The following theorem proves that orthonormal bases of tensor product spaces are obtained with separable products of two orthonormal bases. It provides a simple procedure for transforming bases for one-dimensional signals into separable bases for multidimensional signals.

Theorem A.6. *Let $\mathbf{H} = \mathbf{H}_1 \otimes \mathbf{H}_2$. If $\{e_n^1\}_{n\in\mathbb{N}}$ and $\{e_n^2\}_{n\in\mathbb{N}}$ are two Riesz bases respectively of $\mathbf{H}_1$ and $\mathbf{H}_2$ then $\{e_n^1 \otimes e_m^2\}_{(n,m)\in\mathbb{N}^2}$ is a Riesz basis of $\mathbf{H}$. If the two bases are orthonormal then the tensor product basis is also orthonormal.*

Example A.5. *A product of functions $f \in \mathbf{L^2}(\mathbb{R})$ and $g \in \mathbf{L^2}(\mathbb{R})$ defines a tensor product:*

$$f(x_1)\, g(x_2) = f \otimes g(x_1, x_2).$$

Let $\mathbf{L^2}(\mathbb{R}^2)$ be the space of $h(x_1, x_2)$ such that

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |h(x_1, x_2)|^2 \, dx_1 \, dx_2 < +\infty.$$

One can verify that $\mathbf{L^2}(\mathbb{R}^2) = \mathbf{L^2}(\mathbb{R}) \otimes \mathbf{L^2}(\mathbb{R})$. Theorem A.6 proves that if $\{\psi_n(t)\}_{n\in\mathbb{N}}$ is an orthonormal basis of $\mathbf{L^2}(\mathbb{R})$, then $\{\psi_{n_1}(x_1)\, \psi_{n_2}(x_2)\}_{(n_1,n_2)\in\mathbb{N}^2}$ is an orthonormal basis of $\mathbf{L^2}(\mathbb{R}^2)$.

Example A.6. *A product of discrete signals $f \in \ell^2(\mathbb{Z})$ and $g \in \ell^2(\mathbb{Z})$ also defines a tensor product:*

$$f[n_1]\, g[n_2] = f \otimes g[n_1, n_2].$$

The space $\ell^2(\mathbb{Z}^2)$ of images $h[n_1, n_2]$ such that

$$\sum_{n_1=-\infty}^{+\infty} \sum_{n_2=-\infty}^{+\infty} |h[n_1, n_2]|^2 < +\infty$$

is also decomposed as a tensor product $\ell^2(\mathbb{Z}^2) = \ell^2(\mathbb{Z}) \otimes \ell^2(\mathbb{Z})$. Orthonormal bases can thus be constructed with separable products.

A.6 Random Vectors and Covariance Operators

A class of signals can be modeled by a random process (random vector) whose realizations are the signals in the class. Finite discrete signals f are represented by a random vector Y, where $Y[n]$ is a random variable for each $0 \leqslant n < N$. For a review of elementary probability theory for signal processing, the reader may consult [52, 55].

Covariance Operator If $p(x)$ is the probability density of a random variable X, the expected value is

$$\mathsf{E}\{X\} = \int x\, p(x)\, dx$$

and the variance is $\sigma^2 = \mathsf{E}\{|X - \mathsf{E}\{X\}|^2\}$. The covariance of two random variables X_1 and X_2 is

$$\mathrm{Cov}(X_1, X_2) = \mathsf{E}\Big\{\Big(X_1 - \mathsf{E}\{X_1\}\Big)\Big(X_2 - \mathsf{E}\{X_2\}\Big)^*\Big\}. \tag{A.19}$$

The covariance matrix of a random vector Y is composed of the N^2 covariance values

$$R_Y[n, m] = \mathrm{Cov}\Big(Y[n], Y[m]\Big).$$

It defines the covariance operator K_Y which transforms any $h[n]$ into

$$K_Y h[n] = \sum_{m=0}^{N-1} R_Y[n, m]\, h[m].$$

For any h and g

$$\langle Y, h\rangle = \sum_{n=0}^{N-1} Y[n]\, h^*[n] \quad \text{and} \quad \langle Y, g\rangle = \sum_{n=0}^{N-1} Y[n]\, g^*[n]$$

are random variables and

$$\mathrm{Cov}\Big(\langle Y, h\rangle, \langle Y, g\rangle\Big) = \langle K_Y g, h\rangle. \tag{A.20}$$

The covariance operator thus specifies the covariance of linear combinations of the process values. If $\mathsf{E}\{Y[n]\} = 0$ for all $0 \leqslant n < N$ then $\mathsf{E}\{\langle Y, h\rangle\} = 0$ for all h.

Karhunen-Loève Basis The covariance operator K_Y is self-adjoint because $R_Y[n, m] = R_Y^*[m, n]$ and positive because

$$\langle K_Y h, h\rangle = \mathsf{E}\{|\langle Y, h\rangle - \mathsf{E}\{\langle Y, h\rangle\}|^2\} \geqslant 0. \tag{A.21}$$

This guarantees the existence of an orthogonal basis $\{e_k\}_{0 \leqslant k < N}$ that diagonalizes K_Y:

$$K_Y e_k = \sigma_k^2\, e_k.$$

This basis is called a *Karhunen-Loève basis* of Y, and the vectors e_k are the *principal directions.* The eigenvalues are the variances

$$\sigma_k^2 = \langle K_Y e_k, e_k\rangle = \mathsf{E}\{|\langle Y, e_k\rangle - \mathsf{E}\{\langle Y, e_k\rangle\}|^2\}. \tag{A.22}$$

Wide-Sense Stationarity We say that Y is *wide-sense stationary* if

$$\mathrm{Cov}(Y[n], Y[m]) = R_Y[n,m] = R_Y[n-m]. \tag{A.23}$$

The covariance between two points depends only on the distance between these points. The operator K_Y is then a convolution whose kernel $R_Y[k]$ is defined for $-N < k < N$. A wide-sense stationary process is *circular stationary* if $R_Y[n]$ is N periodic:

$$R_Y[n] = R_Y[N+n] \quad \text{for } -N \leqslant n \leqslant 0. \tag{A.24}$$

This condition implies that a periodic extension of $Y[n]$ on $\mathbb{Z}$ remains wide-sense stationary on $\mathbb{Z}$. The covariance operator K_Y of a circular stationary process is a discrete circular convolution. Section 3.3.1 proves that the eigenvectors of circular convolutions are the discrete Fourier vectors

$$\left\{ e_k[n] = \frac{1}{\sqrt{N}} \exp\left(\frac{i2\pi kn}{N}\right)\right\}_{0\leqslant k<N}.$$

The discrete Fourier basis is therefore the Karhunen-Loève basis of circular stationary processes. The eigenvalues (A.22) of K_Y are the discrete Fourier transform of R_Y and are called the *power spectrum*

$$\sigma_k^2 = \hat{R}_Y[k] = \sum_{n=0}^{N-1} R_Y[n] \exp\left(\frac{-i2k\pi n}{N}\right) \quad . \tag{A.25}$$

The following theorem computes the power spectrum after a circular convolution.

Theorem A.7. *Let Z be a wide-sense circular stationary random vector. The random vector $Y[n] = Z \circledast h[n]$ is also wide-sense circular stationary and its power spectrum is*

$$\hat{R}_Y[k] = \hat{R}_Z[k]\, |\hat{h}[k]|^2. \tag{A.26}$$

A.7 Diracs

Diracs are useful in making the transition from functions of a real variable to discrete sequences. Symbolic calculations with Diracs simplify computations, without worrying about convergence issues. This is justified by the theory of distributions [60, 63]. A Dirac δ has a support reduced to $t = 0$ and associates to any continuous function ϕ its value at $t = 0$

$$\int_{-\infty}^{+\infty} \delta(t)\, \phi(t)\, dt = \phi(0). \tag{A.27}$$

Weak Convergence A Dirac can be obtained by squeezing an integrable function g such that $\int_{-\infty}^{+\infty} g(t)\, dt = 1$. Let $g_s(t) = s^{-1} g(s^{-1}t)$. For any continuous function ϕ

$$\lim_{s\to 0} \int_{-\infty}^{+\infty} g_s(t)\, \phi(t)\, dt = \phi(0) = \int_{-\infty}^{+\infty} \delta(t)\, \phi(t)\, dt. \tag{A.28}$$

A Dirac can thus formally be defined as the limit $\delta = \lim_{s\to 0} g_s$, which must be understood in the sense of (A.28). This is called *weak convergence*. A Dirac is not a function since it is zero at $t \neq 0$ although its "integral" is equal to 1. The integral at the right of (A.28) is only a symbolic notation which means that a Dirac applied to a continuous function ϕ associates its value at $t = 0$.

General distributions are defined over the space $\mathbf{C}_0^\infty$ of *test functions* which are infinitely continuously differentiable with a compact support. A distribution d is a linear form that associates to any $\phi \in \mathbf{C}_0^\infty$ a value that is written $\int_{-\infty}^{+\infty} d(t)\phi(t)dt$. It must also satisfy some weak continuity properties [60, 63] that we do not discuss here, and which are satisfied by a Dirac. Two distributions d_1 and d_2 are equal if

$$\forall \phi \in \mathbf{C}_0^\infty \quad , \quad \int_{-\infty}^{+\infty} d_1(t)\, \phi(t)\, dt = \int_{-\infty}^{+\infty} d_2(t)\, \phi(t)\, dt. \tag{A.29}$$

Symbolic Calculations The symbolic integral over a Dirac is a useful notation because it has the same properties as a usual integral, including change of variables and integration by parts. A translated Dirac $\delta_\tau(t) = \delta(t-\tau)$ has a mass concentrated at τ and

$$\int_{-\infty}^{+\infty} \phi(t)\,\delta(t-u)\,dt = \int_{-\infty}^{+\infty} \phi(t)\,\delta(u-t)\,dt = \phi(u).$$

This means that $\phi \star \delta(u) = \phi(u)$. Similarly $\phi \star \delta_\tau(u) = \phi(u-\tau)$.

A Dirac can also be multiplied by a continuous function ϕ and since $\delta(t-\tau)$ is zero outside $t=\tau$, it follows that

$$\phi(t)\,\delta(t-\tau) = \phi(\tau)\,\delta(t-\tau).$$

The derivative of a Dirac is defined with an integration by parts. If ϕ is continuously differentiable then

$$\int_{-\infty}^{+\infty} \phi(t)\,\delta'(t)\,dt = -\int_{-\infty}^{+\infty} \phi'(t)\,\delta(t)\,dt = -\phi'(0).$$

The k^{th} derivative of δ is similarly obtained with k integrations by parts. It is a distribution that associates to $\phi \in \mathbf{C}^k$

$$\int_{-\infty}^{+\infty} \phi(t)\,\delta^{(k)}(t)\,dt = (-1)^k\,\phi^{(k)}(0).$$

The Fourier transform of δ associates to any $\mathrm{e}^{-i\omega t}$ its value at $t=0$:

$$\hat{\delta}(\omega) = \int_{-\infty}^{+\infty} \delta(t)\,\mathrm{e}^{-i\omega t}\,dt = 1,$$

and after translation $\hat{\delta}_\tau(\omega) = \mathrm{e}^{-i\tau\omega}$. The Fourier transform of the Dirac comb $c(t) = \sum_{n=-\infty}^{+\infty} \delta(t-nT)$ is therefore $\hat{c}(\omega) = \sum_{n=-\infty}^{+\infty} \mathrm{e}^{-inT\omega}$. The Poisson formula (2.4) proves that

$$\hat{c}(\omega) = \frac{2\pi}{T} \sum_{k=-\infty}^{+\infty} \delta\left(\omega - \frac{2\pi k}{T}\right).$$

This distribution equality must be understood in the sense (A.29).