# Wavelets: Mathematical Preliminaries

Pieter W. Hemker, Tom H. Koornwinder, and Nico M. Temme

**Abstract.** This paper contains some mathematical preliminaries needed for a good understanding of wavelets. The first section discusses Hilbert spaces. The second section deals with Fourier transformation. Finally, the third section gives a short introduction to Riesz bases and frames.

## §1 Hilbert space theory

A good introduction to this topic is Rudin [3]. We summarize the main elements needed for wavelets. In Rudin [4] the background for measure and integration theory can be found. We have also used the paper Heil & Walnut [2] for preparing this document.

### 1.1 Conventions

- $\mathbb{N}$ is the set of positive integers;
- $\mathbb{Z}$ is the set of integers;
- $\mathbb{R}$ is the set of real numbers;
- $\mathbb{C}$ is the set of complex numbers. The modulus of a complex number $z \in \mathbb{C}$ is denoted by $|z|$, the complex conjugate by $\overline{z}$.

Unless otherwise indicated, integration is always with respect to Lebesgue measure. A property is said to hold *almost everywhere (a.e.)* if the set of points where it fails has Lebesgue measure zero. All functions considered are complex-valued, unless otherwise indicated.

**Definition 1.1.**

(1) *The support* $\operatorname{supp}(f)$ *of a continuous function* $f$ *on* $\mathbb{R}^d$ *is the closure in* $\mathbb{R}^d$ *of* $\{x \in \mathbb{R}^d \mid f(x) \neq 0\}$.

Wavelets: An Elementary Treatment of Theory and Applications
Tom H. Koornwinder (ed.), pp. 13–26.

ISBN 981-02-1388-3

(2) *The characteristic function of a set* $E \subset \mathbb{R}^d$ *is* $\chi_E(x) := \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{if } x \notin E. \end{cases}$

(3) *The Kronecker delta is* $\delta_{x,y} := \begin{cases} 1, & \text{if } x = y, \\ 0, & \text{if } x \neq y. \end{cases}$

(4) *If* $x \in \mathbb{R}^d$ *then* $|x| := (x_1^2 + \cdots + x_d^2)^{\frac{1}{2}}$ *is the norm of* $x$. *If* $x, y \in \mathbb{R}^d$ *then* $x.y := x_1 y_1 + \cdots + x_d y_d$ *is the inner product of* $x$ *and* $y$.

### 1.2 Normed linear spaces

Let $V$ be a normed complex linear space, with *norm* of $x$ denoted by $\|x\|$. Every normed linear space is a *metric space* with respect to the *distance function* $d(x, y) := \|x - y\|$.

**Definition 1.2.**

(1) *A sequence of elements* $\{x_n\}_{n=1,2,\ldots}$ *of* $V$ *is called a Cauchy sequence if for every* $\varepsilon > 0$ *we can find* $N \in \mathbb{N}$ *such that* $\|x_n - x_m\| < \varepsilon$ *if* $n, m > N$.

(2) *A normed linear space* $V$ *is said to be complete if every Cauchy sequence has a limit in* $V$. *That is, if* $\{x_n\}$ *is a Cauchy sequence in* $V$, *then there exists an element* $x \in V$ *such that* $\lim_{n\to\infty} \|x - x_n\| = 0$ *(equivalently written as* $\lim_{n\to\infty} x_n = x$*). A complete normed linear space is called a Banach space.*

Let $V$ be a complex linear space. Let an inner product be defined on the Cartesian product $V \times V$. That is, corresponding to each $x, y \in V$ there is a complex number, called the *inner product* of $x$ and $y$, and denoted by $(x, y)$, such that

(a) $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$,

(b) $(\alpha x, y) = \alpha(x, y)$, for any complex number $\alpha$,

(c) $(x, y) = \overline{(y, x)}$,

(d) $(x, x) > 0$ for $x \neq 0$.

Any linear space with an inner product is called an *inner product space.* Every inner product space is a *normed linear space* with respect to the norm $\|x\| = \sqrt{(x, x)}$.

A linear inner product space $V$ that is complete under the norm induced by the inner product is known as a *Hilbert space.* (Thus $V$ is also a Banach space with respect to this norm.)

**Definition 1.3.** *For a Lebesgue measururable function* $f$ *on* $\mathbb{R}^d$ *put*

$$\|f\|_p := \Big(\int_{\mathbb{R}^d} |f(x)|^p \, dx\Big)^{1/p}, \quad 1 \leq p < \infty,$$

$$\|f\|_\infty = \operatorname{ess\,sup}_{x \in \mathbb{R}^d} |f(x)| := \inf\{\lambda \in \mathbb{R} \mid |f(x)| \leq \lambda \text{ almost everywhere}\}.$$

*Thus* $0 \leq \|f\|_p \leq \infty$. *Then* $L^p(\mathbb{R}^d)$ $(1 \leq p \leq \infty)$ *is defined as the space of all Lebesgue measurable functions* $f$ *on* $\mathbb{R}^d$ *for which* $\|f\|_p < \infty$, *with the convention*

*that two functions $f, g \in L^p(\mathbb{R}^d)$ for which $\|f-g\|_p = 0$ (or, equivalently, $f(x) = g(x)$ outside a set of Lebesgue measure zero) will not be distinguished from each other in $L^p(\mathbb{R}^d)$.*

It is well known that, for $1 \le p \le \infty$, $L^p(\mathbb{R}^d)$ is a Banach space with norm $\|\cdot\|_p$, and that $L^2(\mathbb{R}^d)$ is a Hilbert space with inner product

$$(f,g) := \int f(x)\,\overline{g(x)}\,dx.$$

The *Hölder inequality* states that $(f,g) \le \|f\|_p \, \|g\|_q$, where $q$ is the conjugate exponent of $p$, *i.e.*, $1/p + 1/q = 1$. In particular we have the *Cauchy-Schwarz inequality* which states that, in any Hilbert space, $|(f,g)| \le \|f\|_2 \, \|g\|_2$.

Next we give a formulation of the important Fubini theorem, but without taking care of all subtleties. (See Rudin [4, Theorem 7.8] for details.) Below, the unitiated reader may read $\mathbb{R}$ for $X$ and $Y$ and Lebesgue measure for $\mu$ and $\nu$.

**Theorem 1.4.** (Fubini's theorem) *Let $X$ and $Y$ be measure spaces with $\sigma$-finite measures $\mu$ respectively $\nu$. Let $\mu \times \nu$ be the product measure on $X \times Y$.*

(i) *If $f$ is a nonnegative measurable function on $X \times Y$ then*

$$\int_{X\times Y} f\, d(\mu\times\nu) = \int_X \Big(\int_Y f(x,y)\,d\nu(y)\Big)\,d\mu(x) = \int_Y \Big(\int_X f(x,y)\,d\mu(x)\Big)\,d\nu(y), \tag{1.1}$$

*where possibly all three members of (1.1) are equal to $\infty$.*

(ii) *If $f$ is a measurable function on $X \times Y$ such that one of the members of (1.1), with $f$ replaced by $|f|$, is finite, then $f$ itself satisfies (1.1), with all members of (1.1) well-defined and finite.*

### 1.3 Hilbert spaces

**Definition 1.5.** *Let $\mathcal{H}$ be a Hilbert space. Let $\{y_n\}$ ($n \in \mathbb{N}$) be a sequence of elements of $\mathcal{H}$ and $\{x_\alpha\}_{\alpha\in A}$ a countable system of elements of $\mathcal{H}$ ($A$ being a countable index set).*

(1) *We say that $y_n$ converges to $y \in \mathcal{H}$ as $n \to \infty$, and we write $y_n \to y$, if $\lim_{n\to\infty} \|y - y_n\| = 0$.*

(2) *We write $\sum_{n=1}^{\infty} y_n = s$, and say that the series converges conditionally to $s$, if $\sum_{k=1}^{n} x_k \to s$ as $n \to \infty$.*

(3) *The series $\sum_{\alpha\in A} x_\alpha$ converges unconditionally if for every arrangement of its terms as a sequence $\{z_n\}_{n=1,2,\ldots}$ the series $\sum_{n=1}^{\infty} z_n$ converges conditionally. (Then all these series converge to the same sum.)*

(4) *The series $\sum_{\alpha\in A} x_\alpha$ converges absolutely if $\sum_{\alpha\in A} \|x_\alpha\| < \infty$.*

(5) *The span of $\{x_\alpha\}$ in $\mathcal{H}$ is the set of all finite linear combinations of the $x_\alpha$.*

(6) $\{x_\alpha\}$ *is an orthogonal system if* $(x_\alpha, x_\beta) = 0$ *whenever* $\alpha \neq \beta$.
(7) $\{x_\alpha\}$ *is an orthonormal system if it is orthogonal and* $\|x_\alpha\| = 1$ *for all* $\alpha$ *or, equivalently, if* $(x_\alpha, x_\beta) = \delta_{\alpha,\beta}$ *for all* $\alpha, \beta$.
(8) $\{x_\alpha\}$ *is complete if* $\mathrm{Span}\{x_\alpha\}$ *is dense in* $\mathcal{H}$ *or, equivalently, if the only element* $x \in \mathcal{H}$ *orthogonal to every* $x_\alpha$ *is* $x = 0$.

**Remark 1.6.**

- A series $\sum_{\alpha \in A} x_\alpha$ converges unconditionally to $s$ iff, for each $\epsilon > 0$, there exists a finite subset $B$ of $A$ such that $\|s - \sum_{\alpha \in C} x_\alpha\| < \epsilon$ for all finite subsets $C \supset B$ of $A$.
- Absolute convergence of a series implies unconditional convergence but not conversely.
- In the following, convergence of a series will mean unconditional convergence. Quite often, the index set will be the set $\mathbb{Z}$ of all integers.

**Theorem 1.7.** (The Plancherel Theorem) *Given an orthonormal system* $\{e_n\}$ *in a Hilbert space* $\mathcal{H}$, *it can be shown that the following statements are equivalent:*

$$\begin{aligned} &(1) \quad \{e_n\} \text{ is complete.} \\ &(2) \quad \sum_n |(x, e_n)|^2 = \|x\|^2 \text{ for all } x \in \mathcal{H}. \\ &(3) \quad x = \sum_n (x, e_n)\, e_n \text{ for all } x \in \mathcal{H}. \end{aligned} \tag{1.2}$$

An orthonormal system satisfying these equivalent conditions is called an *orthonormal basis.* Statement (2) of (1.2) is referred to as the *Plancherel formula* for orthonormal bases. In any expansion $x = \sum_n c_n e_n$ the coefficients $c_n$ are unique and given by statement (3) of (1.2).

Every Hilbert space has an orthonormal basis, but not necessarily countable. We will always assume that the Hilbert space under consideration has a countable orthonormal basis or, equivalently, that it is *separable.*

A standard example of a separable Hilbert space is the space $\ell^2$ of sequences $\{c_n\}_{n \in \mathbb{Z}}$ such that $\sum_n |c_n|^2 < \infty$. The inner product of two sequences $\{c_n\}$ and $\{d_n\}$ is then given by $\sum_n c_n \overline{d_n}$. All separable Hilbert spaces are isomorphic to $\ell^2$.

## 1.4 Operators

**Definition 1.8.** *Assume* $\mathcal{H}$ *and* $\mathcal{K}$ *are Hilbert spaces with norms* $\|\cdot\|_{\mathcal{H}}$, $\|\cdot\|_{\mathcal{K}}$, *and inner products* $(\cdot,\cdot)_{\mathcal{H}}$, $(\cdot,\cdot)_{\mathcal{K}}$ *respectively, and that* $S$ *is a linear operator* $S\colon \mathcal{H} \to \mathcal{K}$.
(1) $S$ *is 1-1 or injective if* $x \neq y$ *implies* $Sy \neq Sx$.
(2) *The range of* $S$ *is* $\mathrm{Range}(S) := \{Sx \mid x \in \mathcal{H}\}$.
(3) $S$ *is onto or surjective if* $\mathrm{Range}(S) = \mathcal{K}$.
(4) $S$ *is bijective if it is both injective and surjective.*

(5) *The norm of $S$ is defined by* $\|S\| := \sup\{\|Sx\|_{\mathcal{K}} \mid x \in \mathcal{H},\ \|x\|_{\mathcal{H}} = 1\}$. *We have* $\|Sx\|_{\mathcal{K}} \leq \|S\|\,\|x\|_{\mathcal{H}}$.
(6) *$S$ is bounded if* $\|S\| < \infty$ *and continuous if* $x_n \to x$ *implies* $Sx_n \to Sx$. *These two properties of a bounded operator are equivalent.*
(7) *$S$ is open if there is a positive constant $C$ such that* $\|Sx\|_{\mathcal{K}} \geq C\|x\|_{\mathcal{H}}$ *for all* $x \in \mathcal{H}$.
(8) *The adjoint of a bounded operator $S$ is the unique operator* $S^*: \mathcal{K} \to \mathcal{H}$ *such that* $(Sx, y)_{\mathcal{K}} = (x, S^*y)_{\mathcal{H}}$ *for all* $x \in \mathcal{H}$ *and* $y \in \mathcal{K}$. *It is easy to show that* $\|S^*\| = \|S\|$.
(9) *A bijective operator has an inverse* $S^{-1}: \mathcal{K} \to \mathcal{H}$ *defined by setting* $S^{-1}y = x$ *if* $Sx = y$.
(10) *We say that a bounded operator $S$ is boundedly invertible if $S$ is bijective with bounded inverse. In this case* $\|S^{-1}\|^{-1}\|x\|_{\mathcal{H}} \leq \|Sx\|_{\mathcal{K}} \leq \|S\|\,\|x\|_{\mathcal{H}}$ *for all* $x \in \mathcal{H}$.
(11) *$S$ is an isometry or norm-preserving operator if* $\|Sx\|_{\mathcal{K}} = \|x\|_{\mathcal{H}}$ *for all* $x \in \mathcal{H}$. *This is equivalent to* $(Sx, Sy)_{\mathcal{K}} = (x, y)_{\mathcal{H}}$ *for all* $x, y \in \mathcal{H}$.
(12) *$S$ is a unitary operator of $\mathcal{H}$ if it is a bijective isometry* $S: \mathcal{H} \to \mathcal{H}$. *Equivalently, $S$ is a bounded operator of $\mathcal{H}$ such that* $SS^* = I = S^*S$.

**Definition 1.9.** *Assume $\mathcal{H}$ is a Hilbert space with norm $\|\cdot\|$ and inner product $(\cdot,\cdot)$, and that $S, T: \mathcal{H} \to \mathcal{H}$ are bounded linear operators.*
(1) *$S$ is self-adjoint if* $S^* = S$.
(2) *$S$ is positive, denoted by* $S \geq 0$, *if* $(Sx, x) \geq 0$ *for all* $x \in \mathcal{H}$. *All positive operators are self-adjoint.*
(3) *We say that* $S \geq T$ *if* $S - T \geq 0$.
(4) *$S$ is positive definite if* $(Sx, x) > 0$ *for all* $x \neq 0$ *or if, equivalently, $S$ is positive and boundedly invertible.*

Each positive operator $S$ has a well-defined square root $S^{\frac{1}{2}}$. This is the unique positive operator $T$ such that $T^2 = S$.

Each invertible operator $T$ has *polar decomposition* $T = AU$, where $A$ is positive definite and $U$ is unitary. Both are uniquely determined by $T$: we have $A = (TT^*)^{\frac{1}{2}}$ and $U = (TT^*)^{-\frac{1}{2}}T$.

## §2 Fourier transformation

Good introductions to this topic are Dym & MacKean [1] and Stein & Weiss [5].

### 2.1 General properties

The *Fourier transform* $\mathcal{F}(f) = \widehat{f}$ of a function $f \in L^1(\mathbb{R}^d)$ is defined as

$$\widehat{f}(\xi) := \int_{\mathbb{R}^d} f(x)\, e^{-i\,\xi\cdot x}\, dx, \quad \xi \in \mathbb{R}^d. \tag{2.1}$$

Then $\widehat{f} \in C_0(\mathbb{R}^d)$, *i.e.*, $\widehat{f}$ is continuous and $\widehat{f}(\xi)$ tends to 0 as $|\xi|$ tends to $\infty$.

If $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ then $\widehat{f} \in L^2(\mathbb{R}^d)$ and *Plancherel's formula*

$$\int_{\mathbb{R}^d} |f(x)|^2 \, dx = (2\pi)^{-d} \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2 \, d\xi \tag{2.2}$$

is valid. Thus the Fourier transform $\mathcal{F}$ is an isometry from the dense subspace $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ of $L^2(\mathbb{R}^d)$ into $L^2(\mathbb{R}^d, (2\pi)^{-d} d\xi)$. This isometry has a unique continuous extension to an isometry $\mathcal{F}: f \mapsto \widehat{f}$ from $L^2(\mathbb{R}^d)$ into $L^2(\mathbb{R}^d, (2\pi)^{-d} d\xi)$. This defines the Fourier transform on $L^2(\mathbb{R}^d)$. In practice, if $f \in L^2(\mathbb{R}^d)$ then we find $\widehat{f}$ by first choosing a sequence of functions $f_n \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ such that $f_n \to f$ in $L^2(\mathbb{R}^d)$. Then $\widehat{f}_n \to \widehat{f}$ with convergence in $L^2(\mathbb{R}^d)$. One way to choose $f_n$ is to set $f_n := f \chi_{[-n,n]^d}$. Thus, if we write

$$\text{l.i.m.}_{n\to\infty} f_n(x) = f(x) \quad \text{when } f_n \to f \text{ converges in } L^2(\mathbb{R}^d)$$

then

$$\widehat{f}(\xi) = \text{l.i.m.}_{n\to\infty} \int_{[-n,n]^d} f(x) \, e^{-i\,\xi.x} \, dx \tag{2.3}$$

for $f \in L^2(\mathbb{R}^d)$.

The Plancherel formula (2.2) implies *Parseval's formula*

$$\int_{\mathbb{R}^d} f(x) \overline{g(x)} \, dx = (2\pi)^{-d} \int_{\mathbb{R}^d} \widehat{f}(\xi) \overline{\widehat{g}(\xi)} \, d\xi, \quad f, g \in L^2(\mathbb{R}^d). \tag{2.4}$$

The Fourier transform is a bijection from $L^2(\mathbb{R}^d)$ onto itself. It can be formally inverted by

$$(\mathcal{F}^{-1} g)(x) = (2\pi)^{-d} (\mathcal{F} g)(-x). \tag{2.5}$$

This means for $f \in L^2(\mathbb{R}^d)$ that

$$f(x) = (2\pi)^{-d} \, \text{l.i.m.}_{n\to\infty} \int_{[-n,n]^d} \widehat{f}(\xi) \, e^{i\,\xi.x} \, d\xi. \tag{2.6}$$

If, moreover, $\widehat{f} \in L^1(\mathbb{R}^d)$, this inversion formula can be written as

$$f(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} \widehat{f}(\xi) \, e^{i\,\xi.x} \, d\xi. \tag{2.7}$$

If $f \in L^1(\mathbb{R}^d)$ such that also $\widehat{f} \in L^1(\mathbb{R}^d)$ then both $f$ and $\widehat{f}$ are continuous and formulas (2.1) and (2.7) are valid.

The class $\mathcal{S} = \mathcal{S}(\mathbb{R}^d)$ of *rapidly decreasing $C^\infty$-functions* on $\mathbb{R}^d$ consists of all $C^\infty$-functions $\psi$ on $\mathbb{R}^d$ such that all partial derivatives $D^\alpha \psi(x)$ tend to 0 in absolute

value more rapidly than each inverse power of $|x|$ as $|x|$ tends to $\infty$. The space $\mathcal{S}(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$ and $L^1(\mathbb{R}^d)$. The Fourier transform maps $\mathcal{S}$ bijectively onto itself. If two functions $f$ and $g$ on $\mathbb{R}^d$, each being the product of a $L^1$-function and a polynomial on $\mathbb{R}^d$, satisfy

$$\int_{\mathbf{R}^d} f(x)\,\widehat{\psi}(x)\,dx = \int_{\mathbf{R}^d} g(x)\,\psi(x)\,dx \quad \text{for all } \psi \in \mathcal{S}, \tag{2.8}$$

then we call $g$ the *Fourier transform in the weak sense* of $f$. If $f \in L^2(\mathbb{R}^d)$ or $\in L^1(\mathbb{R}^d)$ then the ordinary Fourier transform $\widehat{f}$ is also the Fourier transform in the weak sense. Conversely we have:

**Lemma 2.1.** *Let $f$ and $g$ be functions on $\mathbb{R}^d$, each being the product of a $L^1$-function and a polynomial on $\mathbb{R}^d$. Suppose that they satisfy* (2.8). *Then, if $f \in L^1(\mathbb{R}^d)$ or $L^2(\mathbb{R}^d)$ then $g = \widehat{f}$.*

### 2.2 Translation, modulation and dilation

**Definition 2.2.** *Given a function $f \in L^2(\mathbb{R}^d)$ we define the following operators*

| | | |
|---|---|---|
| Translation: | $T_a f(x) = f(x-a)$, | for $a \in \mathbb{R}^d$ ; |
| Modulation: | $E_a f(x) = e^{ia\cdot x} f(x)$, | for $a \in \mathbb{R}^d$ ; |
| Dilation: | $D_a f(x) = \lvert a\rvert^{-d/2} f(x/a)$, | for $a \in \mathbb{R} \setminus \{0\}$ . |

Each of these is a unitary operator of $L^2(\mathbb{R}^d)$, and we have

$$(f, T_a g) = (T_{-a} f, g); \quad (f, E_a g) = (E_{-a} f, g); \quad (f, D_a g) = (D_{1/a} f, g).$$

In connection with the Fourier transform we have the following elementary formulas:

$$\widehat{T_a f} = E_{-a}\widehat{f}; \quad \widehat{E_a f} = T_a \widehat{f}; \quad \widehat{D_a f} = D_{1/a}\widehat{f}. \tag{2.9}$$

### 2.3 Convolution

The *convolution product* of two functions $f, g$ on $\mathbb{R}^d$ is given by

$$(f * g)(x) := \int_{\mathbf{R}^d} f(y)\,g(x-y)\,dy \tag{2.10}$$

whenever this formula has meaning. We mention two cases where this is the case. First, if $f \in L^1(\mathbb{R}^d)$ and $g \in L^p(\mathbb{R}^d)$ $(1 \le p \le \infty)$ then $f * g \in L^p(\mathbb{R}^d)$,

$$\|f * g\|_p \le \|f\|_1 \,\|g\|_p\,, \tag{2.11}$$

and, for $p = 1$ or 2,

$$(f * g)\widehat{\ }(\xi) = \widehat{f}(\xi)\,\widehat{g}(\xi). \tag{2.12}$$

Second, if $f, g \in L^2(\mathbb{R}^d)$ then $f * g \in C_0(\mathbb{R}^d)$ and

$$\|f * g\|_\infty \le \|f\|_2 \,\|g\|_2\,. \tag{2.13}$$

### 2.4 Mean and variance

We specialize to $d = 1$. Suppose that $f \in L^2(\mathbb{R})$ such that also $f'$ and $x \mapsto xf(x)$ are in $L^2(\mathbb{R})$. We define the *mean* and *variance* for the normalized probability distributions obtained from $|f(x)|^2\,dx$ and $|\widehat{f}(\xi)|^2\,d\xi$:

$$\mu(f) := (\|f\|_2)^{-2} \int_{\mathbf{R}} x\,|f(x)|^2\,dx, \tag{2.14}$$

$$\mu(\widehat{f}) = (\|\widehat{f}\|_2)^{-2} \int_{\mathbf{R}} \xi\,|\widehat{f}(\xi)|^2\,d\xi, \tag{2.15}$$

$$\sigma(f) := (\|f\|_2)^{-1} \left[\int_{\mathbf{R}} (x-\mu(f))^2\,|f(x)|^2\,dx\right]^{\frac{1}{2}}, \tag{2.16}$$

$$\sigma(\widehat{f}) = (\|\widehat{f}\|_2)^{-1} \left[\int_{\mathbf{R}} (\xi-\mu(\widehat{f}))^2\,|\widehat{f}(\xi)|^2\,d\xi\right]^{\frac{1}{2}}. \tag{2.17}$$

The interval $[\mu(f)-\sigma(f), \mu(f)+\sigma(f)]$ gives a qualitative indication of the set where $f$ attains its most substantial nonzero values. The interval $[\mu(\widehat{f})-\sigma(\widehat{f}), \mu(\widehat{f})+\sigma(\widehat{f})]$ plays a similar role for $\widehat{f}$. The two variances $\sigma(f)$ and $\sigma(\widehat{f})$ cannot be arbitrarily small. We have *Heisenberg's inequality* (cf. for instance Dym & McKean [1, §2.8])

$$\sigma(f)\,\sigma(\widehat{f}) \geq \tfrac{1}{2}. \tag{2.18}$$

In the quantum mechanical interpretation this means that position and impulse cannot be measured simultaneously in an arbitrarily precise way. If $\mu(f) = 0$ and $\mu(\widehat{f}) = 0$ then equality is attained in (2.18) iff $f(x) = \text{const.}\, e^{-\lambda x^2}$ for some $\lambda > 0$.

### 2.5 Poisson summation formula

This formula describes the relation between a series built from an underlying function and a series built from the Fourier transform of the same function. We specialize to $d = 1$.

**Theorem 2.3.** *Let $f$ be a rapidly decreasing $C^\infty$-function on $\mathbb{R}$. Then we have*

$$\sum_{n=-\infty}^{\infty} e^{-ina} f(x+nb) = \frac{1}{|b|} \sum_{m=-\infty}^{\infty} \widehat{f}\left(\frac{2\pi m + a}{b}\right) e^{ix(2\pi m+a)/b},$$

*where $a, b, x \in \mathbb{R}$, $b \neq 0$.*

A special case is

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{m=-\infty}^{\infty} \widehat{f}(2\pi m).$$

Similar forms for sine or cosine transforms follow from taking $f$ odd or even, respectively.

A nice example, that leads to the theory of theta functions, is the relation

$$\sum_{n=-\infty}^{\infty} e^{-\pi n^2 s} = \frac{1}{\sqrt{s}} \sum_{m=-\infty}^{\infty} e^{-\pi m^2 s^{-1}}, \quad s > 0,$$

which follows from the choice

$$f(x) = \exp(-\pi s x^2), \quad \widehat{f}(\xi) = \frac{1}{\sqrt{s}} e^{-\xi^2/(4\pi s)}.$$

## §3 Riesz bases and frames

Background information on this topic can be found in Young [6].

### 3.1 Riesz bases

**Definition 3.1.** *A Riesz basis of a Hilbert space $\mathcal{H}$ is a system $\{x_n\}$ in $\mathcal{H}$ such that there is an orthonormal basis $\{e_n\}$ of $\mathcal{H}$ and a bounded and boundedly invertible linear operator $T$ of $\mathcal{H}$ with the property that $Te_n = x_n$ for all $n$.*

**Proposition 3.2.** *A countable system $\{x_n\}$ in a Hilbert space $\mathcal{H}$ is a Riesz basis if and only if the two following properties hold.*

(a) *For every $x \in \mathcal{H}$ there are unique coefficients $c_n$ such that*

$$x = \sum_n c_n x_n \quad \textit{(unconditional convergence)}. \tag{3.1}$$

(b) *There are positive constants $A$, $B$ such that, for each $x \in \mathcal{H}$ with expansion (3.1),*

$$A \|x\|^2 \leq \sum_n |c_n|^2 \leq B \|x\|^2. \tag{3.2}$$

Riesz bases can of course be orthonormalized by the Gram-Schmidt process. This process heavily depends on the way the Riesz basis is arranged as a sequence. The next theorem describes a construction of orthonormal basis from a Riesz basis which does not depend on the arrangement and which may better pass properties of the vectors of the Riesz basis to the vectors of the orthonormal basis.

**Theorem 3.3.** *Let $\{x_n\}$ be a Riesz basis of a Hilbert space $\mathcal{H}$. Then there is a unique orthonormal basis $\{f_n\}$ such that $x_n = Af_n$ for some positive definite operator $A$.*

**Proof:** Certainly, $x_n = Te_n$ for some orthonormal basis $\{e_n\}$ and some invertible operator $T$. Now use the existence and uniqueness of the polar decomposition $T = AU$ (cf. §1.3, $A$ is positive definite, $U$ is unitary) and the one-to-one correspondence between unitary operators $U$ and orthonormal bases $\{f_n\}$ via $f_n = Ue_n$. ■

The following construction of an orthonormal basis from a Riesz basis is maybe better known. For convenience assume that the index $n$ in the Riesz basis $x_n$ runs over $\mathbb{Z}$. Then vectors $x \in \mathcal{H}$ can be identified with sequences $\{c_n\} \in \ell^2$ by (3.1). (However, this does not identify the inner products on $\mathcal{H}$ and $\ell^2$ with each other.) This also implies an identification of operators $X$ of $\mathcal{H}$ with operators $\mathbf{X}$ of $\ell^2$ (*i.e.*, with matrices wich have row and column indices running over $\mathbb{Z}$): If $\mathbf{X} = (X_{k,l})$ then $Xx_l = \sum_k X_{k,l}x_k$. Associate with $\{x_n\}$ the *Gram matrix* $\mathbf{G} = (G_{k,l})$ given by

$$G_{k,l} := (x_l, x_k). \tag{3.3}$$

This is a positive definite matrix of which we can take the square root. This yields another positive definite matrix $\mathbf{B} = (B_{k,l})$ and hence, by the above mentioned identification, a bounded operator $B$ on $\mathcal{H}$ with bounded inverse.

**Theorem 3.4.** *The operator $B$ is positive definite and $\{B^{-1}x_n\}$ is an orthonormal basis, the same as constructed in Theorem* 3.3.

**Proof:** By definition, $x_n = Te_n$ for some invertible operator $T$ and some orthonormal basis $\{e_n\}$. Hence

$$G_{k,l} = (x_l, Te_k) = (T^*x_l, e_k).$$

Then

$$Gx_l = \sum_k G_{k,l}x_k = \sum_k (T^*x_l, e_k)\, Te_k = T\Big(\sum_k (T^*x_l, e_k)\, e_k\Big) = TT^*x_l.$$

Hence $G = TT^*$, so $G$ is a positive definite operator and $B = G^{\frac{1}{2}}$ is also positive definite. Let $T$ have polar decomposition $T = AU$ with $A$ positive definite and $U$ unitary. Then $G = TT^* = A^2$. Hence $B = A = TU^{-1}$, so $B^{-1}x_l = UT^{-1}x_l = Ue_l$, so $\{B^{-1}x_l\}$ is an orthonormal basis since $\{e_l\}$ is an orthonormal basis and $U$ is unitary. Since $B$ is positive definite, the last part of the theorem follows from the uniqueness statement in Theorem 3.3. ■

**Remark 3.5.** Preserve notation of the last theorem and proof. Consider the operator $\widetilde{G}$ which has matrix $(G_{k,l})$ with respect to the orthonormal basis $\{e_n\}$. Then

$$\widetilde{G}e_l = \sum_k G_{k,l}e_k = \sum_k (T^*Te_l, e_k)\, e_k = T^*Te_l.$$

Hence $\widetilde{G} = T^*T$, while we have seen that $G = TT^*$.

### 3.2 Frames

**Definition 3.6.** *A system $\{x_n\}$ ($n$ running over a countable index set) in a Hilbert space $\mathcal{H}$ is a frame if there exist numbers $A, B > 0$ such that for all $x \in \mathcal{H}$ we have*

$$A\,\|x\|^2 \leq \sum_n |(x, x_n)|^2 \leq B\,\|x\|^2. \tag{3.4}$$

*The numbers $A, B$ are called frame bounds. The frame is tight if $A = B$. The frame is exact if it ceases to be a frame whenever any single element is deleted from the sequence.*

Notice that the $x_n$ are not necessarily linearly independent. The definition however, implies that the mapping $x \to \{(x, x_n)\}$ is injective. If $n$ runs over $\mathbb{Z}$ then Definition 3.1 can be equivalently formulated as saying that the mapping $x \to \{(x, x_n)\}$ is open and bounded from $\mathcal{H}$ to $\ell^2$.

**Remark 3.7.** Let $\mathcal{H}$ be a Hilbert space with orthonormal basis $\{e_1, e_2, e_3, \ldots\}$.

- $\{e_1, e_2/2, e_3/3, \ldots\}$ is an orthogonal sequence, not a frame.
- $\{e_1, e_2/\sqrt{2}, e_2/\sqrt{2}, e_3/\sqrt{3}, e_3/\sqrt{3}, e_3/\sqrt{3}, \ldots\}$ is a tight inexact frame.
- $\{e_1/\sqrt{2}, e_2/\sqrt{2}, (e_1 + e_2)/2, (e_1 - e_2)/2\}$ is a tight inexact frame in $\mathbb{C}^2$.
- $\{2e_1, e_2, e_3, \ldots\}$ is a non-tight exact frame with $A = 1$, $B = 4$.
- In a finite dimensional Hilbert space the exact frames are precisely the bases.

Any Riesz basis $\{x_n\}$ is a frame. Indeed, let $x_n = Te_n$ with $\{e_n\}$ being an orthonormal basis. Then

$$(x, x_k) = (x, Te_k) = (T^*x, e_k).$$

Hence

$$\sum_k |(x, x_k)|^2 = \sum_k |(T^*x, e_k)|^2 = (T^*x, T^*x) = (TT^*x, x).$$

Now (3.4) is evidently satisfied. It is also quickly seen that this frame is exact. It can conversely be shown that any exact frame is a Riesz basis.

Let $\{x_n\}$ be a frame. Define the *frame operator* $S$ by

$$Sx = \sum_k (x, x_k)x_k, \quad x \in \mathcal{H}. \tag{3.5}$$

Then $(Sx, x) = \sum_k |(x, x_k)|^2$ and formula (3.4) is equivalent to stating that $S$ is a positive definite operator satisfying $AI \leq S \leq BI$. Note that $S = TT^*$ in the above example of Riesz bases.

**Theorem 3.8.** *Let $\{x_n\}$ be a frame in $\mathcal{H}$ and let $S$ be the corresponding frame operator.*
(1) *$S$ is invertible and $B^{-1}I \leq S^{-1} \leq A^{-1}I$.*
(2) *$\{S^{-1}x_n\}$ is a frame with bounds $B^{-1}, A^{-1}$, called the dual frame of $\{x_n\}$.*
(3) *Every $x \in \mathcal{H}$ can be written*

$$x = \sum (x, x_n) S^{-1} x_n = \sum (x, S^{-1} x_n) x_n.$$

This theorem shows that, for a set of elements $\{x_n\}$, the frame property implies reconstructability. Via two intermediate results it is next shown in Heil & Walnut [2] that the 'only if' part of the following theorem is valid, while the 'if' part is obvious.

**Theorem 3.9.** *Let $\{x_n\}$ be a frame in $\mathcal{H}$. Then it is an exact frame if and only if $\{x_n\}$ and $\{S^{-1}x_n\}$ are biorthonormal, i.e., if and only if*

$$(x_m, S^{-1} x_n) = \delta_{m,n}.$$

### 3.3 Approximation by frames

Given a frame $\{x_n\}$, we know that the mapping $f \to \{(f, x_n)\}$ is injective, therefore we may try to recover $f$ from a sequence $\{(f, x_n)\}$. In fact we find the following theorem.

**Theorem 3.10.** *Let $\{x_n\}$ be a frame with frame bounds given by (3.4), then*

$$f = \frac{2}{A+B} \sum_n (f, x_n) x_n + Rf, \tag{3.6}$$

*where the operator $R$ satisfies the error bound*

$$\|R\| \leq \frac{B-A}{B+A}. \tag{3.7}$$

**Proof:** We can write (3.6) as

$$f = \frac{2}{A+B} Sf + Rf,$$

where $S$ is given by (3.5). Hence

$$Rf = (I - \frac{2}{A+B} S) f.$$

Thus

$$\|Rf\| \le \|I - \frac{2}{A+B} S\| \, \|f\|.$$

Since $AI \le S \le BI$, we have

$$-\frac{B-A}{B+A} I \le I - \frac{2}{A+B} S \le \frac{B-A}{B+A} I.$$

This implies (3.7). ■

**Remark 3.11.**

- Clearly, if the frame is tight, then $\|R\| = 0$ and the approximation is exact.
- Suppose that we had defined $Rf$ by $f = cSf + Rf$ with $c > 0$. Then we would have arrived at $Rf = (I - cS)f$ and $(1 - cB)I \le I - cS \le (1 - cA)I$. Hence $f = cSf + Rf$ with $c > 0$. Then we would have obtained $Rf = (I - cS)f$ and $(1 - cB)I \le I - cS \le (1 - cA)I$. Hence

  $$\|I - cS\| \le \max\{|1 - cA|, |1 - cB|\}.$$

  The right hand side of the above inequality, as a function of $c$, achieves its maximum for $c := 2/(A + B)$.
- If the approximation is not exact, then we have $0 < A < B$ and $(B - A)/(A + B) < 1$. Therefore we can apply the approximation formula recursively to obtain corrections to the approximation. The error is shrinking, and this iterative procedure converges to $f$. Indeed, write $Xf := 2(A + B)^{-1}Sf$. Then

  $$f = Xf + Rf = Xf + XRf + R^2 f = \cdots = X(1 + R + \cdots + R^{n-1})f + R^n f$$

  and $\|R^n f\| \to 0$. This can be written in algorithmic form as follows. First put

  $$f^{(1)} = \frac{2}{A+B} \sum_n (f, x_n)\, x_n,$$

  and then, for $i = 1, 2, \ldots$,

  $$\begin{aligned} e_n^{(i)} &= (f, x_n) - (f^{(i)}, x_n), \\ e^{(i)} &= \frac{2}{A+B} \sum_n e_n^{(i)} x_n, \\ f^{(i+1)} &= f^{(i)} + e^{(i)}. \end{aligned}$$

**Acknowledgments.** Work for this paper by the second author was partially done at CWI, Amsterdam.

## References

1. Dym, H. and H. P. Mckean, *Fourier Series and Integrals*, Academic Press, New York, 1972.
2. Heil, C. E. and D. F. Walnut, Continuous and discrete wavelet transforms, *SIAM Review* **31** (1989), 628–666.
3. Rudin, W., *Functional Analysis*, McGraw-Hill, New York, 1973.
4. Rudin, W., *Real and Complex Analysis*, McGraw-Hill, New York, 1974.
5. Stein, E. M. and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton, 1971.
6. Young, R. M., *An Introduction to Nonharmonic Fourier Series*, Academic Press, New York, 1980.

*Pieter W. Hemker* and *Nico M. Temme*
CWI
P.O. Box 4079
1009 AB Amsterdam
The Netherlands
pieth@cwi.nl, nicot@cwi.nl

*Tom H. Koornwinder*
University of Amsterdam
Faculty of Mathematics and Computer Science
Plantage Muidergracht 24
1018 TV Amsterdam
The Netherlands
thk@fwi.uva.nl