Mathematics of Signal Representations

Math 4355 – Course Notes

Bernhard G. Bodmann
bgb@math.uh.edu

Contents

Chapter 1

Inner Product Spaces

Defining properties and examples

1.1 Definition. *An* ***inner product*** *for a complex vector space V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ which is sesqui-linear and positive definite. This means, it has the following properties:*

1. $\overline{\langle v, w \rangle} = \langle w, v \rangle$ *for all* $v, w \in V$;
2. $\langle cv, w \rangle = c\langle v, w \rangle$ *for all* $v, w \in V$ *and* $c \in \mathbb{C}$;
3. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ *for all* $u, v, w \in V$;
4. $\langle v, v \rangle > 0$ *for all* $v \in V$, $v \neq 0$.

When a vector space V has been equipped with an inner product, we also refer to it as an ***inner product space****. We also define the* ***norm*** $\|v\| = \sqrt{\langle v, v \rangle}$ *for all $v \in V$. A sequence $\{v_n\}_{n \in \mathbb{N}}$* ***converges*** *to a vector w* ***in norm*** *if* $\lim_{n \to \infty} \|v_n - w\| = 0$.

1.2 Example. The vector space of all trigonometric polynomials, given by the set of functions

$$V = \left\{ p : [0,1] \to \mathbb{C}, p(t) = \sum_{k=-N}^{N} c_k e^{2\pi i k t}, N \in \mathbb{N}, \text{ all } c_k \in \mathbb{C} \right\} \tag{1.1}$$

can be equipped with the inner product

$$\langle v, w \rangle = \int_0^1 v(t)\overline{w(t)}\, dt \,. \tag{1.2}$$

The sesqui-linearity (Properties 1 to 3) follows from the linearity of the integral. To check the positive definiteness, we compute the square norm for a trigonometric polynomial $v(t) = \sum_{k=-N}^{N} c_k e^{2\pi ikt}$ with degree $2N+1 \in \mathbb{N}$,

$$\begin{aligned}
\langle v, v\rangle &= \int_0^1 |v(t)|^2\, dt \\
&= \int_0^1 \sum_{k=-N}^{N} c_k e^{2\pi ikt} \sum_{l=-N}^{N} \overline{c_l} e^{-2\pi ilt}\, dt \\
&= \sum_{k,l=-N}^{N} c_k \overline{c_l} \int_0^1 e^{2\pi i(k-l)t}\, dt = \sum_{k=-N}^{N} |c_k|^2 .
\end{aligned}$$

The last sum is zero if and only if $c_k = 0$ for all $k \in \{-N, -N+1, \ldots, N-1, N\}$, which means that $v(t) = 0$ for all $t \in [0,1]$.

The example of trigonometric polynomials is a vector space that does not have a finite basis, that is, a finite, linearly independent set for which finite linear combinations can produce any vector in V. This is a simple consequence of the fact that a finite set of trigonometric polynomials has a maximal degree. Any monomial with a higher degree cannot be obtained from a linear combination within this set.

1.3 Exercise. Recall the Cauchy property of sequences. A Cauchy sequence $\{v_n\}_{n\in\mathbb{N}}$ in a normed vector space satisfies that for any $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that for all $m, n > N$, $\|v_n - v_m\| < \epsilon$. Show that the space of trigonometric polynomials is not closed, that is, there are sequences of polynomials which have the Cauchy property with respect to the norm induced by the inner product, but they do not converge to a polynomial.

To remedy this problem, one could identify each polynomial with the (finite) sequence of its coefficients, and define an inner product in terms of the coefficients. This way, polynomials are embedded in the larger space of square-summable sequences. We will show in Exercise 1.10 that all Cauchy sequences converge in this larger space.

1.4 Example. Let $l^2(\mathbb{Z})$ be the vector space of all sequences $(x_n)_{n\in\mathbb{Z}}$ with $\sum_{k=-\infty}^{\infty} |x_n|^2 < \infty$. For $x, y \in l^2(\mathbb{Z})$, we define

$$\langle x, y\rangle = \sum_{n=-\infty}^{\infty} x_n \overline{y_n} .$$

We also denote $\|x\| = \sqrt{\langle x, x\rangle}$.

To see that the inner product is indeed defined on all pairs of vectors from $l^2(\mathbb{Z})$, we note that for $x, y \in l^2(\mathbb{Z})$, the series for the inner product is term by term dominated by an absolutely convergent series,

$$\left| \sum_{k=-\infty}^{\infty} x_k \overline{y_k} \right| \leq \sum_{k=-\infty}^{\infty} |x_k y_k| \leq \sum_{k=-\infty}^{\infty} \left(\frac{1}{2}|x_k|^2 + \frac{1}{2}|y_k|^2 \right) .$$

Thinking of a trigonometric polynomial as a sequence of coefficients, finitely many of which are nonzero, motivates to consider 'functions' corresponding to sequences of coefficients which are merely square-summable. Such functions could then be thought of as limits of Cauchy sequences of trigonometric polynomials (obtained from truncating the coefficients). The question of whether these limits can indeed be interpreted as functions, and in which precise sense they are limits of trigonometric polynomials is the central theme of the next chapter on Fourier series. Writing the inner product for these limits in the same form as for trigonometric polynomials motivates the informal definition of $L^2([0,1])$, the space of square-integrable functions on $[0,1]$. We can make this definition more general by using complex exponentials of the form $e^{2\pi int/(b-a)}$, and obtain the space of square-integrable functions on an interval $[a,b]$.

1.5 Definition. *Let $a, b \in \mathbb{R}$, $a < b$, then we define the vector space of square-integrable functions*

$$L^2([a,b]) = \left\{ f : [a,b] \to \mathbb{C}, f(t) = \sum_{k=-\infty}^{\infty} c_k e^{2\pi ikt/(b-a)}, c \in l^2(\mathbb{Z}) \right\}$$

and for two such square-integrable functions f and g, we write

$$\langle f, g \rangle = \int_a^b f(t)\overline{g(t)} dt .$$

1.6 Remark. This cannot define an inner product for *functions*, because if $f(a) = 1$, $f(t) = 0$ for all $a < t \leq b$, then we have $\langle f, f \rangle = 0$ but f is not the zero function! However, one can show that the inner product space obtained from Cauchy sequences of trigonometric polynomials amounts to identifying two functions when they differ on a set that does not contribute in an integral. In this case, we say that the two functions are equal **almost everywhere**.

1.7 Example. Sets that do not contribute in integrals are those that can be covered with an at most countable number of intervals having a total length that can be made arbitrarily small.

One example is the set $\mathbb{Q}$ containing all rational numbers. Since these numbers are countable, we can enumerate them with a sequence $\{q_n\}_{n\in\mathbb{N}}$. Given $\epsilon > 0$, choosing intervals of length 2^{-n} centered at each q_n covers the rationals, and the total length of all intervals is $\sum_{j=1}^{\infty} \epsilon/2^n = \epsilon$. In fact, this construction applies to any countable set, which shows that none of them contribute in integrals.

Another example is the so-called Cantor set. It is given as an intersection of countably many sets obtained from an iterative procedure. The first set is $C_1 = [0,1]$. The next is obtained by removing the middle third, $C_2 = [0,1/3] \cup [2/3,1]$. At each step, we remove the middle third. The total length of the intervals contained in C_n is thus $(2/3)^{n-1}$, which converges to zero. Each number in $C = \cap_{n=1}^{\infty} C_n$ is uniquely determined by the infinite sequence of binary decisions keeping track of which "third" (left or right) contains the number when passing from C_{n-1} to C_n. Therefore, the set C is not countable, as proved by Cantor's diagonal argument. If they were, we could write the binary sequences underneath each other, and then create another sequence by picking the numbers on the diagonal. Switching all "0"s and "1"s then creates a sequence that is different from all of the enumerated ones, thus the enumeration cannot contain all binary sequences.

The space of sequences can be thought of as the space of digitized signals, given by coefficients stored in a computer. The space of square-integrable functions, on the other hand, can be thought of as the space of analog signals. By identifying trigonometric polynomials with their sequences of coefficients, we have tacitly introduced a map between analog and digital signals which is compatible with the inner products on both spaces. We will investigate this map more closely.

Inequalities

Two fundamental inequalities that hold on inner product spaces are the Cauchy-Schwarz inequality and the triangle inequality.

1.8 Theorem. *If V is a vector space with inner product $\langle \cdot, \cdot \rangle$, then for all $x, y \in V$*

$$|\langle x, y\rangle| \leq \|x\|\|y\| . \tag{1.3}$$

1.9 Theorem. *If V is a vector space with inner product $\langle \cdot, \cdot \rangle$, then for all $x, y \in V$,*

$$\|x + y\| \leq \|x\| + \|y\| . \tag{1.4}$$

1.10 Exercise. Show that each Cauchy sequence in $l^2(\mathbb{Z})$ converges in norm to a square-summable sequence.

Orthogonality and basis expansions

1.11 Definition. *Let V be a vector space with an inner product. We say that two vectors $x, y \in V$ are orthogonal, abbreviated $x \perp y$, if $\langle x, y\rangle = 0$. A set $\{e_1, e_2, \dots e_N\}$ is called orthonormal if $\|e_i\| = 1$ and $\langle e_i, e_j\rangle = 0$ for all $i \neq j$. We abbreviate this with Kronecker's δ-symbol as $\langle e_i, e_j\rangle = \delta_{i,j}$. We then call $\{e_1, e_2, \dots e_N\}$ an orthonormal basis for its linear span. Given an infinite orthonormal set $\{e_n\}_{n\in\mathbb{Z}}$, we say that it is an orthonormal basis for all vectors that are obtained from summing the basis vectors with a square-summable sequence of coefficients. Finally, two subspaces V_1, V_2 are called orthogonal, abbreviated $V_1 \perp V_2$, if all pairs (x, y) with $x \in V_1$ and $y \in V_2$ are orthogonal.*

1.12 Example. Let V_0 be the complex subspace of $L^2([-\pi, \pi])$ given by

$$V_0 = \{f(x) = c_1 \cos x + c_2 \sin x \text{ for } c_1, c_2 \in \mathbb{C}\}\,.$$

Then the set $\{e_1, e_2\}$,

$$e_1(x) = \frac{1}{\sqrt{\pi}} \cos x \text{ and } e_2(x) = \frac{1}{\sqrt{\pi}} \sin x\,,$$

is an orthonormal basis for V_0. Strictly speaking, a vector in this subspace specified by c_1 and c_2 is not the function

$$f(x) = c_1 \cos x + c_2 \sin x$$

but the equivalence class of all functions that are equal to f for almost every $x \in [-\pi, \pi]$. However, to simplify notation, we will take the liberty to speak of each function as if it were the vector given by its equivalence class.

Another subspace of of $L^2([0, 1])$ is the space of functions which are almost everywhere constant on $[0, 1/2)$ and $[1/2, 1]$. It has the orthonormal basis $\{\phi, \psi\}$ with

$$\phi(x) = 1 \text{ and } \psi(x) = \begin{cases} 1, & 0 \le x < 1/2 \\ -1, & 1/2 \le x \le 1 \end{cases}$$

The normalization is straightforward to check. The orthogonality can be verified by splitting the domain of the integral in the inner product,

$$\int_0^1 \phi(t)\psi(t)dt = \int_0^{1/2} 1dt + \int_{1/2}^1 (-1)dt = 0\,.$$

1.13 Theorem. *Let V_0 be a subspace of an inner product space V, and $\{e_1, e_2, \dots e_N\}$ an orthonormal basis for V_0. Then for all $v \in V_0$,*

$$v = \sum_{k=1}^{N} \langle v, e_k \rangle e_k \, .$$

Proof. Since $\{e_k\}_{k=1}^N$ is a basis for V_0 as a vector space, we can write

$$v = \sum_{k=1}^{N} \alpha_k e_k$$

with some unique choice of coefficients $\{\alpha_j\}_{j=1}^N$. In order to isolate the value of each coefficient, we take the inner product with e_k, $k \in \{1, 2, \dots N\}$, on each side of this identity, and use the linearity of the inner product as well as the orthonormality of the basis,

$$\langle v, e_k \rangle = \sum_{l=1}^{N} \alpha_l \langle e_l, e_k \rangle = \alpha_k \, .$$

□

Orthogonal projections

1.14 Question. What is the result

$$\hat{v} = \sum_{k=1}^{N} \langle v, e_k \rangle e_k$$

if $v \notin V_0$?

1.15 Theorem. *Let V_0 be an inner product space, V_0 an N-dimensional subspace with an orthonormal basis $\{e_1, e_2, \dots e_N\}$. Then for $v \in V$,*

$$\hat{v} = \sum_{j=1}^{N} \langle v, e_k \rangle e_k$$

satisfies

$$\langle v - \hat{v}, w_0 \rangle = 0$$

for all $w_0 \in V_0$.

Proof. Since $w_0 = \sum_{k=1}^{N} \beta_j e_k$ with some coefficients $\{\beta_k\}_{k=1}^{N}$ and the inner product is linear, we only need to check that for all indices k,

$$\langle v - \hat{v}, e_k \rangle = 0 \,.$$

This is true because of orthonormality of the basis,

$$\langle v - \sum_{l=1}^{N} \langle v, e_l \rangle e_l, e_k \rangle = \langle v, e_k \rangle - \sum_{l=1}^{N} \langle v, e_l \rangle \langle e_l, e_k \rangle = 0 \,.$$

□

Since the difference vector $v - \hat{v}$ is orthogonal to V_0, we call $\hat{v}$ the **orthogonal projection** of v onto V_0.

1.16 Exercise. Let ϕ and ψ be the functions in $L^2([0,1])$ as defined in Example 1.12. Project the function $f(x) = x$ onto the space V_0 for which ϕ and ψ form an orthonormal basis.

If a vector x in an inner product space V is perpendicular to all $y \in V_0$, we write $y \perp V_0$ or $y \in V_0^{\perp}$.

1.17 Theorem. *Let V_0 be a finite dimensional subspace of an inner product space V. Then each $v \in V$ has a unique way of being expressed as*

$$v = v_0 + v_1 \,,$$

where $v_0 \in V_0$ and $v_1 \perp V_0$. We write $V = V_0 \oplus V_0^{\perp}$.

Proof. Take v and project orthogonally onto V_0. Let $v_1 = v - v_0$, then $v = v_0 + v_1$ and $v_1 \in V_0^{\perp}$ by the preceding theorem. Conversely, given v_0 and v_1 with these properties, then v_0 must be the orthogonal projection of v onto V_0. □

A least squares algorithm

1.18 Theorem. *Let V_0 be a finite-dimensional subspace of an inner product space V. Then for any $v \in V$, its orthogonal projection $\hat{v}$ onto V_0 has the least-squares property*

$$\|v - \hat{v}\|^2 = \min_{w \in V_0} \|v - w\|^2 \,.$$

Proof. Consider for given $w \in V_0$ the square-distance function

$$f(t) = \|\hat{v} + tw - v\|^2, t \in \mathbb{R}$$

Since $\hat{v} - v$ and w are orthogonal,

$$\begin{aligned} f(t) &= \langle \hat{v} + tw - v, \hat{v} + tw - v \rangle \\ &= \langle \hat{v} - v, \hat{v} - v \rangle + t^2 \langle w, w \rangle \\ &= \|\hat{v} - v\|^2 + t^2 \|w\|^2 \end{aligned}$$

and the minimum is achieved at $t = 0$. This means that $\hat{v}$ is the least squares approximation. □

1.19 Theorem. *Let V be an inner product space, V_0 be a finite-dimensional subspace spanned by a vector-space basis $\{z_1, z_2, \dots z_q\}$ Given $y \in V$, then its orthogonal projection $\hat{y}$ onto V_0 has the unique expansion*

$$\hat{y} = \sum_{k=1}^{q} \alpha_k z_k$$

with coefficients $\{\alpha_k\}_{k=1}^{q}$ which solve the linear system

$$\langle y, z_l \rangle = \sum_{k=1}^{q} \alpha_k \langle z_k, z_l \rangle$$

for all $l \in \{1, 2, \dots q\}$.

1.20 Theorem. *Let V be an inner product space with finite-dimensional, mutually orthogonal subspaces V_1 and V_2. Given $y \in V$, then its orthogonal projection $\hat{y}$ onto $V_1 \oplus V_2$ is $\hat{y} = y_1 + y_2$, where y_1 and y_2 are the orthogonal projections onto V_1 and V_2.*

1.21 Remark. This means that introducing an additional subspace V_2 that is orthogonal to V_1 improves the approximation of the vector y by summing its orthogonal projections onto V_1 and V_2.

There is no need to re-compute the coefficients for the approximation in V_1 when V_2 is introduced.

Chapter 2

Fourier Series

Fourier series as expansion in an orthonormal basis

2.1 Exercise. Given $V_0 \subset L^2([0,\pi])$ which has the orthonormal basis $\{e_j\}_{j=1}^5$ of functions $e_j(x) = \sqrt{\frac{2}{\pi}}\sin(jx)$. Compute the projection of the constant function $f(x) = C$, $C \in \mathbb{R}$, onto V_0.

2.2 Theorem. *The set* $\{\ldots, \frac{\cos(2x)}{\sqrt{\pi}}, \frac{\cos(x)}{\sqrt{\pi}}, \frac{1}{\sqrt{2\pi}}, \frac{\sin(x)}{\sqrt{\pi}}, \frac{\sin(2x)}{\sqrt{\pi}}, \ldots\}$ *is an orthonormal set in* $L^2([-\pi,\pi])$.

2.3 Theorem. *If a function is given as a series,*

$$f(x) = a_0 + \sum_{k=1}^{\infty}(a_k\cos(kx) + b_k\sin(kx))$$

which converges with respect to the norm in $L^2([-\pi,\pi])$, *then*

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx\,,$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)dx\,,$$

and

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)dx\,.$$

2.4 Theorem. *If f is an even, square integrable function given in the form of a series as in the preceding theorem, then $b_n = 0$ for all $n \in \mathbb{N}$. If f is odd, then $a_n = 0$ for all $n \in \mathbb{N}$.*

2.5 Exercise. With the help of a change of variables, $y = a + (b-a)x/(2\pi)$, find an expression for the coefficients of

$$f(x) = a_0 + \sum_{k=1}^{\infty} (a_k \cos(2\pi k x/(b-a)) + b_k \sin(2\pi k x/(b-a)))\,.$$

For an integrable function f on $[-\pi, \pi]$, one could define the coefficients $\{a_n\}_{n=0}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ as in Theorem 2.3. The question is then: Does the Fourier series with these coefficients converge, and in which sense?

Types of convergence

Identifying vectors in $L^2([a,b])$ with functions motivates several different notions of convergence.

2.6 Definition. *A sequence $\{f_n\}_{n=1}^{\infty}$ in $L^2([a,b])$ converges* ***in the square mean*** *to $f \in L^2([a,b])$ if $\|f_n - f\| \to 0$. The convergence is* ***pointwise*** *if for all $t \in [a,b]$, $\lim_{n\to\infty} f_n(t) = f(t)$. It is* ***uniform*** *if*

$$\sup_{x\in[a,b]} |f_n(x) - f(x)| \to 0\,.$$

2.7 Exercise. Find sequences of functions on $[0,1]$ with either of the following convergence properties

1. $f_n \to 0$ in the square mean, but not pointwise.
2. $f_n \to 0$ pointwise, but not in the square mean.
3. $f_n \to 0$ pointwise and in the square mean, but not uniformly.

2.8 Exercise. Does the sequence of functions $\{f_n\}_{n=1}^{\infty}$ with values $f_n(x) = nx^n e^{-nx}$ for $x \in \mathbb{R}$ converge uniformly on $[-\pi, \pi]$?

Convergence of Fourier series

2.9 Lemma. *If f is piecewise continuous and bounded on $[a,b]$, then*

$$\lim_{k\to\infty} \int_a^b f(x)\cos(kx)\,dx = \lim_{k\to\infty} \int_a^b f(x)\sin(kx)dx = 0\,.$$

2.10 Theorem. *Assuming f is 2π-periodic, piecewise continuous and bounded, and $f'(x)$ exists for some $x \in [-\pi, \pi]$, then the Fourier series*

$$S_N(x) = a_0 + \sum_{k=1}^{N} \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx))$$

converges to

$$\lim_{N \to \infty} S_N(x) = f(x) .$$

2.11 Theorem. *Assuming f is 2π-periodic, piecewise continuous and bounded, left and right differentiable at $x \in [-\pi, \pi]$, then the Fourier series*

$$S_N(x) = a_0 + \sum_{k=1}^{N} \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx))$$

converges to

$$\lim_{N \to \infty} S_N(x) = \frac{1}{2}(\lim_{t \to x^-} f(t) + \lim_{t \to x^+} f(t))$$

What if we do this for a function f which is only defined on $[-\pi, \pi]$, which is left differentiable at π and right differentiable at $-\pi$? The series then converges to the periodic extension of f.

2.12 Definition. *The **periodic extension** of f defined on $[-\pi, \pi)$ is the function g such that $g(x) = f(x)$ for $-\pi \leq x < \pi$ and $g(x + 2\pi) = g(x)$ for all $x \in \mathbb{R}$.*

2.13 Exercise. Compute the Fourier coefficients for $f(x) = x$ on $[-\pi, \pi)$. verify that at the jump discontinuity, the Fourier series converges to the average of the left and right hand side limits.

Uniform convergence

2.14 Remark. Since each partial sum of a Fourier series is a trigonometric polynomial (a continuous function), if the Fourier series converges uniformly, then the limit is also a continuous function.

This is the motivation for studying uniform convergence.

2.15 Theorem. *If the Fourier coefficients $\{a_n, b_n\}$ of a function satisfy*

$$\sum_{n=1}^{\infty} (|a_n| + |b_n|) < \infty$$

then the series converges uniformly.

2.16 Corollary. *If f is periodic, continuous, twice continuously differentiable on $(-\pi, \pi)$ and f'' is a bounded function,*

$$\sup_{x\in[-\pi,\pi]} |f''(x)| \leq M, M > 0$$

then the Fourier series of f converges uniformly to f.

Convergence in square mean

2.17 Theorem. *Let f be square integrable on $[-\pi, \pi]$, then the partial sums of the Fourier series*

$$S_N(x) = a_0 + \sum_{k=1}^{N}\sum_{k=1}^{\infty}(a_k \cos(kx) + b_k \sin(kx))$$

converge in square mean to f,

$$\lim_{N\to\infty} \int_{-\pi}^{\pi} |(f - S_N)(x)|^2 dx = 0\,.$$

2.18 Theorem. *Let f be square integrable on $[-\pi, \pi]$, and*

$$f(x) = a_0 + \sum_{k=1}^{\infty}(a_k \cos(kx) + b_k \sin(kx))$$

then we have the equality

$$\frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx = 2|a_0|^2 + \sum_{k=1}^{\infty}(|a_k|^2 + |b_k|^2)\,.$$

In complex notation, if

$$f(x) = \sum_{k=-\infty}^{\infty} \alpha_k e^{ikx}$$

then

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx = \sum_{k=-\infty}^{\infty} |\alpha_k|^2\,.$$

2.19 Corollary. *If f and g are square integrable on $[-\pi, \pi]$,*

$$f(x) = \sum_{k=-\infty}^{\infty} \alpha_k e^{ikx}\,,$$

and

$$g(x) = \sum_{k=-\infty}^{\infty} \beta_k e^{ikx},$$

then

$$\langle f, g\rangle = 2\pi \sum_{k=-\infty}^{\infty} \alpha_k \overline{\beta_k}.$$

Thus, the map from $L^2([-\pi,\pi])$ to $l^2(\mathbb{Z})$ which maps a function f to its Fourier coefficients preserves inner products.

Gibbs phenomenon

2.20 Exercise. Consider the function

$$f(x) = \begin{cases} \pi - x, & 0 \le x \le \pi \\ -\pi - x, & -\pi \le x < 0 \end{cases}.$$

1. Compute the Fourier series of f.
2. Denote the Nth partial sum of the Fourier series by S_N, and let $g_N(x) = S_N(x) - f(x)$. Using the formula for the Dirichlet kernel, show that $g_N'(x) = \frac{\sin((N+1/2)x}{2\pi\sin(x/2)}$.
3. Compute the value of g_N at the first critical point to the right of $x = 0$.
4. Express the limit of this value as $N \to \infty$ in the form of an integral.

Step-function approximation

2.21 Definition. *We call intervals of the form $[k2^{-j}, (k+1)2^{-j})$, $j, k \in \mathbb{Z}$* ***half-open, dyadic intervals****. For $j \in \mathbb{Z}$, let $V_j([0,1])$ denote the space of functions which are constant on each dyadic interval of length 2^{-j} contained in $[0,1]$. If we identify each function in $V_j([0,1])$ with all the functions that are almost everywhere equal to it, then we can think of $V_j([0,1])$ as a subspace of $L^2([0,1])$.*

2.22 Proposition. *Let f be a square integrable function on $[0,1]$. The projection P_jf onto $V_j([0,1])$, $j \in \{0,1,2,\dots\}$ is specified by the values*

$$P_jf(k2^{-j}) = 2^j \int_{k2^{-j}}^{(k+1)2^{-j}} f(x)dx, 0 \le k \le 2^j - 1.$$

2.23 Remark. The approximation of f by $P_j f$ amounts to piecewise averaging of f on dyadic intervals of a given length. For this reason, there are no overshoots, and there is no Gibbs phenomenon! The price we pay is that $P_j f$ is not continuous, unless f is constant.

One could ask if there is a way to preserve smoothness and avoid the occurrence of the Gibbs phenomenon. We will see a way to approximate functions by projecting on spaces with a degree of smoothness that can be chosen to be "between" that of the piecewise constant functions and the bandlimited ones. These approximation spaces will be discussed in the chapter on multiresolution analysis. Numerical experiments with these approximations show that an increase in the smoothness of these spaces leads to a re-emergence of the Gibbs phenomenon.

Chapter 3

Fourier Transform

Definition and elementary properties

3.1 Fact. If $f \in L^2(\mathbb{R})$, then

$$\hat{f}(\omega) = \lim_{L\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-L}^{L} f(t)e^{-i\omega t}dt$$

exists for almost all $\omega \in \mathbb{R}$, that is, up to a set which does not count under the integral. Moreover, $\hat{f} \in L^2(\mathbb{R})$ and

$$f(t) = \lim_{\Omega\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} \hat{f}(\omega)e^{i\omega t}d\omega\,,$$

again, up to a set of $t \in \mathbb{R}$ which does not count in integrals.

3.2 Theorem (Plancherel). *Let $f, g \in L^2(\mathbb{R})$. Then denoting $F[f] = \hat{f}$ and $F[g] = \hat{g}$, we have*

$$\langle F[f], g\rangle = \langle f, F^{-1}[g]\rangle$$

3.3 Corollary. *Choosing $g = F[h]$, $h \in L^2(\mathbb{R})$, we obtain*

$$\langle F[f], F[h]\rangle = \langle f, F^{-1}[F[h]]\rangle = \langle f, h\rangle\,.$$

So, we have preservation of the norm and, by the polarization identity, of the inner product under the Fourier transform,

$$\|F[f]\|^2 = \|f\|^2\,.$$

3.4 Proposition. *Let $f, h \in L^2(\mathbb{R})$, $h(t) = f(bt)$ for $b > 0$. Then $\hat{h}(\omega) = \frac{1}{b}\hat{f}(\frac{\omega}{b})$.*

3.5 Example. If

$$f(t) = \begin{cases} 1, & -\pi \leq t \leq \pi \\ 0, & \text{else} \end{cases}$$

then $h(t) = f(bt)$ has the Fourier transform

$$\hat{h}(\omega) = \sqrt{\frac{2}{\pi}} \frac{\sin(\pi\omega/b)}{\omega}.$$

3.6 Proposition. *Let $f, h \in L^2(\mathbb{R})$, $h(t) = f(t-a)$ for some $a \in \mathbb{R}$. Then $\hat{h}(\omega) = e^{-i\omega a}\hat{f}(\omega)$.*

3.7 Proposition. *Let $f \in L^2(\mathbb{R})$. If f is even, then so is $\hat{f}$. If f is odd, then the same holds for $\hat{f}$.*

Sampling and reconstruction

3.8 Definition. *A function $f \in L^2(\mathbb{R})$ is called* ***Ω-bandlimited*** *if $\hat{f}(\omega) = 0$ for almost all ω with $|\omega| > \Omega$.*

3.9 Remark. From Parseval's identity, $\hat{f} \in L^2(\mathbb{R})$, and by $\hat{f}$ vanishing outside of $[-\Omega, \Omega]$, the inequality $|\hat{f}(\omega)| \leq \frac{1}{4} + |\hat{f}(\omega)|^2$ implies

$$\int_{-\Omega}^{\Omega} |\hat{f}(\omega)| d\omega \leq \int_{-\Omega}^{\Omega} (\frac{1}{4} + |\hat{f}(\omega)|^2) d\omega = \frac{\Omega}{2} + \|\hat{f}\|^2 < \infty$$

which means $\hat{f}$ is (absolutely) integrable.

A consequence of this fact and the Fourier inversion is that

$$\lim_{s\to t} f(s) = \lim_{s\to t} \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{i\omega s} d\omega = \int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{i\omega t} d\omega = f(t),$$

by uniform convergence of $e^{i\omega s} \to e^{i\omega t}$ on $[-\Omega, \Omega]$. This means that f, unlike the usual vectors in $L^2(\mathbb{R})$, can be interpreted as a continuous function.

3.10 Theorem. *Let $f \in L^2(\mathbb{R})$ be Ω-bandlimited. Then*

$$f(t) = \sum_{k=-\infty}^{\infty} f(\frac{k\pi}{\Omega}) \frac{sin(\Omega t - k\pi)}{\Omega t - k\pi}$$

and the series on the right-hand side converges in the norm of $L^2(\mathbb{R})$ and uniformly on $\mathbb{R}$.

Convolutions and filters

3.11 Definition. *Let $f, g \in L^2(\mathbb{R})$. Then we denote the* **convolution** *of f and g by*

$$(f * g)(t) = \int_{-\infty}^{\infty} f(t-x)g(x)dx\,.$$

3.12 Example. Take

$$g(x) = \begin{cases} 1/a, & 0 \le x \le a \\ 0, & \text{else} \end{cases}$$

then for any integrable (or square-integrable) f,

$$(f * g)(t) = \int_0^a f(t-x)dx = \int_{t-a}^{t} f(x)dx\,.$$

3.13 Theorem. *Let f, g be integrable functions on $\mathbb{R}$. Then $f * g$ is again integrable and $F[f * g] = \sqrt{2\pi}\hat{f}\hat{g}$.*

If, in addition $f, g \in L^2(\mathbb{R})$, then $F^{-1}[\hat{f}\hat{g}] = \frac{1}{\sqrt{2\pi}} f * g$.

3.14 Remark. Convolving with an integrable function g on $\mathbb{R}$ amounts to multiplying with $\sqrt{2\pi}\hat{g}$ in the frequency domain.

3.15 Definition. *A* **filter** *on $L^2(\mathbb{R})$ is a linear map $L : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ for which there is a bounded function m on $\mathbb{R}$ such that for all $f \in L^2(\mathbb{R})$,*

$$F[Lf] = m\hat{f}\,,$$

or equivalently,

$$Lf = F^{-1}[m\hat{f}]\,.$$

The function m is called the **system function** *of the filter.*

3.16 Definition. *A linear map L on $L^2(\mathbb{R})$ is called time invariant if for all $f, g \in L^2(\mathbb{R})$ which are related by a time shift, $g(x) = f(x-a)$ with a constant $a \in \mathbb{R}$,*

$$Lg(x) = Lf(x-a)\,.$$

In short, it does not matter in which order L and the time shift are applied.

3.17 Proposition. *All filters on $L^2(\mathbb{R})$ are time invariant linear maps.*

3.18 Definition. *We say that a filter L_h which acts on a function f by convolution with an integrable function h has an* ***impulse response*** *h. If $h(t) = 0$ for all $t < 0$, we call L_h a* ***causal filter****.*

3.19 Remark. For a causal filter with integrable impulse response h, we note that if a signal $f(t)$ vanishes for all $t < t_0$, then so does

$$L_h f(t) = f * h(t) = \int_{-\infty}^{\infty} h(t-x)f(x)dx \int_{-\infty}^{\infty} h(-x)f(t+x)dx$$

because $h(-x) = 0$ when $x \geq 0$ and otherwise $f(t+x) = 0$ because $t+x < t_0$.

So, the filtered signal responds to the input, it never anticipates. The design of analog devices (e.g., RLC-circuits) can only provide causal filters. Digital signal processing (see the section "From analog to digital filters" below) has opened the possibility of using non-causal filters which are applied to a digitized (sampled) signal.

3.20 Exercise. Find the system function m for the filter

$$Lf(t) = \frac{1}{2}(f(t) + f(t-a)), \quad a \in \mathbb{R}.$$

3.21 Definition. *A* ***low-pass*** *filter on $L^2(\mathbb{R})$ is a filter with system function m which has the limits*

$$\lim_{\omega \to 0} m(\omega) = 1 \text{ and } \lim_{\omega \to \pm\infty} m(\omega) = 0.$$

3.22 Example (Ideal low-pass filter). Let the system function of a filter be given by

$$m(\omega) = \begin{cases} 1, & |\omega| \leq \Omega \\ 0, & \text{else} \end{cases}.$$

What is the impulse response of this filter? We compute

$$L_h f = f * h \text{ thus } F[L_h f] = \sqrt{2\pi}\hat{f}\hat{h}$$

so $\hat{h} = \frac{1}{\sqrt{2\pi}}$ and applying the inverse Fourier transform gives

$$h(t) = \frac{1}{2\pi}\int_{-\Omega}^{\Omega} e^{i\omega t} d\omega = \frac{1}{\pi\Omega t}\sin(\Omega t)$$

which is not integrable, but we can write

$$f * h(t) = \int_{-\infty}^{\infty} f(x)\frac{1}{\pi\Omega(t-x)}\sin(\Omega(t-x))\,dx$$

because f and h are in $L^2(\mathbb{R})$, and so the "convolution" is pointwise defined. The fact that $f * h \in L^2(\mathbb{R})$ is easy to see via the Plancherel identity, but not obvious in the point-wise expression as convolution.

One could investigate what happens when $\Omega \to \infty$.

3.23 Exercise. Is there a square-integrable function h such that $f * h = f$ for all $f \in L^2(\mathbb{R})$?

3.24 Example (Butterworth filter). The system function

$$m(\omega) = \frac{1}{1 + i\frac{\omega}{\Omega}}$$

is associated with

$$h(t) = \begin{cases} \Omega e^{-\Omega t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

which is the impulse response of a causal filter of the Butterworth type. Another example of a filter of this type has the system function

$$m(\omega) = \frac{1}{1 + 2i\frac{\omega}{\Omega} - 2\frac{\omega^2}{\Omega^2} - i\frac{\omega^3}{\Omega^3}} = \frac{1}{(1 + i\frac{\omega}{\Omega})(1 + i\frac{\omega}{\Omega} - \frac{\omega^2}{\Omega^2})}$$

which has absolute value

$$|m(\omega)| = \frac{1}{\sqrt{1 - \omega^6/\Omega^6}}.$$

The defining properties of a Butterworth filter are the following:

1. The filter is causal.
2. The system function is the inverse of a (complex) polynomial.
3. $|m(\omega)|^2 = \frac{1}{1+\omega^n/\Omega^n}$ for some $n \in \mathbb{N}, \Omega > 0$.

As $n \to \infty$, the system function approaches that of an ideal filter!

From analog to digital filters

We now examine filtering for bandlimited signals. Given a filter with a system function m and a bandlimited function f, can we express the sampled values of Lf in terms of those of f?

3.25 Remark. For filtering an Ω-bandlimited function, only the restriction of the system function m to $[-\Omega, \Omega]$ matters, because $\hat{f}$ vanishes outside of this interval.

We can thus expand m in a Fourier series,

$$m(\omega) = \sum_{k \in \mathbb{Z}} \alpha_k e^{-i\pi k\omega/\Omega},$$

where we have changed the sign in the exponent because it is a series for a function on the frequency domain.

3.26 Theorem. *Given an Ω-bandlimited function f and a filter L with system function m whose restriction to $[-\Omega, \Omega]$ has Fourier coefficients $\{\alpha_k\}_{k \in \mathbb{Z}}$. Then*

$$Lf(\frac{k\pi}{\Omega}) = \sum_{l=-\infty}^{\infty} f(\frac{(k-l)\pi}{\Omega})\alpha_l .$$

The upshot is that the convolution is replaced by a series formula for the sampled values of f.

3.27 Definition. *For two sequences $x, y \in l^2(\mathbb{Z})$, we define the discrete convolution as*

$$(x * y)_k = \sum_{l=-\infty}^{\infty} x_l y_{k-l} .$$

3.28 Remark. If a (bounded) system function has a large number of continuous, square integrable derivatives then the impulse response decays fast.

However, when implementing this filter digitally, that is, for Ω-bandlimited functions, then the decay of the coefficients $\{\alpha_k\}$ depends on the smoothness of the **periodization** of m restricted to $[-\Omega, \Omega]$.

For this reason, often filters are modified by smoothing the periodization around $\pm\Omega$.

Chapter 4

Haar Wavelets

Spaces of piecewise constant functions

We begin with the observation that functions that are constant on all intervales $[n, n+1)$, $n \in \mathbb{Z}$, can be written as the pointwise convergent series

$$f(x) = \sum_{k=-\infty}^{\infty} a_k \phi(x-k)$$

where

$$\phi(x) = \begin{cases} 1, & 0 \leq x < 1 \\ 0, & \text{else} \end{cases} .$$

4.1 Definition. *We define the space of square-integrable integer-wide step functions as*

$$V_0 = \{f(x) = \sum_{k=-\infty}^{\infty} a_k \phi(x-k), a \in \ell^2(\mathbb{Z})\} .$$

4.2 Exercise. We remark that the translates $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ form an orthonormal basis for V_0.

4.3 Question. Knowing the values of a function f at one point in each interval $[k, k+1)$ determines the function completely. How can we have a function space with more details?

4.4 Answer. Take $\{2^{j/2}\phi(2^j x - k)\}_{k\in\mathbb{Z}}$ as an orthonormal basis instead of $\{\phi(x-k)\}_{k\in\mathbb{Z}}$.

4.5 Definition. *The space of square-integrable step functions of width* 2^{-j}*, denoted by* V_j*, is the subspace of* $L^2(\mathbb{R})$ *with the orthonormal basis*

$$\{2^{j/2}\phi(2^j x - k)\}_{k\in\mathbb{Z}}.$$

4.6 Remark. Functions in this space have possible discontinuities at $x = 2^{-j}k$, $k \in \mathbb{Z}$. This implies that sampling the function values at 2^j evenly-spaced points in the interval $[k, k+1)$ determines the function on this interval.

We also note that for $j > 0$, we have the inclusions $V_{-j} \subset V_{-j+1} \subset \cdots \subset V_0 \subset V_1 \subset \cdots \subset V_{j-1} \subset V_j \subset V_{j+1}$.

4.7 Proposition. *For any square integrable function* f*,* $f \in V_0$ *if and only if* $f(2^j x) \in V_j$*, or equivalently* $f \in V_j$ *if and only if* $f(2^{-j}x) \in V_0$*.*

4.8 Question. Is there an orthonormal basis for layers of detail? We would like to have a basis of translates for $V_1 \cap V_0^{\perp}$.

Try

$$\psi(x) = \phi(2x) - \phi(2x-1)$$

then

$$\int_{-\infty}^{\infty} \phi(x)\psi(x)dx = \int_0^{1/2} 1dx - \int_{1/2}^{1} 1dx = 0$$

and because ψ is supported in $[0,1]$, it is orthogonal to all $\phi(x-k)$!

Indeed, the translates of ψ form a basis for the detail spaces that bridge between V_0 and V_1. More generally, we can define a subspace of V_{j+1} which is orthogonal to V_j.

4.9 Theorem. *Let* W_j *be the span of all functions in* $L^2(\mathbb{R})$ *such that*

$$f(x) = \sum_{k\in\mathbb{Z}} a_k \psi(2^j x - k).$$

Then $W_j = V_j^{\perp} \cap V_{j+1}$*.*

Haar decomposition

Now we can perform a recursive splitting. Each $f_j \in V_j$ is expressed uniquely as the sum

$$f_j = w_{j-1} + f_{j-1}$$

where $w_{j-1} \in W_{j-1}$ and $f_{j-1} \in V_{j-1}$. This orthogonal splitting is abbreviated by

$$V_j = W_{j-1} \oplus V_{j-1} .$$

Iterating the splitting gives

$$V_j = W_{j-1} \oplus W_{j-2} \oplus V_{j-2}$$

and so on.

If instead we let $j \to \infty$ and keep the last term in this direct sum decomposition fixed, say V_0, then we obtain a unique representation of each vector as a series of vectors from W_j, $j \geq 0$, and V_0.

4.10 Theorem. *For each $f \in L^2(\mathbb{R})$, denote by w_j the orthogonal projection of f onto W_j. Then*

$$f = f_0 + \sum_{j=0}^{\infty} w_j$$

with vectors that are orthogonal and a series that converges in norm. In short,

$$L^2(\mathbb{R}) = V_0 \oplus W_0 \oplus W_1 \oplus W_2 \oplus \cdots$$

4.11 Question. Suppose we have $f_j(x) = \sum_{k \in \mathbb{Z}} a_k \phi(2^j x - k)$, given by the values $\{a_k\}$. How do we compute the coefficients with respect to the orthonormal basis of V_j given by

$$\{\phi(x-k)\}_{k \in \mathbb{Z}} \text{ and } \{2^{l/2}\psi(2^l x - k)\}_{k \in \mathbb{Z}, 0 \leq l \leq j-1} ?$$

4.12 Lemma. *For the Haar scaling function ϕ and the wavelet ψ,*

$$\phi(2^j x) = \frac{1}{2}(\psi(2^{j-1}x) + \phi(2^{j-1}x))$$

and

$$\phi(2^j x - 1) = \frac{1}{2}(\phi(2^{j-1}x) - \psi(2^{j-1}x)) .$$

So we can use this to convert $\sum_k a_k \phi(2^j x - k) \in V_j$ into $\sum_k (c_k \phi(2^{j-1}x - k) + d_k \psi(2^{j-1}x - k))$.

4.13 Exercise. Show that

$$f(x) = 2\phi(4x) + 2\phi(4x-1) + \phi(4x-2) - \phi(4x-3) = 2\phi(2x) + \psi(2x-1) .$$

4.14 Theorem. *Given a square integrable function*

$$f_j(x) = \sum_k a_k^{(j)} \phi(2^j x - k)$$

then

$$f_j(x) = \sum_k b_k^{(j-1)} \psi(2^{j-1}x - k) + \sum_k a_k^{(j-1)} \phi(2^{j-1}x - k)$$

with

$$b_k^{(j-1)} = \frac{a_{2k}^{(j)} - a_{2k+1}^{(j)}}{2}$$

and

$$a_k^{(j-1)} = \frac{a_{2k}^{(j)} + a_{2k+1}^{(j)}}{2}.$$

We note that both of these expressions are obtained from a digital filter applied to $\{a_l^{(j)}\}_{l\in\mathbb{Z}}$.

We can repeat this procedure iteratively to obtain a coefficient tree containing $\{b_k^{(j)}\}_{k\in\mathbb{Z}}$, $\{b_k^{(j-1)}\}_{k\in\mathbb{Z}}$, $\{b_k^{(j-2)}\}_{k\in\mathbb{Z}}$, ... and finally $\{a_k^{(0)}\}_{k\in\mathbb{Z}}$.

Filters and diagrams

4.15 Definition. *For any sequence $\{x_k\}_{k\in\mathbb{Z}}$ and $\{h_k\}_{k\in\mathbb{Z}}$, both in $\ell^2(\mathbb{Z})$, we define the* ***digital/discrete filter*** *of x by*

$$(Hx)_k = (h * x)_k = \sum_{n\in\mathbb{Z}} x_{k-n} h_n .$$

4.16 Definition. *The* ***downsampling operator*** *D acts on a square-summable sequence $\{x_k\}_{k\in\mathbb{Z}}$ by*

$$(Dx)_k = x_{2k} .$$

With these two operations, we can express the analysis and reconstruction algorithm.

4.17 Remark. Let

$$h = (\dots, 0, 0, \dots, 0, -\frac{1}{2}, \underbrace{\frac{1}{2}}_{k=0}, 0, 0, \dots)$$

and let

$$l = (\dots, 0, 0, \dots, 0, \frac{1}{2}, \underbrace{\frac{1}{2}}_{k=0}, 0, 0, \dots).$$

Then

$$(Hx)_k = (h * x)_k = \frac{1}{2}x_k - \frac{1}{2}x_{k+1}$$

and

$$(Lx)_k = (l * x)_k = \frac{1}{2}x_k + \frac{1}{2}x_{k+1}\,.$$

Therefore,

$$b_k^{(j-1)} = \frac{1}{2}(a_{2k}^{(j)} - a_{2k+1}^{(j)}) = (DHa^{(j)})_k$$

and

$$a_k^{(j-1)} = (DLa^{(j)})_k\,.$$