Three Introductory Lectures on Fourier Analysis and Wavelets

Willard Miller

August 22, 2002

Contents

Chapter 1

Lecture I

1.1 Introduction

Let $f(t)$ be a real-valued function defined on the real line R and square integrable:

$$\int_{-\infty}^{\infty} f^2(t)dt < \infty.$$

Think of $f(t)$ as the value of a signal at time t. We want to analyse this signal in ways other than the time-value form $t \to f(t)$ given to us. In particular we will analyse the signal in terms of frequency components and various combinations of time and frequency components. Once we have analysed the signal we may want to alter some of the component parts to eliminate some undesirable features or to compress the signal for more efficient transmission and storage. Finally, we will reconstitute the signal from its component parts.

The three steps are:

- **Analysis.** Decompose the signal into basic components. We will think of the signal space as a vector space and break it up into a sum of subspaces, each of which captures a special feature of a signal.

- **Processing** Modify some of the basic components of the signal that were obtained through the analysis. Examples:

  1. audio compression
  2. video compression
  3. denoising

4. edge detection

- **Synthesis** Reconstitute the signal from its (altered) component parts. An important requirement we will make is **perfect reconstruction**. If we don't alter the component parts, we want the synthesised signal to agree exactly with the original signal.

Remarks:

- Some signals are discrete, e.g., only given at times $t_j = j, \quad j = 0, \pm 1, \pm 2, \cdots$. We will represent these as step functions.

- Audio signals (telephone conversations) are of arbitrary length but video signals are of fixed finite length, say 2π. Thus a video signal can be represented by a function $f(t)$ defined for $-\pi \leq t < \pi$. Mathematically, we can extend f to the real line by requiring that it be periodic

$$f(t) = f(t + 2\pi)$$

  or that it vanish outside the interval $-\pi \leq t < \pi$.

We will look at several methods for signal analysis:

- Fourier series

- The Fourier integral (very briefly)

- Windowed Fourier transforms (very briefly)

- Continuous wavelet transforms (very briefly)

- Discrete wavelet transforms (Haar and Daubechies wavelets)

All of these methods are based on the decomposition of the Hilbert space of square integrable functions into orthogonal subspaces. We will first review a few ideas from the theory of vector spaces.

1.2 Vector Spaces with Inner Product.

1.2.1 Definitions

Review of the following concepts:

1. vector space
2. subspace
3. linear independence
4. basis and dimension

Definition 1 *A vector space V over the field of real numbers R is a collection of elements (vectors) with the following properties:*

- *For every pair $u, v \in V$ there is defined a unique vector $w = u + v \in V$ (the sum of u and v)*
- *For every $\alpha \in R$, $u \in V$ there is defined a unique vector $z = \alpha u \in V$ (product of α and u)*
- *Commutative, Associative and Distributive laws*

  1. $u + v = v + u$
  2. $(u + v) + w = u + (v + w)$
  3. *There exists a vector* $\Theta \in V$ *such that* $u + \Theta = u$ *for all* $u \in V$
  4. *For every* $u \in V$ *there is a* $-u \in V$ *such that* $u + (-u) = \Theta$
  5. $1u = u$ *for all* $u \in V$
  6. $\alpha(\beta u) = (\alpha\beta)u$ *for all* $\alpha, \beta \in F$
  7. $(\alpha + \beta)u = \alpha u + \beta u$
  8. $\alpha(u + v) = \alpha u + \alpha v$

Definition 2 *A non-empty set W in V is a subspace of V if $\alpha u + \beta v \in W$ for all $\alpha, \beta \in R$ and $u, v \in W$.*

Note that W is itself a vector space over R.

Lemma 1 *Let $u_1, u_2, \cdots, u_m$ be a set of vectors in the vector space V. Denote by $[u_1, u_2, \cdots, u_m]$ the set of all vectors of the form $\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_m u_m$ for $\alpha_i \in F$. The set $[u_1, u_2, \cdots, u_m]$ is a subspace of V.*

PROOF: Let $u, v \in [u_1, u_2, \cdots, u_m]$. Thus,

$$u = \sum_{i=1}^{m} \alpha_i u_i, \qquad v = \sum_{i=1}^{m} \beta_i u_i$$

so

$$\alpha u + \beta v = \sum_{i=1}^{m} (\alpha\alpha_i + \beta\beta_i) u_i \in [u_1, u_2, \cdots, u_m].$$

Q.E.D.

Definition 3 *The elements $u_1, u_2, \cdots, u_p$ of V are linearly independent if the relation $\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_p u_p = \Theta$ for $\alpha_i \in F$ holds only for $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$. Otherwise $u_1, \cdots, u_p$ are linearly dependent*

Definition 4 *V is n-dimensional if there exist n linearly independent vectors in V and any $n+1$ vector in V are linearly dependent.*

Definition 5 *V is finite-dimensional if V is n-dimensional for some integer n. Otherwise V is infinite dimensional.*

Remark: If there exist vectors $u_1, \cdots, u_n$, linearly independent in V and such that every vector $u \in V$ can be written in the form

$$u = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n, \qquad \alpha_i \in F,$$

($\{u_1, \cdots, u_n\}$ *spans* V), then V is n-dimensional. Such a set $\{u_1, \cdots, u_n\}$ is called a *basis* for V.

Theorem 1 *Let V be an n-dimensional vector space and $u_1, \cdots, u_n$ a linearly independent set in V. Then $u_1, \cdots, u_n$ is a basis for V and every $u \in V$ can be written uniquely in the form*

$$u = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_n u_n.$$

PROOF: let $u \in V$. then the set $u_1, \cdots, u_n, u$ is linearly dependent. Thus there exist $\alpha_1, \cdots, \alpha_{n+1} \in F$, not all zero, such that

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n + \alpha_{n+1} u = \Theta.$$

If $\alpha_{n+1} = 0$ then $\alpha_1 = \cdots = \alpha_n = 0$. Impossible! Therefore $\alpha_{n+1} \neq 0$ and

$$u = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_n u_n, \qquad \beta_i = -\frac{\alpha_i}{\alpha_{n+1}}.$$

Now suppose

$$u = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_n u_n = \gamma_1 u_1 + \gamma_2 u_2 + \cdots + \gamma_n u_n.$$

Then

$$(\beta_1 - \gamma_1) u_1 + \cdots + (\beta_n - \gamma_n) u_n = \Theta.$$

But the u_i form a linearly independent set, so $\beta_1 - \gamma_1 = 0, \cdots, \beta_n - \gamma_n = 0$. Q.E.D.

Examples 1 • *V_n, the space of all real n-tuples $(\alpha_1, \cdots, \alpha_n)$, $\alpha_i \in R$. Here, $\Theta = (0, \cdots, 0)$. A standard basis is:*

$$u_1 = (1, 0 \cdots, 0), \quad u_2 = (0, 1, 0, \cdots, 0), \cdots, u_n = (0, 0, \cdots, 1).$$

- *V_∞, the space of all real infinity-tuples*

$$(\alpha_1, \alpha_2, \cdots, \alpha_n, \cdots).$$

  This is an infinite-dimensional space.

- *$C^{(n)}[a, b]$: Set of all real-valued functions with continuous derivatives of orders $0, 1, 2, \cdots n$ on the closed interval $[a, b]$ of the real line. Let $t \in [a, b]$, i.e., $a \leq t \leq b$ with $a < b$. Vector addition and scalar multiplication of functions $u, v \in C^{(n)}[a, b]$ are defined by*

$$[u + v](t) = u(t) + v(t) \qquad [\alpha u](t) = \alpha u(t).$$

  The zero vector is the function $\Theta(t) \equiv 0$. The space is infinite-dimensional.

- *$S(J)$: Space of all real-valued step functions on the (bounded or unbounded) interval J on the real line. s is a step function on J if there are a finite number of non-intersecting bounded intervals $J_1, \cdots, J_m$ and complex numbers $c_1, \cdots, c_m$ such that $s(t) = c_k$ for $t \in J_k$, $k = 1, \cdots, m$ and $s(t) = 0$ for $t \in J - \cup_{k=1}^m J_k$. Vector addition and scalar multiplication of step functions $s_1, s_2 \in S(J)$ are defined by*

$$[s_1 + s_2](t) = s_1(t) + s_2(t) \qquad [\alpha s_1](t) = \alpha s_1(t).$$

  (One needs to check that $s_1 + s_2$ and αs_1 are step functions.) The zero vector is the function $\Theta(t) \equiv 0$. The space is infinite-dimensional.

1.2.2 Inner product spaces

Review of the following concepts:

1. inner product
2. Schwarz inequality
3. norm

Definition 6 *A vector space $\mathcal{H}$ over R is an inner product space (pre-Hilbert space) if to every ordered pair $u, v \in \mathcal{H}$ there corresponds a scalar $(u, v) \in R$ such that*

1. $(u, v) = (v, u)$
2. $(u + v, w) = (u, w) + (v, w)$
3. $(\alpha u, v) = \alpha(u, v)$*, for all* $\alpha \in R$
4. $(u, u) \geq 0$*, and* $(u, u) = 0$ *if and only if* $u = 0$

Note: $(u, \alpha v) = \alpha(u, v)$

Definition 7 *let $\mathcal{H}$ be an inner product space with inner product (u, v). The norm $||u||$ of $u \in \mathcal{H}$ is the non-negative number $||u|| = \sqrt{(u, u)}$.*

Theorem 2 *Schwarz inequality. Let $\mathcal{H}$ be an inner product space and $u, v \in \mathcal{H}$. Then*

$$|(u, v)| \leq ||u||\ ||v||.$$

Equality holds if and only if u, v are linearly dependent.

PROOF: We can suppose $u, v \neq \Theta$. Set $w = u + \alpha v$, for $\alpha \in R$. The $(w, w) \geq 0$ and $= 0$ if and only if $u + \alpha v = 0$. hence

$$(w, w) = (u + \alpha v, u + \alpha v) = ||u||^2 + |\alpha|^2\ ||v||^2 + 2\alpha(v, u) \geq 0.$$

Set $\alpha = -(u, v)/||v||^2$. Then

$$||u||^2 + \frac{|(u, v)|^2}{||v||^2} - 2\frac{(u, v)^2}{||v||^2} \geq 0.$$

Thus $(u, v)^2 \leq ||u||^2\ ||v||^2$. Q.E.D.

Theorem 3 *Properties of the norm. Let $\mathcal{H}$ be an inner product space with inner product (u, v). Then*

- $||u|| \geq 0$ *and* $||u|| = 0$ *if and only if* $u = 0$.

- $||\alpha u|| = |\alpha| \; ||u||$.
- *Triangle inequality.* $||u+v|| \leq ||u|| + ||v||$.

  PROOF:

$$||u+v||^2 = (u+v, u+v) = ||u||^2 + (u,v) + (v,u) + ||v||^2$$

$$\leq ||u||^2 + 2||u|| \; ||v|| + ||v||^2 = (||u|| + ||v||)^2.$$

Examples:

- R_n This is the space of real n-tuples V_n with inner product

  $$(u, v) = \sum_{i=1}^{n} \alpha_i \beta_i$$

  for vectors

  $$u = (\alpha_1, \cdots, \alpha_n), \qquad v = (\beta_1, \cdots, \beta_n), \qquad \alpha_i, \beta_i \in R.$$

  Note that (u, v) is just the dot product. In particular for R_3 (Euclidean 3-space) $(u, v) = ||u||\ ||v|| \cos\phi$ where $||u|| = \sqrt{\alpha_i^2 + \alpha_2^2 + \alpha_3^2}$ (the length of u), and $\cos\phi$ is the cosine of the angle between vectors u and v. The triangle inequality $||u + v|| \leq ||u|| + ||v||$ says in this case that the length of one side of a triangle is less than or equal to the sum of the lengths of the other two sides.

- ℓ^2, the space of all real infinity-tuples

  $$u = (\cdots, \alpha_{-1}, \alpha_0, \alpha_1, \cdots, \alpha_n, \cdots).$$

  such that $\sum_{i=-\infty}^{\infty} \alpha_i^2 < \infty$. Here, $(u, v) = \sum_{i=-\infty}^{\infty} \alpha_i \beta_i$. (Need to verify that this is a vector space.)

- $L_0^2[a, b]$: Set of all real-valued functions $u(t)$ on the closed interval $[a, b]$ of the real line, such that $\int_a^b u(t)^2\, dt < \infty$, (Riemann integral). We define an inner product by

  $$(u, v) = \int_a^b u(t)v(t)\, dt, \qquad u, v \in L^2[a, b].$$

  Note: There are problems here. Strictly speaking, this isn't an inner product. Indeed the nonzero function $u(0) = 1, u(t) = 0$ for $t > 0$ belongs to $L_0^2[0, 1]$, but $||u|| = 0$. However the other properties of the inner product hold.

- $S^2(J)$: Space of all real-valued step functions on the (bounded or unbounded) interval J on the real line. s is a *step function* on J if there are a finite number of non-intersecting bounded intervals $J_1, \cdots, J_m$ and numbers $c_1, \cdots, c_m$ such that $s(t) = c_k$ for $t \in J_k$, $k = 1, \cdots, m$ and $s(t) = 0$ for

$t \in J - \cup_{k=1}^{m}$. Vector addition and scalar multiplication of step functions $s_1, s_2 \in S(J)$ are defined by

$$[s_1 + s_2](t) = s_1(t) + s_2(t) \quad [\alpha s_1](t) = \alpha s_1(t).$$

(One needs to check that $s_1 + s_2$ and αs_1 are step functions.) The zero vector is the function $\Theta(t) \equiv 0$. Note also that the product of step functions, defined by $s_1 s_2(t) \equiv s_1(t)s_2(t)$ is a step function, as is $|s_1|$. We define the integral of a step function as $\int_J s(t)dt \equiv \sum_{k=1}^{m} c_k \ell(J_k)$ where $\ell(J_k) =$ length of $J_k = b - a$ if $J_k = [a, b]$ or $[a, b)$, or $(a, b]$ or (a, b). Now we define the inner product by $(s_1, s_2) = \int_J s_1(t)s_2(t)dt$. Finally, we adopt the rule that we identify $s_1, s_2 \in S(J)$, $s_1 \sim s_2$ if $s_1(t) = s_2(t)$ except at a finite number of points. (This is needed to satisfy property 4. of the inner product.) Now we let $S^2(J)$ be the space of equivalence classes of step functions in $S(J)$. Then $S^2(J)$ is an inner product space.

- REMARK. An inner product space $\mathcal{H}$ is called a *Hilbert space* if it is closed in the norm, i.e., if every sequence $\{u_n\}$, Cauchy in the norm, converges to an element of $\mathcal{H}$: $\lim_{n\to\infty} ||u_n - u|| = 0$. In a manner analogoue to the completion of the rational numbers to obtain the real numbers, every inner product space can be completed to a Hilbert space. The completion of $L^2(R)$ (Riemann integral) is the Hilbert space of Lebesgue square integrable functions. In the following we shall assume that we have completed $L^2(R)$, so that every cauchy sequence in the norm converges.

1.2.3 Orthogonal projections

Definition 8 *Two vectors u, v in an inner product space $\mathcal{H}$ are called orthogonal, $u \perp v$, if $(u, v) = 0$. Similarly, two sets $\mathcal{M}, \mathcal{N} \subset \mathcal{H}$ are orthogonal, $\mathcal{M} \perp \mathcal{N}$, if $(u, v) = 0$ for all $u \in \mathcal{M}$, $v \in \mathcal{N}$.*

Definition 9 *Let $\mathcal{S}$ be a nonempty subset of the inner product space $\mathcal{H}$. We define $\mathcal{S}^{\perp} = \{u \in \mathcal{H} : u \perp \mathcal{S}\}$*

Definition 10 *The set of vectors $\{e_1, e_2, \cdots, e_m\}$ (where m could be infinite) for $e_j \in \mathcal{H}$ is called orthonormal (ON) if*

$$(e_i, e_j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Given an ON set $\{e_1, e_2, \cdots, e_m\}$ let

$$\mathcal{W} = \left\{ w \in \mathcal{H} : w = \sum_{i=1}^{m} \alpha_i e_i, \quad \alpha_i \in R \right\}.$$

Then $\mathcal{W}$ is a subspace of $\mathcal{H}$. Note:

1. If m is infinite we must have

$$||w||^2 = (w, w) = \sum_{i=1}^{\infty} \alpha_i^2 < \infty.$$

2. If $w = \sum_{i=1}^{m} \alpha_i e_i \in \mathcal{W}$ then

$$(w, e_j) = (\sum_{i=1}^{m} \alpha_i e_i, e_j) = \sum_{i=1}^{m} \alpha_i (e_i, e_j) = \alpha_j.$$

   (True even if m is infinite, but the property $\sum_{i=1}^{\infty} \alpha_i^2 < \infty$ is needed to justify the term-by-term evaluation of the infinite sum.

3. If $w \in \mathcal{W}$ then it is uniquely represetable in the form

$$w = \sum_{i=1}^{m} (w, e_i) e_i.$$

   The set $\{e_j\}$ is called an *ON basis* for $\mathcal{W}$.

Definition 11 *Let $u \in \mathcal{H}$. We say that the vector $u' = \sum_{i=1}^m (u, e_i)e_i \in \mathcal{W}$ is the projection of u on $\mathcal{W}$.*

Theorem 4 *If $u \in \mathcal{H}$ there exist unique vectors $u' \in \mathcal{W}$, $u'' \in \mathcal{W}^\perp$ such that $u = u' + u''$. We write $\mathcal{H} = \mathcal{W} \oplus \mathcal{W}^\perp$.*

PROOF:

1. Existence: Let $\{e_1, e_2, \cdots, e_m\}$ be an ON basis for $\mathcal{W}$, set $u' = \sum_{i=1}^m (u, e_i)e_i \in \mathcal{W}$ and $u'' = u - u'$. Now $(u'', e_i) = (u, e_i) - (u, e_i) = 0$, $1 \leq i \leq m$, so $(u'', v) = 0$ for all $v \in \mathcal{W}$. Thus $u'' \in \mathcal{W}^\perp$.

2. Uniqueness: Suppose $u = u' + u'' = v" + v''$ where $u', v' \in \mathcal{W}$, $u'', v'' \in \mathcal{W}^\perp$. Then $u' - v' = v'' - u'' \in \mathcal{W} \cap \mathcal{W}^\perp \Longrightarrow (u' - v', u' - v') = 0 = ||u' - v'||^2 \Longrightarrow u' = v', u'' = v''$. Q.E.D.

Corollary 1 *Bessel's Inequality. Let $\{e_1, \cdots, e_m\}$ be an ON set in $\mathcal{H}$. If $u \in \mathcal{H}$ then $||u||^2 \geq \sum_{i=1}^m (u, e_i)^2$.*

PROOF: Set $W = [e_1, \cdots, e_m]$. Then $u = u' + u''$ where $u' \in \mathcal{W}$, $u'' \in \mathcal{W}^\perp$, and $u' = \sum_{i=1}^m (u, e_i)e_i$. Therefore $||u||^2 = (u' + u'', u' + u'') = ||u'||^2 + ||u''||^2 \geq ||u'||^2 = \sum_{i=1}^m |(u, e_i)|^2$. Q.E.D.

Note that this inequality holds even if m is infinite. If m is infinite then we must have that the terms $(u, e_i)^2$ go to zero as $i \to \infty$ in order that the infinite sum of squares converge.

Corollary 2 *Riemann-Lebesgue Lemma. If $u \in \mathcal{H}$ and $\{e_n : n = 1, 2, \cdots\}$ is an ON set in $\mathcal{H}$ then*

$$\lim_{n \to \infty} (u, e_n) = 0.$$

The projection of $u \in \mathcal{H}$ onto the subspace $\mathcal{W}$ has invariant meaning, i.e., it is basis independent. Also, it solves an important minimization problem: u' is the vector in $\mathcal{W}$ that is closest to u.

Theorem 5 $\min_{v \in \mathcal{W}} ||u - v|| = ||u - u'||$ *and the minimum is achieved if and only if $v = u'$.*

PROOF: let $v \in \mathcal{W}$ and let $\{e_1, e_2, \cdots, e_m\}$ be an ON basis for $\mathcal{W}$. Then $v = \sum_{i=1}^m \alpha_i e_i$ for $\alpha_i = (v, e_i)$ and $||u - v|| = ||u - \sum_{i=1}^m \alpha_i e_i||^2 = (u - \sum_{i=1}^m \alpha_i e_i, u - \sum_{i=1}^m \alpha_i e_i) = ||u||^2 - \sum_{i=1}^m \alpha_i (u, e_i) - \sum_{i=1}^m \alpha_i (e_i, u) + \sum_{i=1}^m \alpha_i^2 = ||u - \sum_{i=1}^m (u, e_i)e_i||^2 + \sum_{i=1}^m |(u, e_i) - \alpha_i|^2 \geq ||u - u'||^2$. Equality is obtained if and only if $\alpha_i = (u, e_i)$, for $1 \leq i \leq m$. Q.E.D.

1.3 Fourier Series

1.3.1 Real Fourier series

Let $L^2[0, 2\pi]$ be the inner product space of Riemann square-integrable functions on the interval $[0, 2\pi]$. Here the inner product is

$$(u, v) = \int_0^{2\pi} u(t)v(t)\, dt, \qquad u, v \in L^2[0, 2\pi].$$

(This satisfies the condition $||v|| = 0 \Leftrightarrow v \equiv 0$ provided we identify all functions with the same interals.) It is convenient to assume that $L^2[0, 2\pi]$ consists of square-integrable functions on the unit circle, rather than on an interval of the real line. Thus we will replace every function $f(t)$ on the interval $[0, 2\pi]$ by a function $f^*(t)$ such that $f^*(0) = f^*(2\pi)$ and $f^*(t) = f(t)$ for $0 < t < 2\pi$. Then we will extend f^* to all $-\infty < t < \infty$ be requiring periodicity: $f^*(t + 2\pi) = f^*(t)$. This will not affect the values of any integrals over the interval $[0, 2\pi]$. Thus, from now on our functions will be assumed $2\pi - periodic$.

Consider the set

$$e_0(t) = \frac{1}{\sqrt{2\pi}}, \quad e_{2j}(t) = \frac{1}{\sqrt{\pi}} \cos jt, \quad e_{2j-1}(t) = \frac{1}{\sqrt{\pi}} \sin jt$$

for $j = 1, 2, \cdots$. It is easy to check that $\{e_n(t) : n = 0, 1, 2, \cdots\}$ is an ON set in $L^2[0, 2\pi]$. Let $\mathcal{W}$ be the subspace of $L^2[0, 2\pi]$ consisting of all vectors $g = \sum_{n=0}^{\infty} \alpha_n e_n(t)$ such that $\sum_{n=0}^{\infty} \alpha_n^2 < \infty$.

Definition 12 *Given $f \in L^2[0, 2\pi]$ the* **Fourier series** *of f is the projection $f_{\rm proj}(t)$ of f on $\mathcal{W}$:*

$$f_{\rm proj}(t) = \sum_{n=0}^{\infty} (f, e_n) e_n(t).$$

In terms of sines and cosines this is usually written

$$\begin{aligned} f(t) &\sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt), \qquad a_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos nt\, dt \\ b_n &= \frac{1}{\pi} \int_0^{2\pi} f(t) \sin nt\, dt \end{aligned} \tag{1.1}$$

where,

$$a_0 = \sqrt{\frac{2}{\pi}}(f, e_0), \quad a_n = \sqrt{\frac{1}{\pi}}(f, e_{2n}), \quad b_n = \sqrt{\frac{1}{\pi}}(f, e_{2n-1})$$

with Bessel inequality

$$\frac{1}{\pi}\int_0^{2\pi} f(t)^2 dt \geq \frac{a_0^2}{2} + \sum_{n=1}^{\infty}(a_n^2 + b_n^2).$$

We will prove (partially) the following basic results:

Theorem 6 *Parseval's equality. (Convergence in the norm) Let $f \in L^2[0, 2\pi]$. Then*

$$\frac{1}{\pi}\int_0^{2\pi} |f(t)|^2 dt = \frac{a_0^2}{2} + \sum_{n=1}^{\infty}(a_n^2 + b_n^2).$$

This is equivalent to the statement that $\mathcal{W} = L^2[0, 2\pi]$, i.e., that $\{e_n(t)\}$ is an ON basis for $L^2[0, 2\pi]$.

Let $f \in L^2[0, 2\pi]$ and remember that we are assuming that all such functions satisfy $f(t + 2\pi) = f(t)$. We say that f is *piecewise continuous* on $[0, 2\pi]$ if it is continuous except for a finite number of discontinuities. Furthermore, at each t the limits $f(t + 0) = \lim_{h\to 0, h>0} f(t + h)$ and $f(t - 0) = \lim_{h\to 0, h>0} f(t - h)$ exist. NOTE: At a point t of continuity of f we have $f(t+0) = f(t-0)$, whereas at a point of discontinuity $f(t + 0) \neq f(t - 0)$ and $f(t + 0) - f(t - 0)$ is the magnitude of the jump discontinuity.

Theorem 7 *Suppose*

- $f(t)$ *is periodic with period* 2π.
- $f(t)$ *is piecewise continuous on* $[0, 2\pi]$.
- $f'(t)$ *is piecewise continuous on* $[0, 2\pi]$.

Then the Fourier series of $f(t)$ converges to $\frac{f(t+0)+f(t-0)}{2}$ at each point t.

PROOF: We modify f, if necessary, so that

$$f(t) = \frac{f(t + 0) + f(t - 0)}{2}$$

at each point t. This condition affects the definition of f only at a finite number of points of discontinuity. It doesn't change any integrals and the values of the Fourier coefficients.

Expanding f in a Fourier series (real form) we have

$$
\begin{aligned}
f(t) &\sim \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt) = S(t), \qquad a_n = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos nt\, dt \\
b_n &= \frac{1}{\pi}\int_0^{2\pi} f(t)\sin nt\, dt
\end{aligned} \tag{1.2}
$$

Let

$$
S_k(t) = \frac{a_0}{2} + \sum_{n=1}^{k}(a_n \cos nt + b_n \sin nt)
$$

be the k-th partial sum of the Fourier series. This is a finite sum, a *trigonometric polynomial*, so it is well defined for all $t \in R$. Now we have

$$
S(t) = \lim_{k\to\infty} S_k(t),
$$

if the limit exists. We will recast this finite sum as a single integral. Substituting the expressions for the Fourier coefficients a_n, b_n into the finite sum we find

$$
S_k(t) = \frac{1}{2\pi}\int_0^{2\pi} f(x)dx + \frac{1}{\pi}\sum_{n=1}^{k}\left(\int_0^{2\pi} f(x)\cos nx\, dx \cos nt + \int_0^{2\pi} f(x)\sin nx\, dx \sin nt\right),
$$

so

$$
\begin{aligned}
S_k(t) &= \frac{1}{\pi}\int_0^{2\pi}\left[\frac{1}{2} + \sum_{n=1}^{k}(\cos nx \cos nt + \sin nx \sin nt\right] f(x)dx \\
&= \frac{1}{\pi}\int_0^{2\pi}\left[\frac{1}{2} + \sum_{n=1}^{k}\cos[n(t-x)]\right] f(x)dx \\
&= \frac{1}{\pi}\int_0^{2\pi} D_k(t-x)f(x)dx.
\end{aligned} \tag{1.3}
$$

We can find a simpler form for the kernel $D_k(t) = \frac{1}{2} + \sum_{n=1}^{k}\cos nt = -\frac{1}{2} + \sum_{m=0}^{k}\cos mt$. The last cosine sum is the real part of the geometric series

$$
\sum_{m=0}^{k}(e^{it})^m = \frac{(e^{it})^{k+1} - 1}{e^{it} - 1}
$$

so

$$
-\frac{1}{2} + \sum_{m=0}^{k}\cos mt = -\frac{1}{2} + \mathrm{Re}\,\frac{(e^{it})^{k+1} - 1}{e^{it} - 1}
$$

$$= \mathrm{Re}\, \frac{(e^{it})^{k+1} - \frac{1}{2}e^{it} - \frac{1}{2}}{e^{it} - 1} = \mathrm{Re}\, \frac{e^{ikt} - e^{i(k+1)t} + \frac{1}{2}(e^{it} - e^{-it})}{4\sin^2 \frac{t}{2}}.$$

Thus,

$$D_k(t) = \frac{\cos kt - \cos(k+1)t}{4\sin^2 \frac{t}{2}} = \frac{\sin(k+\frac{1}{2})t}{2\sin \frac{t}{2}}. \tag{1.4}$$

Note that D has the properties:

- $D_k(t) = D_k(t + 2\pi)$
- $D_k(-t) = D_k(t)$
- $D_k(t)$ is defined and differentiable for all t and $D_k(0) = k + \frac{1}{2}$.

Lemma 2

$$\int_0^{2\pi} D_k(x)dx = \pi$$

PROOF:

$$\int_0^{2\pi} D_k(x)dx = \int_0^{2\pi} (\frac{1}{2} + \sum_{n=1}^{k} \cos nx)dx = \pi.$$

Q.E.D.

Using the Lemma we can write

$$S_k(t) - f(t) = \frac{1}{\pi} \int_0^{2\pi} D_k(t-x)[f(x) - f(t)]dx$$

$$= \frac{1}{\pi} \int_0^{\pi} D_k(x)[f(t+x) + f(t-x) - 2f(t)]dx$$

$$= \frac{1}{\pi} \int_0^{\pi} \frac{[f(t+x) + f(t-x) - 2f(t)]}{2\sin \frac{x}{2}} \sin(k + \frac{1}{2})x \; dx$$

$$= \frac{1}{\pi} \int_0^{\pi} [H_1(t,x) \sin kx + H_2(t,x) \cos kx]dx$$

From the assumptions, H_1, H_2 are square integrable in x. In particular, they are bounded for $x = 0$. Thus, by the Riemann-Lebesgue Lemma the last expression goes to 0 as $k \to \infty$:

$$\lim_{k\to\infty} [S_k(t) - f(t)] = 0$$

. Q.E.D.

1.3.2 Example

Let

$$f(t) = \begin{cases} 0, & t = 0 \\ \frac{\pi - t}{2}, & 0 < t < 2\pi \\ 0, & t = 2\pi. \end{cases}$$

and $f(t + 2\pi) = f(t)$. We have $a_0 = \frac{1}{\pi} \int_0^{2\pi} \frac{\pi - t}{2}\, dt = 0$. and for $n \geq 1$,

$$a_n = \frac{1}{\pi} \int_0^{2\pi} \frac{\pi - t}{2} \cos nt\, dt = \frac{\frac{\pi - t}{2} \sin nt}{n\pi} \Big|_0^{2\pi} + \frac{1}{2\pi n} \int_0^{2\pi} \sin nt\, dt = 0,$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} \frac{\pi - t}{2} \sin nt\, dt = -\frac{\frac{\pi - t}{2} \cos nt}{n\pi} \Big|_0^{2\pi} - \frac{1}{2\pi n} \int_0^{2\pi} \cos nt\, dt = \frac{1}{n}.$$

Therefore,

$$\frac{\pi - t}{2} = \sum_{n=1}^{\infty} \frac{\sin nt}{n}, \qquad 0 < t < 2\pi.$$

By setting $t = \pi/2$ in this expansion we get an alternating series for $\pi/4$:

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots.$$

Parseval's identity gives

$$\frac{\pi^2}{6} = \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

One way that Fourier series can be used for data compression of a signal

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

is that the signal can be approximated by the trigonometric polynomial

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

for some suitable integer N, i.e., $f(t)$ can be replaced by its projection on the subspace generated by the harmonics, $\sin nt$, $\cos nt$ for $n = 0, 1, \cdots, N$. Then just the data $a_0, a_1, \cdots, a_N, b_1, \cdots, b_N$ is transmitted, rather than the entire signal $f(t)$. Once the data is received, the projection $S_N(t)$ can then be synthesized.

1.4 The Fourier Transform

Let $f(t)$ belong to the inner product space $L^2(R)$, where now we permit f to take complex values. The (complex) inner product on this space is defined by

$$(f,g) = \int_{-\infty}^{\infty} f(t)\overline{g}(t)dt, \qquad f,g \in L^2(R)$$

where $\overline{g}(t)$ is the complex conjugate of $g(t)$. This inner product satisfies the usual Schwarz inequality in the form

$$|(f,g)| \le ||f|| \cdot ||g||$$

where $||f||^2 = (f,f)$. We define the Fourier integral of f by

$$\hat{f}(\lambda) = \int_{-\infty}^{\infty} f(t)e^{-i\lambda t}dt, \tag{1.5}$$

if the integral converges. Whether or not the infinite integral converges, we can define the finite integral

$$\hat{f}_N(\lambda) = \int_{-N}^{N} f(t)e^{-i\lambda t}dt. \tag{1.6}$$

and show that the sequence $\hat{f}_N, N = 1, 2, \cdots$ is Cauchy in the norm of $L^2(R)$. Thus it converges to a Lebesgue square-integrable function $\hat{f}(\lambda)$ in the completion of $L^2(R)$ as a Hilbert space: $\hat{f}_N \to \hat{f}$ as $N \to \infty$. Moreover, f can be recovered from its Fourier transform:

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\lambda)e^{i\lambda t}d\lambda \tag{1.7}$$

where the convergence is in the norm of $L^2(R)$ and, if f is sufficiently well behaved as a function, in the pointwise sense. Also we have the *Plancherel identity*

$$2\pi\int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |\hat{f}(\lambda)|^2 d\lambda. \tag{1.8}$$

1.4.1 Example

1. The box function (or rectangular wave)

$$\Pi(t) = \begin{cases} 1 & \text{if } -\pi < t < \pi \\ \frac{1}{2} & \text{if } t = \pm\pi \\ 0 & \text{otherwise.} \end{cases} \tag{1.9}$$

   Then, since $\Pi(t)$ is an even function, we have

$$\hat{\Pi}(\lambda) = \int_{-\infty}^{\infty} \Pi(t) e^{-i\lambda t} dt = \int_{-\infty}^{\infty} \Pi(t) \cos(\lambda t) dt$$
$$= \int_{-\pi}^{\pi} \cos(\lambda t) dt = \frac{2\sin(\pi\lambda)}{\lambda} = 2\pi \text{ sinc } \lambda.$$

   Thus sinc λ is the Fourier transform of the box function. The inverse Fourier transform is

$$\int_{-\infty}^{\infty} \text{sinc}(\lambda) e^{i\lambda t} d\lambda = \Pi(t),$$

   as follows from a limit argument in calculus, or from complex variable theory. Furthermore, we have

$$\int_{-\infty}^{\infty} |\Pi(t)|^2 dt = 2\pi$$

   and

$$\int_{-\infty}^{\infty} |\text{ sinc } (\lambda)|^2 d\lambda = 1$$

   from calculus, so the Plancherel equality is verified in this case. Note that the inverse Fourier transform converges to the midpoint of the discontinuity, just as for Fourier series.

2. We want to compute the Fourier transform of the rectangular box function with support on $[c, d]$:

$$R(t) = \begin{cases} 1 & \text{if } c < t < d \\ \frac{1}{2} & \text{if } t = c, d \\ 0 & \text{otherwise.} \end{cases}$$

   Recall that the box function

$$\Pi(t) = \begin{cases} 1 & \text{if } -\pi < t < \pi \\ \frac{1}{2} & \text{if } t = \pm\pi \\ 0 & \text{otherwise.} \end{cases}$$

has the Fourier transform $\hat{\Pi}(\lambda) = 2\pi \ \mathrm{sinc}\ \lambda.$ but we can obtain R from Π by first translating $t \to s = t - \frac{(c+d)}{2}$ and then rescaling $s \to \frac{2\pi}{d-c}s$:

$$R(t) = \Pi(\frac{2\pi}{d-c}t - \pi\frac{c+d}{d-c}).$$

$$\hat{R}(\lambda) = \frac{4\pi^2}{d-c}e^{i\pi\lambda(c+d)/(d-c)}\mathrm{sinc}(\frac{2\pi\lambda}{d-c}). \qquad (1.10)$$

1.4.2 L^2 convergence of the Fourier transform

Lemma 3

$$2\pi(R_{a,b}, R_{c,d})_{L^2} = (\hat{R}_{a,b}, \hat{R}_{c,d})_{\hat{L}^2}$$

for all real numbers $a \leq b$ and $c \leq d$.

Since any step functions u, v are finite linear combination of indicator functions R_{a_j,b_j} with complex coeficients, $u = \sum_j \alpha_j R_{a_j,b_j}$, $v = \sum_k \beta_k R_{c_k,d_k}$ we have

$$(\hat{u}, \hat{v})_{\hat{L}^2} = \sum_{j,k} \alpha_j \overline{\beta}_k (\hat{R}_{a_j,b_j}, \hat{R}_{c_k,d_k})_{\hat{L}^2}$$

$$= 2\pi \sum_{j,k} \alpha_j \overline{\beta}_k (R_{a_j,b_j}, R_{c_k,d_k})_{L^2} = (u, v)_{L^2}.$$

Thus $\mathcal{F}$ preserves inner product on step functions, and by taking Cauchy sequences of step functions, we have the

Theorem 8 *(Plancherel Formula) Let $f, g \in L^2[-\infty, \infty]$. Then*

$$2\pi(f, g)_{L^2} = (\hat{f}, \hat{g})_{\hat{L}^2}, \qquad ||f||^2_{L^2} = ||\hat{f}||^2_{\hat{L}^2}$$

The pointwise convergence properties of the inverse Fourier transform (and the proofs) are very similar to those for Fourier series:

Theorem 9 *Let f be a complex valued function such that*

- *$f(t)$ is absolutely Riemann integrable on $(-\infty, \infty)$ (hence $f \in L^1[-\infty, \infty]$).*
- *$f(t)$ is piecewise continuous on $(-\infty, \infty)$, with only a finite number of discontinuities in any bounded interval.*
- *$f'(t)$ is piecewise continuous on $(-\infty, \infty)$, with only a finite number of discontinuities in any bounded interval.*
- *$f(t) = \frac{f(t+0)+f(t-0)}{2}$ at each point t.*

Let

$$\hat{f}(\lambda) = \int_{-\infty}^{\infty} f(t) e^{-i\lambda t} dt$$

be the Fourier transform of f. Then

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\lambda) e^{i\lambda t} d\lambda$$

for every $t \in (-\infty, \infty)$.

Remarks:

- Fourier series decompose periodic signals $f(t)$ into frequency harmonics $\sin nt$ and $\cos nt$. The frequency information is given by the data a_n, b_n.
- The frequency coefficients a_n, b_n depend on the values $f(t)$ for *all* t in the interval $[0, 2\pi)$ whereas the convergence of the Fourier series at t_0 depends only on the local behavior of f in an arbitrarily small neighborhood of t_0.
- The Fourier transform decomposes signals $f(t)$ into pure frequency terms $e^{i\lambda t}$. The frequency information is given by the transform function $\hat{f}(\lambda)$.
- The transform function $\hat{f}(\lambda)$ depends on the values of $f(t)$ for *all* $-\infty < t < \infty$ whereas the convergence of the inverse Fourier transform at t_0 depends only on the local behavior of f in an arbitrarily small neighborhood of t_0.
- For compression and transmission of an audio signal, the transform as given would be almost useless. One would have to wait an infinite legth of time to compute the Fourier transform. What is needed is an audio compression filter that analyses and processes the audio signal on the fly, and then retransmits it, say with a one second delay.
- We will now look at several methods that still devide the signal into frequency bands, but that can sample the signal only locally in time to determine the transform coefficients.

Chapter 2

Lecture II

2.1 Windowed Fourier transforms

Let $g \in L_2(R)$ with $||g|| = 1$ and define the time-frequency translation of g by

$$g^{[x_1,x_2]}(t) = e^{2\pi i t x_2} g(t + x_1).$$

Now suppose g is centered about the point (t_0, ω_0) in phase (time-frequency) space, i.e., suppose

$$\int_{-\infty}^{\infty} t|g(t)|^2 dt = t_0, \quad \int_{-\infty}^{\infty} \omega|\hat{g}(\omega)|^2 d\omega = \omega_0$$

where $\hat{g}(\omega) = \int_{-\infty}^{\infty} g(t)e^{-2\pi i\omega t}dt$ is the Fourier transform of $g(t)$. Then

$$\int_{-\infty}^{\infty} t|g^{[x_1,x_2]}(t)|^2 dt = t_0 - x_1, \quad \int_{-\infty}^{\infty} \omega|\hat{g}^{[x_1,x_2]}(t)|^2 d\omega = \omega_0 + x_2$$

so $g^{[x_1,x_2]}$ is centered about $(t_0 - x_1, \omega_0 + x_2)$ in phase space. To analyze an arbitrary function $f(t)$ in $L_2(R)$ we compute the inner product

$$F(x_1, x_2) = \langle f, g^{[x_1,x_2]} \rangle = \int_{-\infty}^{\infty} f(t)\bar{g}^{[x_1,x_2]}(t)dt$$

with the idea that $F(x_1, x_2)$ is sampling the behavior of f in a neighborhood of the point $(t_0 - x_1, \omega_0 + x_2)$ in phase space. As x_1, x_2 range over all real numbers the samples $F(x_1, x_2)$ give us enough information to reconstruct $f(t)$.

As x_1, x_2 range over all real numbers the samples $F(x_1, x_2)$ give us enough information to reconstruct $f(t)$. It is easy to show this directly for functions f

such that $f(t)\overline{g}(t-s) \in L^2[-\infty,\infty]$ for all s. Indeed let's relate the windowed Fourier transform to the usual Fourier transform of f (rescaled for this lecture):

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i\omega t}dt, \qquad f(t) = \int_{-\infty}^{\infty} \tilde{f}(\omega)e^{2\pi i\omega t}d\omega. \tag{2.1}$$

Thus since

$$F(x_1,x_2) = \int_{-\infty}^{\infty} f(t)\overline{g(t+x_1)}e^{-2\pi i t x_2}dt,$$

we have

$$f(t)\overline{g(t+x_1)} = \int_{-\infty}^{\infty} F(x_1,x_2)e^{2\pi i t x_2}dx_2.$$

Multiplying both sides of this equation by $g(t+x_1)$ and integrating over x_1 we obtain

$$f(t) = \frac{1}{||g||^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x_1,x_2)g(t+x_1)e^{2\pi i t x_2}dx_1dx_2. \tag{2.2}$$

This shows us how to recover $f(t)$ from the windowed Fourier transform, if f and g decay sufficiently rapidly at ∞.

However, the set of basis states $g^{[x_1,x_2]}$ is overcomplete: the coefficients $\langle f, g^{[x_1,x_2]}\rangle$ are not independent of one another, i.e., in general there is no $f \in L_2(R)$ such that $\langle f, g^{[x_1,x_2]}\rangle = F(x_1,x_2)$ for an arbitrary $F \in L_2(R^2)$. The $g^{[x_1,x_2]}$ are examples of **coherent states**, continuous overcomplete Hilbert space bases which are of interest in quantum optics and quantum field theory, as well as group representation theory.

Example 1 *Given the function*

$$g(t) = \begin{cases} 1, & |t| \le \frac{1}{2} \\ 0, & |t| \ge \frac{1}{2}, \end{cases}$$

the set $\{g^{[m,n]}\}$ is an ON basis for $L^2(-\infty,\infty)$. Here, m,n run over the integers. Thus $g^{[x_1,x_2]}$ is overcomplete.

Hence it isn't necessary to compute the inner products $\langle f, g^{[x_1,x_2]}\rangle = F(x_1,x_2)$ for every point in phase space. In the windowed Fourier approach one typically samples F at the lattice points $(x_1,x_2) = (ma, nb)$ where a, b are fixed positive numbers and m, n range over the integers. Here, a, b and $g(t)$ must be chosen so that the map $f \longrightarrow \{F(ma,nb)\}$ is one-to-one; then f can be recovered from the lattice point values $F(ma,nb)$. The study of when this can happen is the study of *Weyl-Heisenberg frames*. It is particularly useful when g can be chosen such that $g^{[ma,nb]}$ is an ON basis for L^2.

2.2 Continuous wavelets

Let $\phi \in L_2(R)$ with $||\phi|| = 1$ and define the affine translation of ϕ by

$$\phi^{(a,b)}(t) = a^{-1/2}\phi\left(\frac{t+b}{a}\right)$$

where $a > 0$. (The factor $a^{-1/2}$ is chosen so that $||\phi^{(a,b)}|| = ||\phi||$.

Suppose $\int_{-\infty}^{\infty} t|\phi(t)|^2 dt = \ell$ and $k = \int_0^{\infty} y|\hat{\phi}(y)|^2 dy$. Then ϕ is centered about ℓin position space and about k in momentum space. It follows that

$$\int_{-\infty}^{\infty} t|\phi^{(a,b)}(t)|^2 dt = \ell - b, \quad \int_0^{\infty} y|\hat{\phi}^{(a,b)}(y)|^2 dy = a^{-1}k.$$

The affine translates $\phi^{(a,b)}$ are called **wavelets** and the function ϕ is a **father wavelet**. The map

$$\mathbf{T} : f \longrightarrow \langle f, \phi^{(ab)}\rangle = \int_{-\infty}^{\infty} f(t)\overline{\phi^{(ab)}(t)}dt \equiv F_\phi(a,b)$$

is the **continuous wavelet transform**

In order to invert $\mathbf{T}$ and synthesize f from the transform of a single mother wavelet ϕ we need to require that

$$\int_{-\infty}^{\infty} \phi(t)dt = 0. \tag{2.3}$$

Further, we require that $\phi(t)$ has exponential decay at ∞, i.e., $|\phi(t)| \le Ke^{-k|t|}$ for some $k, K > 0$ and all t. Among other things this implies that $|\hat{\phi}(\omega)|$ is uniformly bounded in ω. Then there is a Plancherel formula.

Theorem 10 *Let $f, g \in L^2[-\infty, \infty]$ and $C = 2\pi \int |\hat{\phi}(\omega)|^2 \frac{d\omega}{|\omega|}$. Then*

$$C\int_{-\infty}^{\infty} f(t)\overline{g}(t)dt = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F_\phi(a,b)\overline{G}_w(a,b)\frac{da\ db}{a^2}. \tag{2.4}$$

The synthesis equation for continuous wavelets is as follows.

Theorem 11

$$f(t) = \frac{1}{C}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F_\phi(a,b)|a|^{-1/2}\overline{\phi}(\frac{t-b}{a})\frac{db\ da}{a^2}. \tag{2.5}$$

To define a lattice we choose two nonzero real numbers $a_0, b_0 > 0$ with $a_0 \neq 1$. Then the lattice points are $a = a_0^m, b = nb_0 a_0^m$, $m, n = 0, \pm 1, \cdots$, so

$$\phi^{mn}(t) = \phi^{(a_0^m, nb_0 a_0^m)}(t) = a_0^{-m/2}\phi(a_0^{-m}t + nb_0).$$

Thus ϕ^{mn} is centered about $\ell - nb_0 a_0^m$ in position space and about $a_0^{-m}k$ in momentum space. (Note that this behavior is very different from the behavior of the windowed Fourier translates $g^{[ma,nb]}$. In the windowed case the support of g in either position or momentum space is the same as the support of $g^{[ma,nb]}$. In the wavelet case the sampling of position-momentum space is on a logarithmic scale. There is the possibility, through the choice of m and n, of sampling in smaller and smaller neighborhoods of a fixed point in position space.)

Again the continuous wavelet transform is overcomplete, as we shall see. The question is whether we can find a subgroup lattice and a function ϕ for which the functions

$$\phi^{mn}(t) = \phi^{(a_0^m, nb_0 a_0^m)}(t) = a_0^{-m/2}\phi(a_0^{-m}t + nb_0)$$

generate an ON basis. We will choose $a_0 = 1/2, b_0 = 1$ and find conditions such that the functions

$$\phi^{mn}(t) = 2^{m/2}\phi(2^m t + n), \qquad m, n = 0, \pm 1, \pm 2, \cdots$$

span L^2. In particular we require that the set $\phi^{0n}(t) = \phi(t+n)$ be orthonormal.

2.3 Discrete wavelets and the multiresolution structure

Our problem is to find a scaling function (or father wavelet) ϕ such that the functions $\phi^{mn}(t) = 2^{m/2}\phi(2^m t + n)$ will generate an ON basis for L^2. In particular we require that the set $\phi^{0n}(t) = \phi(t+n)$ be orthonormal. Then for each fixed m we have that $\{\phi^{mn}\}$ is ON in n. This leads to the concept of a multiresolution structure on L^2.

Definition 13 *Let $\{V_j : j = \cdots, -1, 0, 1, \cdots\}$ be a sequence of subspaces of $L^2[-\infty, \infty]$ and $\phi \in V_0$. This is a multiresolution analysis for $L^2[-\infty,\infty]$ provided the following conditions hold:*

1. *The subspaces are nested: $V_j \subset V_{j+1}$.*
2. *The union of the subspaces generates L^2 : $\overline{\cup_{j=-\infty}^{\infty} V_j} = L^2[-\infty,\infty]$. (Thus, each $f \in L^2$ can be obtained a a limit of a Cauchy sequence $\{s_n : n = 1, 2, \cdots\}$ such that each $s_n \in V_{j_n}$ for some integer j_n.)*
3. *Separation: $\cap_{j=-\infty}^{\infty} V_j = \{0\}$, the subspace containing only the zero function. (Thus only the zero function is common to all subspaces V_j.)*
4. *Scale invariance: $f(t) \in V_j \Longleftrightarrow f(2t) \in V_{j+1}$.*
5. *Shift invariance of V_0: $f(t) \in V_0 \Longleftrightarrow f(t-k) \in V_0$ for all integers k.*
6. *ON basis: The set $\{\phi(t-k) : k = 0, \pm 1, \cdots\}$ is an ON basis for V_0.*

Here, the function $\phi(t)$ is called the scaling function (or the father wavelet).

Of special interest is a multiresolution analysis with a scaling function $\phi(t)$ on the real line that has *compact support*. The functions $\phi(t+k)$ will form an ON basis for V_0 as k runs over the integers, and their integrals with any polynomial in t will be finite.

Example 2 *The Haar scaling function*

$$\phi(t) = \begin{cases} 1 & 0 \le t < 1 \\ 0 & \text{otherwise} \end{cases}$$

defines a multiresolution analysis. Here V_j is the space of piecewise constant functions with possible discontinuities only at the gridpoints $t_j = \frac{k}{2^{nj}}$, $k = 0, \pm 1, \pm 2 \cdots$.

2.3.1 Haar wavelets

The simplest wavelets are the Haar wavelets. They were studied by Haar more than 50 years before wavelet theory came into vogue. We start with the *father wavelet* or *scaling function*. For the Haar wavelets the scaling function is the *box function*

$$\phi(t) = \begin{cases} 1 & \text{if } 0 \le t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

We can use this function and its integer translates to construct the space V_0 of all step functions of the form

$$s(t) = a_k \qquad \text{for } k \le t < k+1,$$

where the a_k are complex numbers such that $\sum_{k=-\infty}^{\infty} |a_k|^2 < \infty$. Note that the $\{\phi(t-k) : \quad k = 0, \pm 1, \cdots\}$ form an ON basis for V_0. Also, the area under the father wavelet is 1:

$$\int_{-\infty}^{\infty} \phi(t)dt = 1.$$

We can approximate signals $f(t) \in L^2[-\infty, \infty]$ by projecting them on V_0 and then expanding the projection in terms of the translated scaling functions. Of course this would be a very crude approximation. To get more accuracy we can change the scale by a factor of 2.

Consider the functions $\phi(2t-k)$. They form a basis for the space V_1 of all step functions of the form

$$s(t) = a_k \qquad \text{for } \frac{k}{2} \le t < \frac{k+1}{2},$$

where $\sum_{k=-\infty}^{\infty}|a_k|^2<\infty$. This is a larger space than V_0 because the intervals on which the step functions are constant are just $1/2$ the width of those for V_0. The functions $\{2^{1/2}\phi(2t-k):\quad k=0,\pm1,\cdots\}$ form an ON basis for V_1.The scaling function also belongs to V_1. Indeed we can expand it in terms of the basis as

$$\phi(t)=\phi(2t)+\phi(2t-1). \tag{2.7}$$

We can continue this rescaling procedure and define the space V_j of step functions at level j to be the Hilbert space spanned by the linear combinations of the functions $\phi(2^jt-k),\quad k=0,\pm1,\cdots$. These functions will be piecewise constant with discontinuities contained in the set

$$\{t=\frac{n}{2^j},\qquad n=0,\pm1,\pm2,\cdots\}.$$

The functions

$$\phi_{jk}(t)=2^{\frac{j}{2}}\phi(2^jt-k),\qquad k=0,\pm1,\pm2,\cdots$$

form an ON basis for V_j. Further we have

$$V_0\subset V_1\subset\cdots\subset V_{j-1}\subset V_j\subset V_{j+1}\subset\cdots$$

and the containment is strict. (Each V_j contains functions that are not in V_{j-1}.) Also, note that the dilation equation (2.7) implies that

$$\phi_{jk}(t)=\frac{1}{\sqrt{2}}[\phi_{j+1,2k}(t)+\phi_{j+1,2k+1}(t)]. \tag{2.8}$$

Since $V_0\subset V_1$, it is natural to look at the orthogonal complement of V_0 in V_1, i.e., to decompose each $s\in V_1$ in the form $s=s_0+s_1$ where $s_0\in V_0$ and $s_1\in V_0^{\perp}$. We write

$$V_1=V_0\oplus W_0,$$

where $W_0=\{s\in V_1:\qquad (s,f)=0 \text{ for all } f\in V_0\}$. It follows that the functions in W_0 are just those in V_1 that are orthogonal to the basis vectors $\phi(t-k)$ of V_0.

Note from the dilation equation that $\phi(t-k)=\phi(2t-2k)+\phi(2t-2k-1)=2^{-1/2}\left(\phi_{1,2k}(t)+\phi_{1,2k+1}(t)\right)$. Thus

$$(\phi_{0k},\phi_{1\ell})=2^{1/2}\int_{-\infty}^{\infty}\phi(t-k)\phi(2t-\ell)dt=\begin{cases}2^{-1/2} & \text{if } \ell=2k,2k+1\\ 0 & \text{otherwise}\end{cases}$$

and

$$s_1(t) = \sum_k a_k \phi(2t-k) \in V_1$$

belongs to W_0 if and only if $a_{2k+1} = -a_{2k}$. Thus

$$s_1 = \sum_k a_{2k} \left[\phi(2t-2k) - \phi(2t-2k-1)\right] = \sum_k a_{2k} w(t-k)$$

where

$$w(t) = \phi(2t) - \phi(2t-1) \qquad (2.9)$$

is the *Haar wavelet*, or *mother wavelet*. You can check that the wavelets $w(t-k), \qquad k = 0 \pm 1, \cdots$ form an ON basis for W_0.

We define functions

$$w_{jk}(t) = 2^{\frac{j}{2}} w(2^j t - k) = 2^{\frac{j}{2}} (\phi(2^{j+1}t - 2k) - \phi(2^{j+1}t - 2(k+1)),$$

$$k = 0, \pm 1, \pm 2, \cdots, \quad j = 1, 2, \cdots.$$

It is easy to prove

Lemma 4 *For fixed j,*

$$(w_{jk}, w_{jk'}) = \delta_{kk'}, \qquad (\phi_{jk}, w_{jk'}) = 0 \qquad (2.10)$$

where $k, k' = 0, \pm 1, \cdots$.

Other properties proved above are

$$\phi_{jk}(t) = \frac{1}{\sqrt{2}} (\phi_{j+1,2k}(t) + \phi_{j+1,2k+1}(t)),$$

$$w_{jk}(t) = \frac{1}{\sqrt{2}} (\phi_{j+1,2k}(t) - \phi_{j+1,2k+1}(t)).$$

Theorem 12 *let W_j be the orthogonal complement of V_j in V_{j+1}:*

$$V_j \oplus W_j = V_{j+1}.$$

The wavelets $\{w_{jk}(t) : \quad k = 0, \pm 1, \cdots\}$ form an ON basis for W_j.

Since $V_j \oplus W_j = V_{j+1}$ for all $j \geq 0$, we can iterate on j to get $V_{j+1} = W_j \oplus V_j = W_j \oplus W_{j-1} \oplus V_{j-1}$ and so on. Thus

$$V_{j+1} = W_j \oplus W_{j-1} \oplus \cdots \oplus W_1 \oplus W_0 \oplus V_0.$$

and any $s \in V_{j+1}$ can be written uniquely in the form

$$s = \sum_{k=0}^{j} w_k + s_0 \qquad \text{where } w_k \in W_k,\ s_0 \in V_0.$$

Theorem 13

$$L^2[-\infty, \infty] = V_0 \oplus \sum_{k=0}^{\infty} W_k = V_0 \oplus W_0 \oplus W_1 \oplus \cdots,$$

so that each $f(t) \in L^2[-\infty, \infty]$ can be written uniquely in the form

$$f = f_0 + \sum_{k=0}^{\infty} w_k, \qquad w_k \in W_k,\ f_0 \in V_0. \tag{2.11}$$

We have a new ON basis for $L^2[-\infty, \infty]$:

$$\{\phi_{0k}, w_{jk'} : \qquad j, \pm k, \pm k' = 0, 1, \cdots\}.$$

Let's consider the space V_j for fixed j. On one hand we have the scaling function basis

$$\{\phi_{j,k} : \qquad \pm k = 0, 1, \cdots\}.$$

Then we can expand any $f_j \in V_j$ as

$$f_j = \sum_{k=-\infty}^{\infty} a_{j,k} \phi_{j,k}. \tag{2.12}$$

On the other hand we have the wavelets basis

$$\{\phi_{j-1,k}, w_{j-1,k'} : \qquad \pm k, \pm k' = 0, 1, \cdots\}$$

associated with the direct sum decomposition

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

Using this basis we can expand any $f_j \in V_j$ as

$$f_j = \sum_{k'=-\infty}^{\infty} b_{j-1,k'} w_{j-1,k'} + \sum_{k=-\infty}^{\infty} a_{j-1,k} \phi_{j-1,k}. \tag{2.13}$$

If we substitute the relations

$$\phi_{j-1,k}(t) = \frac{1}{\sqrt{2}}(\phi_{j,2k}(t) + \phi_{j,2k+1}(t)),$$

$$w_{j-1,k}(t) = \frac{1}{\sqrt{2}}(\phi_{j,2k}(t) - \phi_{j,2k+1}(t))$$

into the expansion (2.13) and compare coefficients of $\phi_{j,\ell}$ with the expansion (2.12), we obtain the fundamental recursions

$$\text{Averages(lowpass)} \quad a_{j-1,k} = \tfrac{1}{\sqrt{2}}(a_{j,2k} + a_{j,2k+1}) \tag{2.14}$$

$$\text{Differences(highpass)} \quad b_{j-1,k} = \tfrac{1}{\sqrt{2}}(a_{j,2k} - a_{j,2k+1}). \tag{2.15}$$

We can iterate this process by inputting the output $a_{j-1,k}$ to the recursion again to compute $a_{j-2,k}, b_{j-2,k}$, etc. At each stage we save the wavelet coefficients $b_{j'k'}$ and input the scaling coefficients $a_{j'k'}$ for further processing, see Figure 2.1.

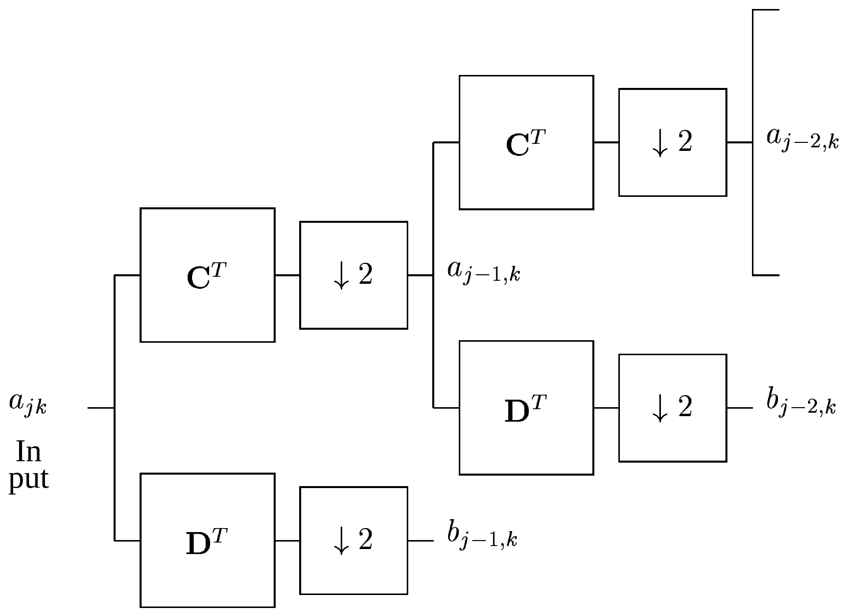

Figure 2.1: Fast Wavelet Transform

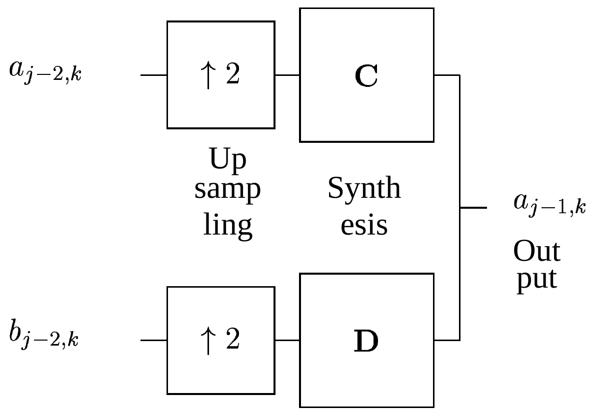

Figure 2.2: Haar wavelet inversion

The output of the final stage is the set of scaling coefficients a_{0k}. Thus our final output is the complete set of coeffients for the wavelet expansion

$$f_j = \sum_{j'=0}^{j} \sum_{k=-\infty}^{\infty} b_{j'k} w_{j'k} + \sum_{k=-\infty}^{\infty} a_{0k} \phi_{0k},$$

based on the decomposition

$$V_{j+1} = W_j \oplus W_{j-1} \oplus \cdots \oplus W_1 \oplus W_0 \oplus V_0.$$

The synthesis recursion is :

$$\begin{aligned} a_{j,2k} &= \frac{1}{\sqrt{2}}(a_{j-1,k} + b_{j-1,k}) \\ a_{j,2k+1} &= \frac{1}{\sqrt{2}}(a_{j-1,k} - b_{j-1,k}). \end{aligned} \tag{2.16}$$

This is exactly the output of the synthesis filter bank shown in Figure 2.2.

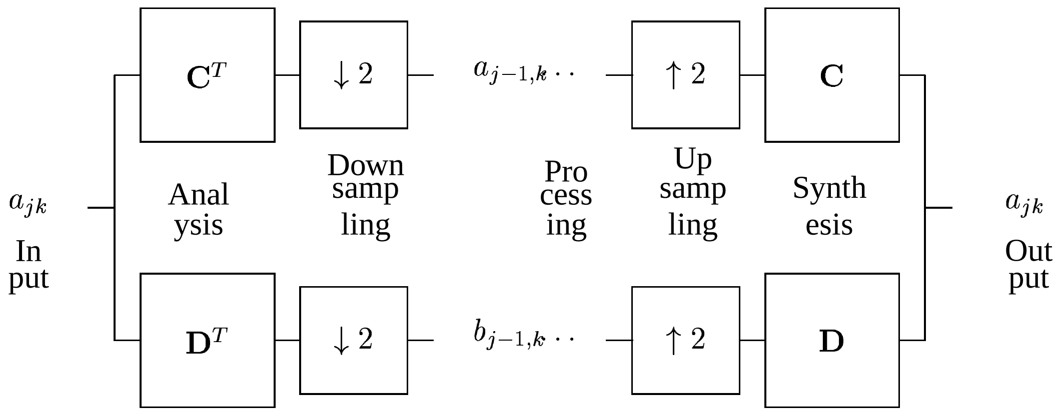

Figure 2.3: Fast Wavelet Transform and Inversion

Thus, for level j the full analysis and reconstruction picture is Figure 2.3.

COMMENTS ON HAAR WAVELETS:

1. For any $f(t) \in L^2[-\infty, \infty]$ the scaling and wavelets coefficients of f are defined by

$$
\begin{aligned}
a_{jk} &= (f, \phi_{jk}) = 2^{j/2} \int_{-\infty}^{\infty} f(t)\phi(2^j t - k)dt \\
&= 2^{j/2} \int_{\frac{k}{2^j}}^{\frac{k}{2^j}+\frac{1}{2^j}} f(t)dt, \qquad (2.17) \\
b_{jk} &= (f, w_{jk}) = 2^{j/2} \int_{-\infty}^{\infty} f(t)\phi(2^{j+1}t - 2k)dt \\
&- 2^{j/2} \int_{-\infty}^{\infty} f(t)\phi(2^{j+1}t - 2k - 1)dt \\
&= 2^{j/2} \int_{\frac{k}{2^j}}^{\frac{k}{2^j}+\frac{1}{2^j}} [f(t) - f(t + \frac{1}{2^{j+1}})]dt. \qquad (2.18)
\end{aligned}
$$

If f is a continuous function and j is large then $a_{jk} \sim 2^{-j/2} f(\frac{k}{2^j})$. (Indeed if f has a bounded derivative we can develop an upper bound for the error of this approximation.) If f is continuously differentiable and j is large, then $b_{jk} \sim -\frac{1}{2^{1+3j/2}} f'(\frac{k}{2^j})$. Again this shows that the a_{jk} capture averages of f (low pass) and the b_{jk} capture changes in f (high pass).

2. Since the scaling function $\phi(t)$ is nonzero only for $0 \le t < 1$ it follows that $\phi_{jk}(t)$ is nonzero only for $\frac{k}{2^j} \le t < \frac{k}{2^j} + \frac{1}{2j}$. Thus the coefficients a_{jk} depend only on the local behavior of $f(t)$ in that interval. Similarly for the wavelet coefficients b_{jk}. This is a dramatic difference from Fourier series or Fourier integrals where each coefficient depends on the global behavior of f. If f has compact support, then for fixed j, only a finite number of the coefficients a_{jk}, b_{jk} will be nonzero. The Haar coefficients a_{jk} enable us to track t intervals where the function becomes nonzero or large. Similarly the coefficients b_{jk} enable us to track t intervals in which f changes rapidly.

3. Given a signal f, how would we go about computing the wavelet coefficients? As a practical matter, one doesn't usually do this by evaluating the integrals (2.17) and (2.18). Suppose the signal has compact support. By translating and rescaling the time coordinate if necessary, we can assume that $f(t)$ vanishes except in the interval $[0, 1)$. Since $\phi_{jk}(t)$ is nonzero only for $\frac{k}{2^j} \le t < \frac{k}{2^j} + \frac{1}{2j}$ it follows that all of the coefficients a_{jk}, b_{jk} will vanish except when $0 \le k < 2^j$. Now suppose that f is such that for a sufficiently

large integer $j = J$ we have $a_{Jk} \sim 2^{-J/2} f(\frac{k}{2^J})$. If f is differentiable we can compute how large J needs to be for a given error tolerance. We would also want to exceed the Nyquist rate. Another possibility is that f takes discrete values on the grid $t = \frac{k}{2^J}$, in which case there is no error in our assumption.

Inputing the values $a_{Jk} = 2^{-J/2} f(\frac{k}{2^J})$ for $= 0, 1, \cdots, 2^J - 1$ we use the recursion

$$\text{Averages (lowpass)} \quad a_{j-1,k} = \frac{1}{\sqrt{2}}(a_{j,2k} + a_{j,2k+1}) \tag{2.19}$$

$$\text{Differences (highpass)} \quad b_{j-1,k} = \frac{1}{\sqrt{2}}(a_{j,2k} - a_{j,2k+1}). \tag{2.20}$$

described above, to compute the wavelet coefficients b_{jk}, $j = 0, 1, \cdots, J - 1, \quad k = 0, 1, \cdots 2^j - 1$ and a_{00}.

The input consists of 2^J numbers. The output consists of $\sum_{j=0}^{J-1} 2^j + 1 = 2^J$ numbers. The algorithm is very efficient. Each recurrence involves 2 multiplications by the factor $\frac{1}{\sqrt{2}}$. At level j there are $2 \cdot 2^j$ such recurrences. thus the total number of multiplications is $2\sum_{j=0}^{J-1} 2 \cdot 2^j = 4 \cdot 2^J - 4 < 4 \cdot 2^J$.

4. The preceeding algorithm is an example of the Fast Wavelet Transform (FWT). It computes 2^J wavelet coefficients from an input of 2^J function values and does so with a number of multipications $\sim 2^J$. Compare this with the FFT which needs $\sim J \cdot 2^J$ multiplications from an input of 2^J function values. In theory at least, the FWT is faster. The Inverse Fast Wavelet Transform is based on (2.16). (Note, however, that the FFT and the FWT compute dfiferent things. They divide the spectral band in different ways. Hence they aren't directly comparable.)

5. The FWT discussed here is based on filters with $N + 1$ taps, where $N = 1$. For wavelets based on more general $N+1$ tap filters (such as the Daubechies filters) , each recursion involves $N+1$ multiplications, rather than 2. Otherwise the same analysis goes through. Thus the FWT requires $\sim 2(N+1)2^J$ multiplications.

6. Haar wavelets are very simple to implement. However they are terrible at approximating continuous functions. By definition, any truncated Haar wavelet expansion is a step function. The Daubechies wavelets to come are continuous and are much better for this type of approximation.

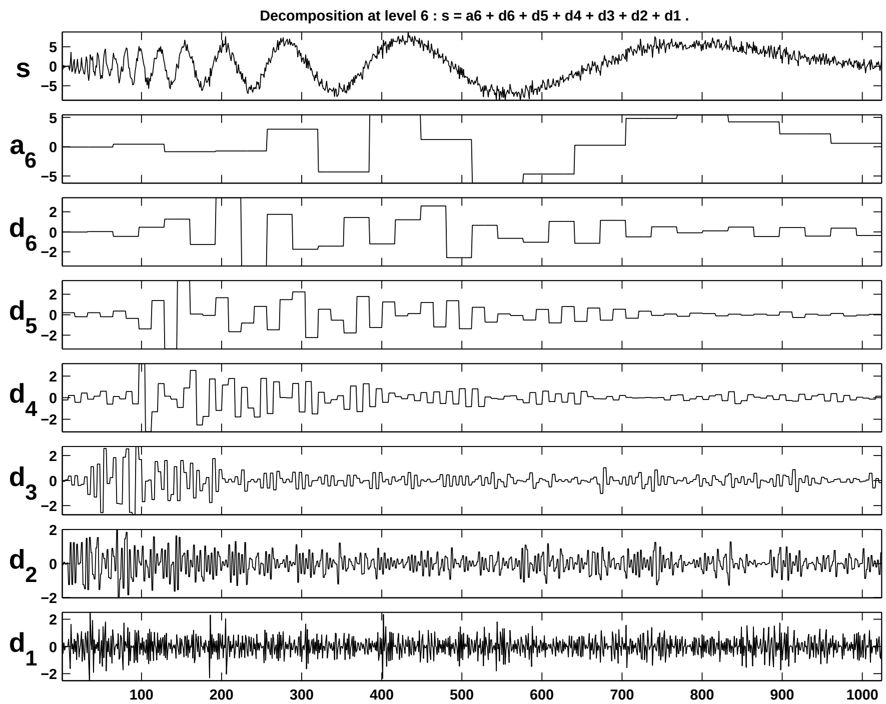

Figure 2.4: Haar Analysis of a Signal

This is output from the Wavelet Toolbox of Matlab. The signal $s = a_0$ is sampled at $1024 = 2^{10}$ points, so $J = 10$ and s is assumed to be in the space V_{10}. The signal is taken to be zero at all points $k/2^{10}$, except for $k = 0, 1, \cdots, 2^{10} - 1$. The approximations a_ℓ (the averages) are the projections of s on the subspaces $V_{10-\ell}$ for $\ell = 1, \cdots, 6$. The lowest level approximation a_6 is the projection on the subspace V_4. There are only 16 distinct values at this lowest level. The approximations d_ℓ (the differences) are the projections of s on the wavelet subspaces $W_{10-\ell}$.

Chapter 3

Lecture III

3.1 Continuous scaling functions with compact support

We continue our exploration of multiresolution analysis for some scaling function $\phi(t)$, with a particular interest in finding such functions that are continuous (or even smoother) and have compact support. Given $\phi(t)$ we can define the functions

$$\phi_{jk}(t) = 2^{\frac{j}{2}}\phi(2^j t - k), \qquad k = 0, \pm 1, \pm 2, \cdots$$

and for fixed integer j they will form an ON basis for V_j. Since $V_0 \subset V_1$ it follows that $\phi \in V_1$ and ϕ can be expanded in terms of the ON basis $\{\phi_{1k}\}$ for V_1. Thus we have the *dilation equation*

$$\phi(t) = \sqrt{2}\sum_k \mathbf{c}(k)\phi(2t-k),$$

or, equivalently,

$$\phi(t) = 2\sum_{k=0}^{N} \mathbf{h}(k)\phi(2t-k)$$

where $\mathbf{h}(k) = \frac{1}{\sqrt{2}}\mathbf{c}(k)$. Since the ϕ_{jk} form an ON set, the coefficient vector $\mathbf{c}$ must be a unit vector in ℓ^2,

$$\sum_k \mathbf{c}(k)^2 = 1.$$

Since $\phi(t) \perp \phi(t-m)$ for all nonzero m, the vector $\mathbf{c}$ satisfies the orthogonality relation:

$$(\phi_{00}, \phi_{0m}) = \sum_k \mathbf{c}(k)\mathbf{c}(k-2m) = \delta_{0m}.$$

Lemma 5 *If the scaling function is normalised so that*

$$\int_{-\infty}^{\infty} \phi(t)dt = 1,$$

then $\sum_{k=0}^{N} \mathbf{c}(k) = \sqrt{2}$.

We can introduce the orthogonal complement W_j of V_j in V_{j+1}.

$$V_{j+1} = V_j \oplus W_j.$$

We start by trying to find an ON basis for the wavelet space W_0. Associated with the father wavelet $\phi(t)$ there must be a mother wavelet $w(t) \in W_0$, with norm 1, and satisfying the *wavelet equation*

$$w(t) = \sqrt{2} \sum_k \mathbf{d}(k)\phi(2t-k),$$

and such that w is orthogonal to all translations $\phi(t-k)$ of the father wavelet. We further require that w is orthogonal to integer translations of itself. Since the ϕ_{jk} form an ON set, the coefficient vector $\mathbf{d}$ must be a unit vector in ℓ^2,

$$\sum_k |\mathbf{d}(k)|^2 = 1.$$

Moreover since $w(t) \perp \phi(t-m)$ for all m, the vector $\mathbf{d}$ satisfies so-called double-shift orthogonality with $\mathbf{c}$:

$$(w, \phi_{0m}) = \sum_k \mathbf{c}(k)\mathbf{d}(k-2m) = 0. \quad (3.1)$$

The requirement that $w(t) \perp w(t-m)$ for nonzero integer m leads to double-shift orthogonality of $\mathbf{d}$ to itself:

$$(w(t), w(t-m)) = \sum_k \mathbf{d}(k)\mathbf{d}(k-2m) = \delta_{0m}. \quad (3.2)$$

However, if the unit coefficient vector $\mathbf{c}$ is double-shift orthogonal then the coefficient vector $\mathbf{d}$ defined by

$$\mathbf{d}(n) = (-1)^n \mathbf{c}(N-n). \quad (3.3)$$

automatically satisfies the conditions (3.1) and (3.2).

The coefficient vector $\mathbf{c}(k)$ must satisfy the following necessary conditions in order to define a multiresolution analysis whose scaling function is continuous and has compact support.

1. $$\sum_k \mathbf{c}(k)^2 = 1$$

2. $$\sum_k \mathbf{c}(k)\mathbf{c}(k-2m) = \delta_{0m}$$

3. $$\sum_{k=0}^{N} \mathbf{c}(k) = \sqrt{2}$$

4. $\mathbf{c}(k) = 0$ unless $0 \le k \le N$ for some finite odd integer N

5. For maximum smoothness of the scaling function with fixed N the filter $C(\omega) = \sum_k \mathbf{c}(k)e^{ik\omega}$ should be *maxflat*, i.e., $C^{(\ell)}(\pi) = 0$ for $k = 0, 1, \cdots, (N-1)/2$.

Results:

1. For $N = 1$ we can easily solve these equations to get $\mathbf{c}(0) = \mathbf{c}(1) = 1/\sqrt{(}2)$, corresponding to the Haar wavelets.

2. For $N = 3$ they are also straightforward to solve The nonzero Daubechies filter coefficients for D_4 ($N = 3$) are $4\sqrt{2}\mathbf{c}(k) = 1+\sqrt{3}, 3+\sqrt{3}, 3-\sqrt{3}, 1-\sqrt{3}$.

3. For $N = 5$ they have just been solved explicitly using a computer algebra package.

4. In general there are no explicit solutions. (We would need to know how to find explicit roots of polynomial equations of arbitrarily high order.) However, in 1989, Ingrid Daubechies exhibited a unique solution for each odd integer N. The coefficients $\mathbf{c}(k)$ can be approximated numerically.

To find compact support wavelets must find solutions $\mathbf{c}(k)$ of the orthogonality relations above, nonzero for a *finite* range $k = 0, 1, \cdots, N$. Then given a solution $\mathbf{c}(k)$ must solve the dilation equation

$$\phi(t) = \sqrt{2} \sum_k \mathbf{c}(k)\phi(2t - k). \tag{3.4}$$

to get $\phi(t)$. Can show that the support of $\phi(t)$ must be contained in the interval $[0, N)$.

- **Cascade Algorithm** One way to try to determine a scaling function $\phi(t)$ from the impulse response vector $\mathbf{c}$ is to iterate the dilation equation. That is, we start with an initial guess $\phi^{(0)}(t)$, the Haar scaling function on $[0, 1)$, and then iterate

  $$\phi^{(i+1)}(t) = \sqrt{2} \sum_{k=0}^{N} \mathbf{c}(k)\phi^{(i)}(2t - k) \tag{3.5}$$

  This is called the *cascade algorithm*.

- **Frequency domain** The frequency domain formulation of the dilation equation is :

  $$\hat{\phi}(\omega) = \left(\sum_k \mathbf{h}(k) e^{-i\omega k/2} \right) \hat{\phi}(\frac{\omega}{2})$$

  where $\mathbf{c}(k) = \sqrt{2}\mathbf{h}(k)$. Thus

  $$\hat{\phi}(\omega) = H(\frac{\omega}{2})\hat{\phi}(\frac{\omega}{2}).$$

  where

  $$H(\omega) = \sum_{k=0}^{N} \mathbf{h}(k) e^{-i\omega k}$$

  Iteration yields the explicit infinite product formula: $\hat{\phi}(\omega)$:

  $$\hat{\phi}(\omega) = \Pi_{j=1}^{\infty} H(\frac{\omega}{2^j}). \tag{3.6}$$

L^2 CONVERGENCE OF THE CASCADE ALGORITHM

We want to find conditions that guarantee that the iterates $\phi^{(i)}(t)$ converge in the $L^2(R)$ norm, i.e., that $\{\phi^{(i)}\}$ is a Cauchy sequence in the norm. Thus, we want to show that for any $\epsilon > 0$ there is an integer N_ϵ such that

$$||\phi^{(i)} - \phi^{(j)}||2 = (\phi^{(i)}, \phi^{(i)}) - 2(\phi^{(i)}, \phi^{(j)}) + (\phi^{(j)}, \phi^{(j)}) < \epsilon$$

whenever $i, j \geq N_\epsilon$. Then, since $L^2(R)$ is closed, there will be a $\phi(t) \in L^2(R)$ such that

$$||\phi - \phi^{(i)}|| \to 0 \quad \text{as } i \to \infty.$$

It isn't difficult to show that this will be the case provided the inner products

$$\mathbf{a}^{(i)}(k) = (\phi^{(i)}_{00}, \phi^{(i)}_{0k}) = \int_{-\infty}^{\infty} \phi^{(i)}(t)\phi^{(i)}(t-k)dt, \tag{3.7}$$

converge to δ_{0k} as $i \to \infty$. We can compute the transformation that relates the inner products

$$\int_{-\infty}^{\infty} \phi^{(i+1)}(t)\phi^{(i+1)}(t-k)dt$$

to the inner products $\int_{-\infty}^{\infty} \phi^{(i)}(t)\phi^{(i)}(t-k)dt$ in successive passages through the cascade algorithm. Note that although $\mathbf{a}^{(i)}$ is an infinite-component vector, since $\phi^{(i)}(t)$ has support limited to the interval $[0, N]$ only the $2N-1$ components $\mathbf{a}^{(i)}(k)$, $k = -N+1, \cdots - 1, 0, 1 \cdots, N-1$ can be nonzero. We can use the cascade recursion to express $\mathbf{a}^{(i+1)}(s)$ as a linear combination of terms $\mathbf{a}^{(i)}(k)$:

$$\mathbf{a}^{(i+1)}(s) = \int_{-\infty}^{\infty} \phi^{(i)(t)}\phi^{(i)}(t+s)dt$$

$$= 4\sum_{k,\ell} \mathbf{h}(k)\mathbf{h}(\ell) \int_{-\infty}^{\infty} \phi^{(i)}(t-k)\phi^{(i)}(2t+2s-\ell)dt$$

$$= 2\sum_{j,\ell} \mathbf{h}(2s-j)\mathbf{h}(\ell-j) \int_{-\infty}^{\infty} \phi^{(i)}(t)\phi^{(i)}(t+\ell)dt.$$

Thus

$$\mathbf{a}^{(i+1)}(s) = 2\sum_{j,\ell} \mathbf{h}(2s+j)\mathbf{h}(\ell+j)\mathbf{a}^{(i)}(\ell) \tag{3.8}$$

In matrix notation this is just

$$\mathbf{a}^{(i+1)} = \mathbf{T}\mathbf{a}^{(i)} = (\downarrow 2)2\mathbf{H}\mathbf{H}^{\mathrm{tr}}\mathbf{a}^{(i)} \tag{3.9}$$

where the matrix elements of the $\mathbf{T}$ matrix (the *transition matrix*) are given by

$$\mathbf{T}_{s\ell} = 2\sum_j \mathbf{h}(2s+j)\mathbf{h}(\ell+j).$$

Although $\mathbf{T}$ is an infinite matrix, the only elements that correspond to inner products of functions with support in $[0, N]$ are contained in the $(2N-1)\times(2N-1)$ block $-N+1 \le s, \ell \le N-1$. When we discuss the eigenvalues and eigenvectors of $\mathbf{T}$ we are normally talking about this $(2N-1)\times(2N-1)$ matrix.

If we apply the cascade algorithm to the inner product vector of the scaling function itself

$$\mathbf{a}(k) = (\phi_{00}, \phi_{0k}) = \int_{-\infty}^{\infty} \phi(t)\phi(t-k)dt,$$

we just reproduce the inner product vector:

$$\mathbf{a}(s) = 2\sum_{j,\ell} \mathbf{h}(2s-j)\mathbf{h}(\ell-j)\mathbf{a}(\ell), \tag{3.10}$$

or

$$\mathbf{a} = \mathbf{T}\mathbf{a} = (\downarrow 2)2\mathbf{H}\mathbf{H}^{\mathrm{tr}}\mathbf{a} \tag{3.11}$$

Since $\mathbf{a}(k) = \delta_{0k}$ in the orthogonal case, this justs says that

$$1 = 2\sum_j \mathbf{h}(j)^2,$$

which we already know to be true. Thus $\mathbf{T}$ always has 1 as an eigenvalue, with associated eigenvector $\mathbf{a}(k) = \delta_{0k}$.

3.2 L^2 convergence

The necessary and sufficient condition for the cascade algorithm to converge in L^2 to a unique solution of the dilation equation is that the transition matrix $\mathbf{T}$ has a non-repeated eigenvalue 1 and all other eigenvalues λ such that $|\lambda| < 1$. Since the only nonzero part of $\mathbf{T}$ is a $(2N-1)\times(2N-1)$ block with very special structure, this is something that can be checked in practice.

Theorem 14 *The infinite matrix* $\mathbf{T} = (\downarrow 2)2\mathbf{H}\mathbf{H}^{\mathrm{tr}}$ *and its finite submatrix* $\mathbf{T}_{2N-1}$ *always have* $\lambda = 1$ *as an eigenvalue. The cascade iteration* $\mathbf{a}^{(i+1)} = \mathbf{T}\mathbf{a}^{(i)}$ *converges in* ℓ^2 *to the eigenvector* $\mathbf{a} = \mathbf{T}\mathbf{a}$ *if and only if the following condition is satisfied:*

- *All of the eigenvalues* λ *of* $\mathbf{T}_{2N-1}$ *satisfy* $|\lambda| < 1$ *except for the simple eigenvalue* $\lambda = 1$.

PROOF: let λ_j be the $2N-1$ eigenvalues of $\mathbf{T}_{2N-1}$, including multiplicities. Then there is a basis for the space of $2N-1$-tuples with respect to which $\mathbf{T}_{2N-1}$ takes the *Jordan canonical form*

$$\tilde{\mathbf{T}}_{2N-1} = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_p & & & & \\ & & & A_{p+1} & & & \\ & & & & A_{p+2} & & \\ & & & & & \ddots & \\ & & & & & & A_{p+q} \end{pmatrix}$$

where the Jordan blocks look like

$$A_s = \begin{pmatrix} \lambda_s & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_s & 1 & \cdots & 0 & 0 \\ \cdots & & & & & \cdots \\ 0 & 0 & 0 & \cdots & \lambda_s & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_s \end{pmatrix}.$$

If the eigenvectors of $\mathbf{T}_{2N-1}$ form a basis, for example if there were $2N-1$ distinct eigenvalues, then with respect to this basis $\tilde{\mathbf{T}}_{2N-1}$ would be diagonal and there would be no Jordan blocks. In general, however, there may not be enough eigenvectors to form a basis and the more general Jordan form will hold, with Jordan blocks. Now suppose we perform the cascade recursion n times. Then the action of the iteration on the base space will be

$$\tilde{\mathbf{T}}^n_{2N-1} = \begin{pmatrix} \lambda_1^n & & & & & & \\ & \ddots & & & & & \\ & & \lambda_p^n & & & & \\ & & & A^n_{p+1} & & & \\ & & & & A^n_{p+2} & & \\ & & & & & \ddots & \\ & & & & & & A^n_{p+q} \end{pmatrix}$$

where

$$A_s^n = \begin{pmatrix} \lambda_s^n & \binom{n}{1}\lambda_s^{n-1} & \binom{n}{2}\lambda_s^{n-2} & \cdots & \binom{n}{m_s-2}\lambda_s^{n-m_s+2} & \binom{n}{m_s-1}\lambda_s^{n-m_s+1} \\ 0 & \lambda_s^n & & \cdots & \binom{n}{m_s-3}\lambda_s^{n-m_s+3} & \binom{n}{m_s-2}\lambda_s^{n-m_s+2} \\ \cdots & & & & & \cdots \\ 0 & 0 & 0 & \cdots & \lambda_s^n & \binom{n}{1}\lambda_s^{n-1} \\ 0 & 0 & 0 & \cdots & 0 & \lambda_s^n \end{pmatrix}.$$

and A_s is an $m_s \times m_s$ matrix and m_s is the multiplicity of the eigenvalue λ_s. If there is an eigenvalue with $|\lambda_j| > 1$ then the corresponding terms in the power matrix will blow up and the cascade algorithm will fail to converge. (Of course if the original input vector has zero components corresponding to the basis vectors with these eigenvalues and the computation is done with perfect accuracy, one might have convergence. However, the slightest deviation, such as due to roundoff error, would introduce a component that would blow up after repeated iteration. Thus in practice the algorithm would diverge. With perfect accuracy and filter coefficients that satisfy double-shift orthogonality, one can maintain orthogonality of the shifted scaling functions at each pass of the cascade algorithm if orthogonality holds for the initial step. However, if the algorithm diverges, this theoretical result is of no practical importance. Roundoff error would lead to meaningless results in successive iterations.)

Similarly, if there is a Jordan block corresponding to an eigenvalue $|\lambda_j| = 1$ then the algorithm will diverge. If there is no such Jordan block, but there is more than one eigenvalue with $|\lambda_j| = 1$ then there may be convergence, but it won't be unique and will differ each time the algorithm is applied. If, however, all eigenvalues satisfy $|\lambda_j| < 1$ except for the single eigenvalue $\lambda_1 = 1$, then in the limit as $n \to \infty$ we have

$$\lim_{n\to\infty} \tilde{\mathbf{T}}_{2N-1}^n = \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix}$$

and there is convergence to a unique limit. Q.E.D.

Example 3 *The nonzero Daubechies filter coefficients for D_4 ($N = 3$) are*

$$4\sqrt{2}\mathbf{c}(k) = 1+\sqrt{3}, 3+\sqrt{3}, 3-\sqrt{3}, 1-\sqrt{3}.$$

The finite T matrix for this filter has the form

$$T = \begin{pmatrix} 0 & \frac{-1}{16} & 0 & 0 & 0 \\ 1 & \frac{9}{16} & 0 & \frac{-1}{16} & 0 \\ 0 & \frac{9}{16} & 1 & \frac{9}{16} & 0 \\ 0 & \frac{-1}{16} & 0 & \frac{9}{16} & 1 \\ 0 & 0 & 0 & \frac{-1}{16} & 0 \end{pmatrix}.$$

The vector

$$\mathbf{a}(k) = \int \phi(t)\phi(t-k)dt$$

is an eigenvector of this matrix with eigenvalue 1*. By looking at column 3 of* T *we can see that this eigenvector is* $\mathbf{a}(k) = \delta_{0k}$*, so we have orthonormal wavelets if the algorithm converges. The Jordan form for* T *is*

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{8} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 1 \\ 0 & 0 & 0 & 0 & \frac{1}{4} \end{pmatrix}.$$

so the eigenvalues of T *are* $1, \frac{1}{2}, \frac{1}{8}\frac{1}{4}, \frac{1}{4}$*. and the algorithm converges to give an* $L^2(R)$ *scaling function* $\phi(t)$*.*

To get the wavelet expansions for functions $f \in L^2$ we can now follow the steps in the construction for the Haar wavelets. The proofs are virtually identical. Since $V_j \oplus W_j = V_{j+1}$ for all $j \geq 0$, we can iterate on j to get $V_{j+1} = W_j \oplus V_j = W_j \oplus W_{j-1} \oplus V_{j-1}$ and so on. Thus

$$V_{j+1} = W_j \oplus W_{j-1} \oplus \cdots \oplus W_1 \oplus W_0 \oplus V_0.$$

and any $s \in V_{j+1}$ can be written uniquely in the form

$$s = \sum_{k=0}^{j} w_k + s_0 \qquad \text{where } w_k \in W_k,\ s_0 \in V_0.$$

Theorem 15

$$L^2[-\infty, \infty] = V_j \oplus \sum_{k=j}^{\infty} W_k = V_j \oplus W_j \oplus W_{j+1} \oplus \cdots,$$

so that each $f(t) \in L^2[-\infty, \infty]$ can be written uniquely in the form

$$f = f_j + \sum_{k=j}^{\infty} w_k, \qquad w_k \in W_k,\ f_j \in V_j. \tag{3.12}$$

We have a family of new ON bases for $L^2[-\infty, \infty]$, one for each integer j:

$$\{\phi_{jk}, w_{j'k'} : \qquad j' = j, j+1, \cdots, \quad \pm k, \pm k' = 0, 1, \cdots\}.$$

Let's consider the space V_j for fixed j. On one hand we have the scaling function basis

$$\{\phi_{j,k} : \qquad \pm k = 0, 1, \cdots\}.$$

Then we can expand any $f_j \in V_j$ as

$$f_j = \sum_{k=-\infty}^{\infty} a_{j,k}\phi_{j,k}. \tag{3.13}$$

On the other hand we have the wavelets basis

$$\{\phi_{j-1,k}, w_{j-1,k'} : \qquad \pm k, \pm k' = 0, 1, \cdots\}$$

associated with the direct sum decomposition

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

Using this basis we can expand any $f_j \in V_j$ as

$$f_j = \sum_{k'=-\infty}^{\infty} b_{j-1,k'} w_{j-1,k'} + \sum_{k=-\infty}^{\infty} a_{j-1,k}\phi_{j-1,k}. \tag{3.14}$$

If we substitute the relations

$$\phi_{j-1,\ell} = \sum_k \mathbf{c}(k - 2\ell)\phi_{jk}(t), \tag{3.15}$$

$$w_{j-1,\ell} = \sum_k \mathbf{d}(k - 2\ell)\phi_{j,k}(t), \tag{3.16}$$

into the expansion (3.13) and compare coefficients of $\phi_{j,\ell}$ with the expansion (3.14), we obtain the fundamental recursions

$$\text{Averages(lowpass)} \quad a_{j-1,k} = \textstyle\sum_n \mathbf{c}(n - 2k)a_{jn} \tag{3.17}$$

$$\text{Differences(highpass)} \quad b_{j-1,k} = \textstyle\sum_n \mathbf{d}(n - 2k)a_{jn}. \tag{3.18}$$

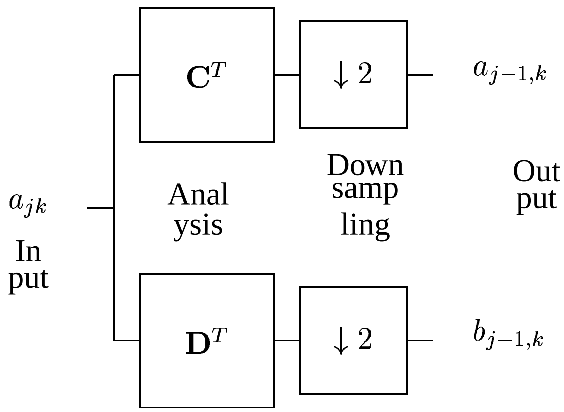

Figure 3.1: Wavelet Recursion

The picture, in complete analogy with that for Haar wavelets, is in Figure 3.1.

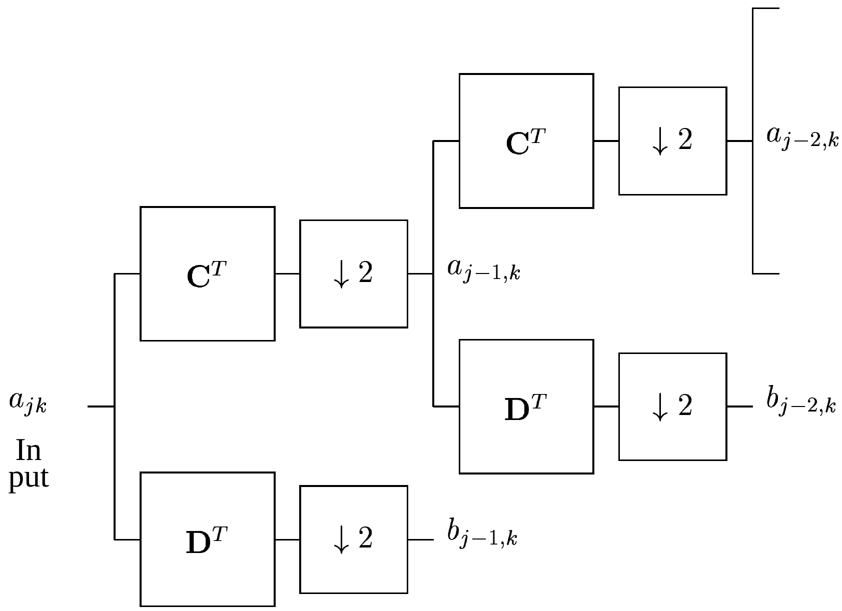

Figure 3.2: General Fast Wavelet Transform

We can iterate this process by inputting the output $a_{j-1,k}$ to the recursion again to compute $a_{j-2,k}$, $b_{j-2,k}$, etc. At each stage we save the wavelet coefficients $b_{j'k'}$ and input the scaling coefficients $a_{j'k'}$ for further processing, see Figure 3.2.

The output of the final stage is the set of scaling coefficients a_{0k}, assuming that we stop at $j = 0$. Thus our final output is the complete set of coeffients for the wavelet expansion

$$f_j = \sum_{j'=0}^{j-1} \sum_{k=-\infty}^{\infty} b_{j'k} w_{j'k} + \sum_{k=-\infty}^{\infty} a_{0k} \phi_{0k},$$

based on the decomposition

$$V_j = W_{j-1} \oplus W_{j-2} \oplus \cdots \oplus W_1 \oplus W_0 \oplus V_0.$$

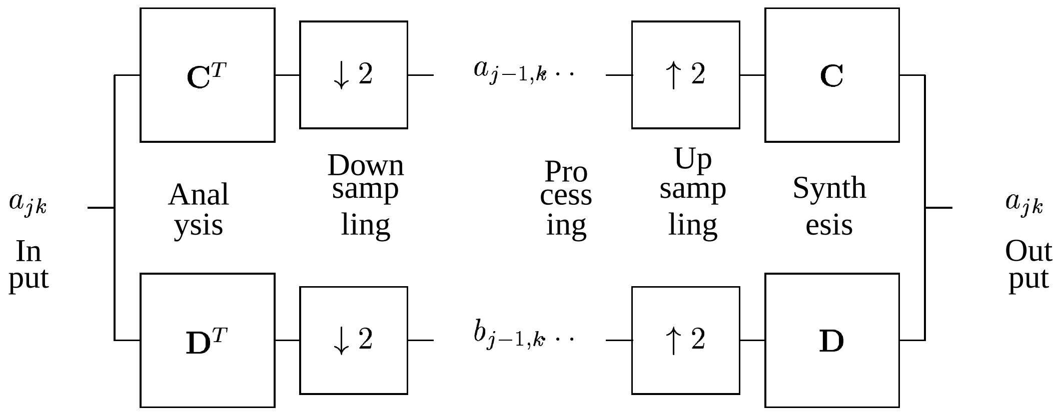

Figure 3.3: General Fast Wavelet Transform and Inversion

For level j the full analysis and reconstruction picture is Figure 3.3.

In analogy with the Haar wavelets discussion, for any $f(t) \in L^2[-\infty, \infty]$ the scaling and wavelets coefficients of f are defined by

$$\begin{aligned} a_{jk} &= (f, \phi_{jk}) = 2^{j/2} \int_{-\infty}^{\infty} f(t)\phi(2^j t - k)dt, \qquad (3.19) \\ b_{jk} &= (f, w_{jk}) = 2^{j/2} \int_{-\infty}^{\infty} f(t)w(2^j t - k)dt \end{aligned}$$

RESULTS:

- Daubechies has found a solution $\mathbf{c}(k)$ and the associated $L^2(R)$ scaling function for each $N = 1, 3, 5, \cdots$. (There are no solutions for even N.) Denote these solutions by $D_M = D_{N+1} = D_{2p}$. D_2 is just the Haar function. Daubechies finds the unique solutions for which the Fourier transform of the impulse response vector $C(\omega)$ has a zero of order p at $\omega = \pi$, where $2p = N + 1$. (At each N this is the maximal possible value for p.)

- Can compute the values of $\phi(t)$ exactly at all dyadic points $t = \sum_n \frac{j_n}{2^n}$, $j_n = \pm 1$.

- $\int \phi(t)dt = 1$, $\sum_k \phi(\frac{k}{2^j}) = 2^j$ for $j = 0, 1, 2, \cdots$.

- Can find explicit expressions

  $$\sum_k \mathbf{y}_{\ell k}\phi(t+k) = t^\ell, \qquad \ell = 0, 1, \cdots, p-1,$$

  so polynomials in t of order $\leq p - 1$ can be expressed in V_0 with no error.

- The support of $\phi(t)$ is contained in $[0, N)$, and $\phi(t)$ is orthogonal to all integer translates of itself. The wavelets $\{w^{mn}\}$ form an ON basis for L^2.

- B-splines fit into this multiresolution framework, though more naturally with biorthogonal wavelets.

- There are matrices

  $$\mathbf{T} = (\downarrow 2)2\mathbf{H}\mathbf{H}^{\rm tr}.$$

  associated with each of the Daubechies solutions whose eigenvalue struture determines the convergence properties of the wavelet expansions. These matrices have beautiful eigenvalue structures.

- There is a smoothness theory for Daubechies D_M. Recall $M = N+1 = 2p$. The smoothness grows with p. For $p = 1$ (Haar) the scaling function is piecewise continuous. For $p = 2$, (D_4) the scaling function is continuous but not differentiable. For $p \geq 3$ we have $s = 1$ (one derivative). For $p = 5, 6, 7, 8$ we have $s = 2$. For $p = 9, 10$ we have $s = 3$. Asymptotically s grows as $0.2075p + \text{ constant}$.

- The constants $\mathbf{c}$ are explicit for $N = 1, 3$. For $N = 5, 7, \cdots$ they must be computed numerically.

CONCLUSION: EXAMPLES AND DEMOS FROM THE WAVELET TOOLBOX OF MATLAB.