Wavelet Notes

W. N. Polyzou
The University of Iowa
Iowa City, IA, 52242

February 11, 2019

Abstract

Notes on using wavelets in scattering calculations

file: usr/wavelets/wavelets.tex

1 Fourier Analysis:

Consider a function $f(x)$, defined for x satisfying $0 \leq x \leq L$, of the form:

$$f(x) = \sum_{n=1}^{N} c_n \sin(\frac{n\pi x}{L}). \tag{1}$$

The coefficients c_n can be found in terms of $f(x)$ by computing the integral:

$$\int_0^L f(x) \sin(\frac{m\pi x}{L})dx = \int_0^L \sum_{n=1}^{N} c_n \sin(\frac{n\pi x}{L}) \sin(\frac{m\pi x}{L})dx. \tag{2}$$

Because the sum is finite, it is possible to change the order of the sum and the integral:

$$\int_0^L f(x)\sin(\frac{m\pi x}{L}) = \sum_{n=1}^{N} c_n \int_0^L \sin(\frac{n\pi x}{L})\sin(\frac{m\pi x}{L})dx. \tag{3}$$

The integral can be done analytically using the trigonometric identity:

$$\sin(a)\sin(b) = \frac{1}{2i}(e^{iax} - e^{-iax})\frac{1}{2i}(e^{ibx} - e^{-ibx}) =$$

$$-\frac{1}{4}(e^{i(a+b)x} + e^{-i(a+b)x} - e^{i(a-b)x} - e^{-i(a-b)x}) =$$

$$\frac{1}{2}\left(\cos((a-b)x) - \cos((a+b)x)\right). \tag{4}$$

Using this identity in the integrals gives

$$\int_0^L \sin(\frac{n\pi x}{L})\sin(\frac{m\pi x}{L})dx =$$

$$\int_0^L \frac{1}{2}\left(\cos(\frac{(m-n)\pi x}{L}) - \cos(\frac{(m+n)\pi x}{L})\right) dx$$

For the case $m \neq n$ this integral is

$$\frac{1}{2}\left(\frac{L}{(m-n)\pi}\sin(\frac{(m-n)\pi x}{L}) - \frac{L}{(m+n)\pi}\sin(\frac{(m+n)\pi x}{L})\right)|_0^L = 0$$

while for the case that $m = n$ the first term is non-zero and gives

$$\int_0^L \sin(\frac{n\pi x}{L})\sin(\frac{n\pi x}{L})dx = \frac{L}{2}.$$

This gives the orthogonality relation

$$\frac{2}{L}\int_0^L \sin(\frac{n\pi x}{L})\sin(\frac{m\pi x}{L})dx = \begin{cases} 1 & : \quad m = n \\ 0 & : \quad m \neq n \end{cases}$$

The integral becomes

$$c_m = \frac{2}{L}\int_0^L f(x)\sin(\frac{m\pi x}{L})dx.$$

If $f(x)$ is an arbitrary function we can define approximate functions $f_N(x)$ by

$$f_N(x) := \sum_{n=1}^{N} c_n \sin(\frac{n\pi x}{L})$$

where the coefficients c_n are given by:

$$c_n := \frac{2}{L}\int_0^L f(x)\sin(\frac{m\pi x}{L})dx.$$

It is natural to ask to what extent f_N approximates f

$$\lim_{N\to\infty}(f(x) - f_N(x)) \overset{?}{\to} 0$$

This is the basic problem of Fourier analysis. In order to answer this both the class of functions and nature of the convergence need to be defined.

It turns our that this works for a large class of functions. This will be discussed later. Let us define

$$\phi(x) = \sqrt{\frac{2}{L}}\sin(\frac{\pi x}{L}).$$

' The Fourier expansion has the form

$$f_N(x) := \sum_{n=1}^{N} c_n \phi(nx).$$

The normalization integral becomes

$$\int_0^L \phi(nx)\phi(mx)dx = \delta_{mn}$$

If we want to approximate a function on the interval $[nL, (n+1)L]$ we can use

$$f_N(x) := \sum_{m=1}^{N} c_m \phi(mx - nL).$$

where the $\phi_n(x)$ is defined to be 0 outside of the interval $[0, L]$. It follows that functions of the form $\phi(mx - nL)$ can be used to approximate a large class of functions.

These expressions look very much like the relations that express vectors in terms of orthogonal unit vectors. They are infinite dimensional versions of these vector relations.

Wavelets attempt to do a similar thing. The idea is to start with a "mother function" (like $\sqrt{\frac{2}{L}}\sin(\frac{\pi x}{L})$) and use scaling and translations to generate a basis.

note: The Fourier sine series can be thought of as defining a basis for odd functions on $[-L, L]$, while Fourier cosine series can be thought of as defining a basis for even functions on $[-L, L]$. Since any function on $[0, L]$ can be extended uniquely to an even or odd function of $[0, L]$, we can represent any function on $[0, L]$ by sines or cosines. What this means in practice, if the interval $[0, L]$ corresponds to half of a wavelength we can use sines or cosines. If the interval corresponds to a full wavelength we need both sines and cosines.

2 Periodic Functions:

The Fourier series provides a useful representation for functions that are periodic. A function $f(t)$ is *periodic* with period T if T is the smallest number for which $f(t+T) = f(t)$ for all t.

The functions $\sin(\frac{n2\pi t}{T})$, $\cos(\frac{n2\pi t}{T})$ are easily seen to satisfy this condition with period T.

One can show (see later) that (almost) any periodic function with period T can be approximated to arbitrary accuracy in the mean, $L^2[0, T]$, by a series of the form:

$$f_N(t) = c_0 + \sum_{n=1}^{N} c_n \cos(\frac{n2\pi t}{T}) + \sum_{n=1}^{N} d_n \sin(\frac{n2\pi t}{T}) \tag{5}$$

where

$$c_0 = \frac{1}{T}\int_0^T f(t)dt \tag{6}$$

$$c_n = \frac{2}{T}\int_0^T f(t)\cos(\frac{n2\pi t}{T})dt \tag{7}$$

$$d_n = \frac{2}{T}\int_0^T f(t)\sin(\frac{n2\pi t}{T})dt. \tag{8}$$

Note that, as was discussed in the first section, for functions defined on $[0, T/2]$ (half of the full wavelength) it is possible to use just sines or cosines.

The nature of the convergence is that if $f(x)$ is any (complex) function that satisfies

$$\|f\| = [\int_0^T f^*(t)f(t)dt]^{1/2} < \infty \tag{9}$$

then for every $.pdfilon > 0$ there is an $N.pdfilon$ such that for any $N > N_{.pdfilon}$

$$\int_0^T (f^*(t) - f_N^*(t))(f(t) - f_N(t))dt < .pdfilon^2 \tag{10}$$

This is called $L^2[0, T]$ or convergence or convergence in the mean.

If the function is continuous and periodic the convergence is uniform and pointwise. The class of functions satisfying $\|f\| < \infty$ is larger because it also includes discontinuous functions, and functions with mild singularities. For these functions the pointwise convergence is replaced by convergence in the mean.

Since many signals are periodic, the Fourier coefficients are useful parameterizations of the signal.

3 Complex Fourier Series:

The functions

$$e^{i\frac{2n\pi x}{L}} = \cos(\frac{2n\pi x}{L}) + i\sin(\frac{2n\pi x}{L}) \tag{11}$$

and

$$e^{-i\frac{2n\pi x}{L}} = \cos(\frac{2n\pi x}{L}) - i\sin(\frac{2n\pi x}{L}) \tag{12}$$

can be use to express $\cos(\frac{2n\pi x}{L})$ and $\sin(\frac{2n\pi x}{L})$ as linear combinations of $e^{-i\frac{n\pi x}{L}}$ and $e^{i\frac{n\pi x}{L}}$.

In this case the Fourier expansion can be written as

$$f(x) = f_r(x) + if_i(x) = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{2n\pi x}{L}}. \tag{13}$$

If we take the complex conjugate of this equation it becomes

$$f^*(x) = f_r(x) - if_i(x) = \sum_{n=-\infty}^{\infty} c_n^* e^{-i\frac{2n\pi x}{L}} =$$

$$\sum_{n=-\infty}^{\infty} c_{-n}^* e^{i\frac{2n\pi x}{L}}. \tag{14}$$

If we use

$$c_n = c_{nr} + ic_{ni} \qquad c_n^* = c_{nr} - ic_{ni} \tag{15}$$

we can separate the real and imaginary parts of these expressions:

$$f_r(x) = \sum_{n=-\infty}^{\infty} [\frac{(c_{nr} + c_{-nr})}{2} \cos(\frac{2n\pi x}{L}) - \frac{c_{ni} - c_{-ni}}{2} \sin(\frac{2n\pi x}{L})] \tag{16}$$

and

$$f_i(x) = \sum_{n=-\infty}^{\infty} [\frac{(c_{nr} - c_{-nr})}{2} \sin(\frac{2n\pi x}{L}) + \frac{c_{ni} + c_{-ni}}{2} \cos(\frac{2n\pi x}{L})] \tag{17}$$

Except for the special case of $n = 0$, the sums over negative and positive values of n give identical values which means that these equations can be written as

$$f_r(x) = c_{0r} + \sum_{n=1}^{\infty}[(c_{nr} + c_{-nr}) \cos(\frac{2n\pi x}{L}) - (c_{ni} - c_{-ni}) \sin(\frac{2n\pi x}{L})] \tag{18}$$

$$f_i(x) = c_{0i} + \sum_{n=1}^{\infty}[(c_{nr} - c_{-nr}) \sin(\frac{2n\pi x}{L}) + (c_{ni} + c_{-ni}) \cos(\frac{2n\pi x}{L})] \tag{19}$$

which is the usual form of the Fourier series in terms of trigonometric functions.

This shows that the complex form of the series has the same content as the real from. The Fourier coefficients in

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{2n\pi x}{L}} \tag{20}$$

can be evaluated

$$c_n = \frac{1}{L}\int_0^L e^{-i\frac{2n\pi x}{L}} f(x)dx \tag{21}$$

If the function $f(x)$ is real, then the coefficients satisfy $c_n^* = c_{-n}$, which is equivalent to setting $c_{nr} = c_{-nr}$ and $c_{ni} = -c_{-ni}$. These conditions also ensure that $f_i(x)$ in the expression above vanishes.

Note that we have freely changed the orders of sums and integrals. This is not generally justified for infinite sums. The proper treatment is us use finite sums and then show that the limit exists and converges to the desired function.

4 Fourier Transform:

The Fourier transform is a continuous version of the Fourier series. The Fourier transform of $f(x)$ is given by

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx \tag{22}$$

The functions $\tilde{f}(k)$ play the role of the coefficients c_n, and the integral over k replaces the sum over n.

The inverse Fourier transform is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \tilde{f}(k) dk \tag{23}$$

This follows from the key result:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dy e^{ik(x-y)} f(y) \tag{24}$$

for $f(x)$ piecewise continuous and absolutely integrable. The origin of the factor 2π is because

$$\int e^{i2\pi kx} dk = \frac{1}{2\pi} \int e^{ikx} dk. \tag{25}$$

In order to prove the key result first note that the second integral only makes sense after the first one is done. To interchange the order of integration we replace the k integral by the limit of integrals over a finite region:

$$f(x) = \lim_{\lambda\to\infty} \frac{1}{2\pi} \int_{-\lambda}^{\lambda} dk \int_{-\infty}^{\infty} dy e^{ik(x-y)} f(y) = \tag{26}$$

$$\lim_{\lambda\to\infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} dy \int_{-\lambda}^{\lambda} dk e^{ik(x-y)} f(y). \tag{27}$$

Because the integrand is absolutely integrable we can change the order of integration to get

$$\lim_{\lambda\to\infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} dy \frac{e^{i\lambda(x-y)} - e^{-i\lambda(x-y)}}{i(x-y)} f(y) = \tag{28}$$

$$\lim_{\lambda\to\infty} \frac{1}{\pi} \int_{-\infty}^{\infty} dy \frac{\sin(\lambda(x-y))}{(x-y)} f(y). \tag{29}$$

To motivate the proof let $t = x - y$:

$$\lim_{\lambda\to\infty} \frac{1}{\pi} \int_{-\infty}^{\infty} dt \frac{\sin(\lambda t)}{t} f(x-t) \tag{30}$$

and $s = \lambda t$

$$\lim_{\lambda\to\infty} \frac{1}{\pi} \int_{-\infty}^{\infty} ds \frac{\sin(s)}{s} f(x - \frac{s}{\lambda}) \tag{31}$$

If we could bring the limit inside of the integral and the integrand was continuous we would get the desired result (after doing the s integral). Since we cannot justify this interchange, and we want to allow for the possibility that the function might have some isolated discontinuities, we have to be more careful.

First separate this into two cases - one where s is positive and one where s is negative. We also undo the $s = \lambda t$ substitution to get:

$$\lim_{\lambda\to\infty} \frac{1}{\pi} [\int_{-\infty}^{0} dt \frac{\sin(\lambda t)}{t} f(x-t)) + \int_{0}^{\infty} dt \frac{\sin(\lambda t)}{t} f(x-t)]. \tag{32}$$

To evaluate this we treat each of the two terms separately. We will need the result

$$\frac{1}{\pi} \int_{-\infty}^{0} ds \frac{\sin(s)}{s} = \frac{1}{\pi} \int_{0}^{\infty} ds \frac{\sin(s)}{s} = \frac{1}{2} \tag{33}$$

which can be checked numerically or by using a tables of integrals.

We consider the second term in (**??**). The treatment of the first term in similar. The second term is

$$\lim_{\lambda\to\infty} \frac{1}{\pi} \int_{0}^{\infty} dt \frac{\sin(\lambda t)}{t} f(x-t).$$

We break the region of integration up into three parts:

$$\int_0^\infty = \int_0^\delta + \int_\delta^T + \int_T^\infty \tag{34}$$

We want to show that we can choose T, λ large enough, and δ small enough that this integral is within a prescribed $.pdfilon$ of the left hand limit of $f(x)/2 = f(x - 0^+)/2$. To this end we fix any $.pdfilon > 0$.

If $f(x)$ is absolutely integrable then

$$|\int_T^\infty dt \frac{\sin(\lambda t)}{t} f(x-t)]| \leq$$

$$\int_T^\infty dt |\frac{f(x-t)}{t}| < \infty$$

In order that this integral converge there must be a sufficiently large T such that

$$\int_T^\infty dt |\frac{f(x-t)}{t}| < \frac{.pdfilon}{3}$$

Note that this choice is independent of λ. This shows that it is possible to choose T sufficiently large so the contribution from $\int_T^\infty$ is less than $.pdfilon/3$

Next consider the choice of δ which is in the integral

$$\frac{1}{\pi} \int_0^\delta dt \frac{\sin(\lambda t)}{t} f(x-t). \tag{35}$$

Since $f(x)$ is assumed to have only isolated discontinuities we can assume the either $f(x-t)$ is continuous at $t=0$ or discontinuous at $t=0$, and continuous in $(0,\delta]$. I will not consider the most general case, I will assume that the derivative of $f(x-t)$ is also continuous on $(0,\delta]$ (i.e. except possibly at $t=0$). Then the mean value theorem implies

$$f(x-t) = f(x-0^+) + t g_x(t)$$

where $g_x(t)$ is continuous in t. With this replacement

$$\frac{1}{\pi} \int_0^\delta dt \frac{\sin(\lambda t)}{t} (f(x-0^+) + t g_x(t)) =$$

$$\frac{f(x-0^+)}{\pi} \int_0^\delta dt \frac{\sin(\lambda t)}{t} + \frac{1}{\pi} \int_0^\delta dt \sin(\lambda t) g_x(t)) \tag{36}$$

The second term is bounded by $\frac{g_{max}\delta}{\pi}$ where g_{max} is the maximum value of $g_x(t)$ for $0 \leq t \leq \delta$. We can choose δ so this is less than $.pdfilon/3$

Give δ and T the middle integral can be made as small as desired by integrating by parts and choosing λ large enough

$$|\frac{1}{\pi} \int_\delta^T dt \frac{\sin(\lambda t)}{t} f(x-t)| \leq \tag{37}$$

$$\frac{1}{\lambda\pi}|\frac{\cos(\lambda\delta)}{\delta}f(x-\delta)-\frac{\cos(\lambda T)}{T}f(x-T)-\frac{1}{\pi}\int_{\delta}^{T}dt\cos(\lambda t)(\frac{f(x-t)}{t^2}-\frac{f'(x-t)}{t})| \tag{38}$$

As long a f' is continuous the expression in the absolute value is bounded and independent of λ. The $1/\lambda$ term makes this expression vanish as $\lambda \to \infty$. It follows that λ can be chosen to be sufficiently large so this is less than $.pdfilon/3$

Putting everything together gives the following

$$|\frac{1}{2\pi}\int_{-\infty}^{\infty}dy\int_{-\lambda}^{\lambda}dke^{ik(x-y)}f(y)-\frac{(f(x-0^+)+f(x+0^+)}{\pi}\int_{0}^{\delta}dt\frac{\sin(\lambda t)}{t}| < 2.pdfilon \tag{39}$$

The factor of 2 comes from the $-\infty \to 0$ integral. Since the $.pdfilon$ can be made as small as desired and the estimate is independent of λ for sufficiently large λ, we can take the lim as $\lambda \to \infty$ which gives

$$\lim_{\lambda\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty}dy\int_{-\lambda}^{\lambda}dke^{ik(x-y)}f(y)=\frac{(f(x-0^+)+f(x+0^+)}{2} \tag{40}$$

We write the final result as

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}dk\int_{-\infty}^{\infty}dye^{ik(x-y)}f(y) \tag{41}$$

or

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}dk\int_{-\infty}^{\infty}dye^{ik(x-y)}=\delta(x-y) \tag{42}$$

where

$$f(x)\int_{-\infty}^{\infty}\delta(x-y)f(y)dy \tag{43}$$

with the understanding that this form is only valid at points of continuity.

This can be extended to square integrable functions, functions without local continuous derivatives (and more general classes of functions) by approximating them in the mean by functions that are also absolutely integrable. This is a technicality that is neither difficult nor important.

exercise 1: show by plotting that $|\sin(x)/x| \leq 1$

exercise 2: show (using mathcad) that $\int_0^{\lambda} dx \sin(x)/x$ approaches $\frac{\pi}{2}$ as λ gets large.

5 Spaces of Functions

It is claimed that the Fourier series can be used to approximate a large class of functions defined on the interval $[0, L]$. This raises a number of questions such as (a) what is the class of functions and (b) what is meant by the term "approximate"?

There are many possible answers to these two questions. In quantum mechanics the appropriate class of functions are complex functions $f(x) = f_r(x) + if_i(x)$ that are square integrable in the sense:

$$\|f\|^2 = \int_0^L f(x)^* f(x) dx = \int_0^L (f_r^2(x) + f_i^2(x)) dx < \infty \tag{44}$$

The answer to what do we mean by approximate is that if $f_N(x)$ is the partial series constructed from the first $2N + 1$ terms in Fourier series:

$$f_N(x) = \sum_{n=-N}^{N} c_n e^{i\frac{2n\pi x}{L}} \tag{45}$$

with

$$c_n = \frac{1}{L} \int_0^L e^{-i\frac{2n\pi x}{L}} f(x) dx \tag{46}$$

then we require that for every $.pdfilon > 0$ there is an $N.pdfilon$ such that for any $N > N_{.pdfilon}$ that

$$\|f - f_N\| < .pdfilon. \tag{47}$$

This means that if we can always choose N large enough so that the integral of the square of the difference between the exact and approximate functions are small.

Note that while this kind of convergence is physically relevant in quantum mechanics, it does not always mean that the difference has to be small at every point. To see this consider a function that has value $1/.pdfilon$ if x is within $\pm .pdfilon^2/2$ of $L/2$ and zero otherwise. The integral of this function is $.pdfilon$ which vanishes as $.pdfilon \to 0$, but at the point $x = L/2$, the value of the function is $1/.pdfilon \to \infty$ as $.pdfilon \to 0$. If $f - f_N$ was this function with $.pdfilon = 1/N$, then $f_N \to f$ in the above sense, but the difference at the point $x = L/2$ is arbitrarily large.

This same point is illustrated with the Gibbs phenomena, the difference between the Fourier approximations and the exact function leads to large oscillations near points where the function has discontinuous derivatives.

The simplest proof of the above result uses a classical result about polynomials due to Weierstrass. The Weierstrass theorem states that any continuous function of $[0, L]$ can be uniformly approximated by a finite degree polynomial, $p_n(x)$. What is essential is that $L < \infty$ and the end points are points of continutiy. In this case the both the class of functions (continuous) and the convergence (uniform) is different. Uniform convergence simply means that for every $.pdfilon > 0$ there is an $N.pdfilon$ such that for any $n > N_{.pdfilon}$

$$|f(x) - p_n(x)| < .pdfilon \tag{48}$$

independent of $x \in [0, L]$. Here the class of functions is smaller but the convergence is stronger. This can be used to eventually derive the result for the $L^2[0, L]$ convergence of the Fourier series to any square integrable function.

The idea of the proof of the Weierstrass theorem is based on some key elements:
1.) First, using an additional polynomial if necessary, extend the function to a continuous function that vanishes outside of a fixed bounded interval.
2.) Pick a set of functions $p_n(t)$ that are polynomials in some sufficiently large region, are continuous and identically zero outside of that region, and have two important properties. The first is that they always integrate to 1. The second is that they decrease strongly with increasing n except at the point $t = 0$. An example would be something like $c_n(\frac{L^2-t^2}{L^2})^n$, where the constant c_n is chosen so the integral is 1.
3.) The integral $\int f(s)p_n(x-s)ds$ defines functions that are polynomials on the appropriate region and are approximations to $f(s)$. The fact that $f(s)$ is bounded and continuous on the interval $[0, L]$ is essential to show there is a sufficiently small interval such that $f(s)$ can be approximated by a constant in any given interval of that size. This is enough to ensure that for sufficiently large N that this polynomial is close to the whole function.
4.) The next step is to convert the Fourier series to a series in two polynomial variables. If the function is not periodic, the discontinuities can be fixed with small square integrable corrections.

6 Wavelets

We begin by considering the continuous wavelet transform. This is like the Fourier transform. Both continuous and discrete wavelets are built from a single function called the **mother function**. We use the notation, $M(x)$, for the mother function

Before considering general properties of a mother function we define translations and scale transformations of $M(x)$

$$M_{t,s}(x) := |s|^{-p} M(\frac{x-t}{s}). \tag{49}$$

The factor p is at our disposal. These functions are the wavelets associated with the mother function $M(x)$. We investigate conditions on the mother function that allow one to expand any function in terms of wavelets.

We can choose p to satisfy certain properties. Note that

$$\int_{-\infty}^{\infty} \left| |s|^{-p} M(\frac{x-t}{s}) \right|^q dx =$$

$$|s|^{1-qp} \int_{-\infty}^{\infty} |M(u)|^q du \tag{50}$$

It follows that if $p = 1/q$ the q-norm

$$\|M\|_q := \left(\int_{-\infty}^{\infty} |M(u)|^q du \right)^{1/q} \tag{51}$$

is preserved under scale transformations. This is a nice feature, but it not of any fundamental importance. With $p = 1/q$:

$$\|M\|_q = \|M_{t,s}\|_q \qquad \text{for all} \qquad s, t \tag{52}$$

The **continuous wavelet transform** of f is defined by the formula:

$$\hat{f}(s,t) := \int_{-\infty}^{\infty} M^*_{s,t}(x) f(x) dx = (M_{s,t}, f) \tag{53}$$

where $*$ indicates the complex conjugate in the case that the mother function is complex. In what follows a $\hat{}$ is used to indicate the wavelet transform of a function.

We investigate the condition on the mother function so this expression can be inverted to express $f(x)$ in terms of $\hat{f}(s,t)$.

Parseval's identity for the Fourier transform ensures that the wavelet transform can be expressed in terms of the function and the mother function or in terms of their Fourier transforms:

$$\hat{f}(s,t) = (M_{s,t}, f) = (\tilde{M}_{s,t}, \tilde{f}) \tag{54}$$

where the $\sim$ indicates the Fourier transform defined by:

$$\tilde{M}_{s,t}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} M_{s,t}(x) dx \tag{55}$$

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx. \tag{56}$$

Note that Parseval's identity states $(f,f) = (\tilde{f}, \tilde{f})$, however using this with $f = g + h$ and $f = g + ih$ gives

$$(g,g) + (h,h) + (g,h) + (h,g) = (\tilde{g},\tilde{g}) + (\tilde{h},\tilde{h}) + (\tilde{g},\tilde{h}) + (\tilde{h},\tilde{g}) \tag{57}$$

and

$$(g,g) + (h,h) + i(g,h) - i(h,g) = (\tilde{g},\tilde{g}) + (\tilde{h},\tilde{h}) + i(\tilde{g},\tilde{h}) - i(\tilde{h},\tilde{g}) \tag{58}$$

which, using the identities $(g,g) = (\tilde{g},\tilde{g})$ and $(h,h) = (\tilde{h},\tilde{h})$, gives the solution to (**??**) and (**??**):

$$(g,h) = (\tilde{g},\tilde{h}) \tag{59}$$

which is the form of Parseval's identity that was used in (**??**).

The Fourier transform of $M_{s,t}(x)$ can be expressed in terms of the Fourier transform of the mother function:

$$\tilde{M}_{s,t}(k) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} |s|^{-p} M(\frac{x-t}{s}) dx =$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-iksu} e^{-ikt} |s|^{-p+1} M(u) du =$$

$$|s|^{1-p} e^{-ikt} \tilde{M}(sk). \tag{60}$$

This is useful because the Fourier transform can be expressed in terms of the Fourier transform of the mother function.

The inner product of the Fourier transforms gives

$$\hat{f}(s,t) = (\tilde{M}_{s,t}, \tilde{f}) =$$

$$\int_{-\infty}^{\infty} \tilde{M}^*_{s,t}(k)\tilde{f}(k)dk$$

$$\int_{-\infty}^{\infty} |s|^{1-p} e^{ikt} \tilde{M}^*(sk)\tilde{f}(k)dk \tag{61}$$

Multiply both sides of (**??**) by $e^{-ik't}$ and integrate over t to get

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ik't}(\tilde{M}_{s,t}, \tilde{f})dt =$$

$$|s|^{1-p}\tilde{M}^*(sk')\tilde{f}(k'), \tag{62}$$

were we have used the representation of the delta function:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i(k'-k)t}dt = \delta(k'-k). \tag{63}$$

The right had side of (**??**) is multiplied by the Fourier transform of the original function. We can't divide by $\tilde{M}^*(sk')$ because it might be zero for some values of k'. Instead, the trick is to eliminate it using the variable s.

Multiply both sides of this equation by $\tilde{M}(sk')$ and a yet to be determined weight function $w(s)$ and integrate over s. This gives

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} w(s)ds \int_{-\infty}^{\infty} dte^{-ik't}\tilde{M}(sk')\hat{f}(s,t) =$$

$$\tilde{f}(k')\int_{-\infty}^{\infty} w(s)ds|s|^{1-p}\tilde{M}^*(sk')\tilde{M}(sk') = \tilde{f}(k')Y(k') \tag{64}$$

where

$$Y(k') = \int_{-\infty}^{\infty} dsw(s)|s|^{1-p}|\tilde{M}(sk')|^2. \tag{65}$$

In order to be able to extract the Fourier transform of the original function, it is sufficient that $Y(k')$ satisfies $0 < A \leq Y(k') \leq B < \infty$ for some numbers A and B. In this case

$$\tilde{f}(k) = \frac{1}{2\pi Y(k)} \int_0^\infty w(s)ds \int_{-\infty}^\infty dte^{-ikt} \tilde{M}(sk)\hat{f}(s,t). \tag{66}$$

It is convenient to rewrite this in terms of the wavelet basis:

$$\tilde{f}(k) = \frac{1}{2\pi Y(k)} \int_{-\infty}^\infty w(s)|s|^{p-1}ds \int_{-\infty}^\infty dt\tilde{M}_{s,t}(k)\hat{f}(s,t). \tag{67}$$

We define the **dual wavelet** by

$$\tilde{M}^{s,t}(k) = \frac{1}{2\pi Y(k)} \tilde{M}_{s,t}(k). \tag{68}$$

The dual wavelet is distinguished from the ordinary wavelet by having the parameters s, t appearing as superscripts rather than subscripts.

The inversion formula can be expressed in terms of the dual wavelet by

$$\tilde{f}(k) = \int_{-\infty}^\infty w(s)s^{p-1}ds \int_{-\infty}^\infty dt\tilde{M}^{s,t}(k)\hat{f}(s,t). \tag{69}$$

In order to recover the original function we take the inverse Fourier transform of this expressions:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty dke^{ikx}\tilde{f}(k) =$$

$$\int_{-\infty}^\infty w(s)|s|^{p-1}ds \int_{-\infty}^\infty dtM^{s,t}(x)\hat{f}(s,t) \tag{70}$$

where

$$M^{s,t}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty dke^{ikx}\tilde{M}^{s,t}(k). \tag{71}$$

In general this is a tedious procedure because the dual wavelet $M^{s,t}(x)$ must be computed using (**??**) and (**??**) for each value of s and t. If the dual wavelet also had a mother function, then it would only be necessary to Fourier transform the "dual mother" and then all of the other Fourier transforms could be expressed in terms of the transform of the dual mother.

The first step is to investigate the structure of the dual wavelets in x-space:

$$M^{s,t}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty dke^{ikx}\tilde{M}^{s,t}(k) =$$

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk e^{ikx}\frac{1}{2\pi Y(k)}|s|^{1-p}e^{-ikt}\tilde{M}(sk) =$$
$$M^{s,0}(x-t)$$

where
$$M^{s,0}(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk e^{ikx}\frac{1}{2\pi Y(k)}|s|^{1-p}\tilde{M}(sk).$$

This shows for a single scale the dual wavelet and its translation can be expressed in terms of a single function. This is not necessarily true for the dual wavelet and the scaled quantity.

$$M^{s,0}(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk e^{ikx}\frac{1}{2\pi Y(k)}|s|^{1-p}\tilde{M}(sk) =$$

$$M^{s,0}(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} du e^{iu\frac{x}{s}}\frac{1}{2\pi Y(u/s)}|s|^{-p}\tilde{M}(u).$$

This fails to be a rescaling of the single function due to the s dependence in the quantity $Y(u)$. It follows that *if we can find a weight function* $w(s)$ such that $Y(u/s) = Y(u)$ then the dual wavelet will satisfy

$$M^{s,0}(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} du e^{iu\frac{x}{s}}\frac{1}{2\pi Y(u/s)}|s|^{-p}\tilde{M}(u) = |s|^{-p}M^{1,0}(x/s). \quad (72)$$

Note that in this case $Y(u)$ is a constant which we denote by Y. The function $M^{1,0}(x)$ serves as the dual mother wavelet.

To show how to choose $w(s)$ note that

$$Y(sk) = \int_{-\infty}^{\infty} dt w(t)|t|^{1-p}|\tilde{M}(tsk)|^2.$$

Let $t' = st$ to get

$$Y(sk) = \int_{-\infty}^{\infty} dt w(t)|t|^{1-p}|\tilde{M}(tsk)|^2$$

$$|s|^{p-2}\int_{-\infty}^{\infty} dt' w(t'/s)|t|'^{1-p}|\tilde{M}(t'k)|^2.$$

This will equal $Y(k)$ provided

$$w(t') = |s|^{p-2}w(t'/s) \qquad w(s) = |s|^{p-2}w(1).$$

With this choice

$$Y(k) = w(1) \int_{-\infty}^{\infty} \frac{dt}{|t|} |\tilde{M}(t)|^2 = Y.$$

Assuming this choice of weight the admissibility condition becomes

$$0 < A \leq Y \leq B < \infty$$

Having computed the constant Y it is now possible to write down an explicit expression for the dual mother wavelet:

$$M^{s,0}(x-t) =$$

$$\frac{|s|^{-p}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} du e^{iu\frac{(x-t)}{s}} \frac{1}{2\pi Y} \tilde{M}(u)$$

Letting $k = u/s$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \frac{1}{2\pi Y^+} |s|^{1-p} e^{ik(x-t)} \tilde{M}(ks)$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \frac{1}{2\pi Y} e^{ikx} \tilde{M}_{s,t}(k).$$

This has the form

$$M^{s,t}(x) = \frac{1}{2\pi} \frac{1}{Y} M_{s,t}(x). \tag{73}$$

Thus the inversion procedure can be summarized by the formulas:

$$f(x) = \int_{-\infty}^{\infty} |s|^{2p-3} ds \int_{-\infty}^{\infty} dt M^{s,t}(x) \hat{f}(s,t) \tag{74}$$

$$Y = \int_{-\infty}^{\infty} \frac{dt}{|t|} |\tilde{M}(t)|^2 \tag{75}$$

$$M^{s,t}(x) = \frac{M_{s,t}(x)}{2\pi Y} \tag{76}$$

$$M_{s,t} = |s|^{-p} M(\frac{x-t}{s}). \tag{77}$$

The only condition on the mother function is $0 < Y < \infty$. This requires that the Fourier transform of the mother function vanishes at the origin.

If we use the representation for the wavelet transform we get a representation of a delta function:

$$\delta(x-y) =$$

$$\int_{-\infty}^{\infty} |s|^{2p-3} ds \int_{-\infty}^{\infty} dt M^{s,t}(x) M^*_{s,t}(y) =$$

$$\frac{1}{2\pi Y} \int_{-\infty}^{\infty} |s|^{2p-3} ds \int_{-\infty}^{\infty} dt M_{s,t}(x) M^*_{s,t}(y).$$

We can also use this to formulate a Parseval's identity for wavelets

$$(f, f) = \frac{1}{2\pi Y} \int_{-\infty}^{\infty} |s|^{2p-3} ds \int_{-\infty}^{\infty} dt |\hat{f}(s,t)|^2. \tag{78}$$

Consider the example of the **Mexican hat** wavelet. The mother function is

$$M(x) = \frac{1}{\sqrt{2\pi}} (x^2 - 1) e^{-x^2/2}$$

To work with the Mexican hat mother function it is useful to derive general properties of Gaussian integrals:

$$\int_{-\infty}^{\infty} e^{-ax^2+bx+c} dx =$$

$$\int_{-\infty}^{\infty} e^{-a(x-\frac{b}{2a})^2+\frac{b^2}{4a}+c} dx.$$

Change variables to $y = \sqrt{a}(x - \frac{b}{2a})$ to obtain:

$$\frac{e^{\frac{b^2}{4a}+c}}{\sqrt{a}} \int_{-\infty}^{\infty} e^{-y^2} dy =$$

$$\sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a}+c}.$$

This can be used to compute the Fourier transform of the Mexican hat mother function:

$$\tilde{M}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} M(x) dx =$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}(x^2-1)e^{-x^2/2-ikx}dx$$

To do the integral insert a parameter a which will be set to 1 at the end of the calculation:

$$(-2\frac{d}{da}-1)\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-x^2a/2-ikx}dx =$$

$$(-2\frac{d}{da}-1)\frac{1}{2\pi}\sqrt{\frac{2\pi}{a}}e^{-\frac{k^2}{2a}} =$$

$$(\frac{1}{a}-\frac{k^2}{a^2}-1)\sqrt{\frac{1}{2\pi a}}e^{-\frac{k^2}{2a}}.$$

In the limit that $a \to 1$ this becomes

$$-\sqrt{\frac{1}{2\pi}}k^2e^{-\frac{k^2}{2}}.$$

Using this expression it is possible to calculate the coefficient Y

$$Y = \int_{-\infty}^{\infty}\frac{dk}{|k|}|\tilde{M}(k)|^2 =$$

$$\int_{-\infty}^{\infty}\frac{dk}{|k|}|\tilde{M}(k)|^2 =$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}|k|^3dke^{-k^2} =$$

$$\frac{1}{\pi}\int_{0}^{\infty}k^3dke^{-k^2}.$$

Inserting a parameter a which will eventually be set to 1 gives

$$\frac{1}{2\pi}(-\frac{d}{da})\int_{0}^{\infty}2kdke^{-ak^2} =$$

$$\frac{1}{2\pi}(-\frac{d}{da})\frac{1}{a}\int_{0}^{\infty}dve^{-v} =$$

$$\frac{1}{2\pi}.$$

This satisfies the essential inequality $0 < Y < \infty$ which ensures the admissibility of the Mexican hat mother function.

The expression for the wavelet transform and its inverse can be written as:

$$\hat{f}(s,t) = |s|^{-p} \int_{-\infty}^{\infty} dx \frac{1}{\sqrt{2\pi}} ((\frac{x-t}{s})^2 - 1) e^{-(\frac{x-t}{s})^2/2} f(x) =$$

$$|s|^{1-p} \int_{-\infty}^{\infty} du \frac{1}{\sqrt{2\pi}} (u^2 - 1) e^{-u^2/2} f(su+t).$$

where $x = su + t$

The inverse is formally given by

$$f(x) = \int_{-\infty}^{\infty} |s|^{2p-3} ds \int_{-\infty}^{\infty} dt \frac{M_{st}(x)}{2\pi Y} \hat{f}(s,t) =$$

$$\int_{-\infty}^{\infty} |s|^{2p-3} ds \int_{-\infty}^{\infty} dt \frac{1}{\sqrt{2\pi}} |s|^{-p} ((\frac{x-t}{s})^2 - 1) e^{-(\frac{x-t}{s})^2/2} \hat{f}(s,t)$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |s|^{p-3} ds \int_{-\infty}^{\infty} dt ((\frac{x-t}{s})^2 - 1) e^{-(\frac{x-t}{s})^2/2} \hat{f}(s,t)$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |s|^{p-3} ds \int_{-\infty}^{\infty} du (u^2 - 1) e^{-u^2/2} \hat{f}(s, su+x)$$

where $t = su + x$.

7 Scaling Functions and Wavelets

The concept of scaling functions is most easily understood using Haar wavelets (these are the ones made out of simple box functions).

In this example the scaling function is given by

$$\phi(x) = \begin{cases} 0 & x \leq 0 \\ 1 & 0 < x \leq 1 \\ 0 & x > 1 \end{cases} . \tag{79}$$

Note that this is normalized so

$$(\phi, \phi) := \int_{-\infty}^{\infty} \phi^*(x)\phi(x) dx = \int_0^1 \phi^*(x) dx = 1. \tag{80}$$

Define the unit translation operator T by

$$(T\psi)(x) = \psi(x-1). \tag{81}$$

This operator moves $\phi(x)$ to the right by 1 unit. The translation operator has the following properties:

$$(T\psi, T\chi) = \int_{-\infty}^{\infty} \psi^*(x-1)\chi(x-1)dx = \tag{82}$$

changing variables $y = x - 1$ gives:

$$\int_{-\infty}^{\infty} \psi^*(y)\chi(y)dy = (\psi, \chi) \tag{83}$$

or

$$(T\psi, T\chi) = (\psi, \chi). \tag{84}$$

If A is a linear operator its **adjoint** $A^\dagger$ is defined by the relation

$$(\psi, A^\dagger \chi) = (A\psi, \chi). \tag{85}$$

It follows that

$$(\psi, T^\dagger \chi) = (T\psi, \chi) = \int_{-\infty}^{\infty} \psi^*(x-1)\chi(x)dx. \tag{86}$$

Changing variables to $y = x - 1$ gives

$$(\psi, T^\dagger \chi) = \int_{-\infty}^{\infty} \psi^*(y)\chi(y+1)dy \tag{87}$$

or

$$(T^\dagger \psi)(x) = \psi(x+1). \tag{88}$$

Since

$$(\psi, \chi) = (T\psi, T\chi) = (\psi, T^\dagger T\chi) \tag{89}$$

we see that $T^\dagger = T^{-1}$. An operator whose adjoint is its inverse is called unitary. Unitary operators preserve inner products.

It follows from the definition of the scaling function $\phi(x)$ that

$$(T^m\phi, T^n\phi) = (\phi, T^{n-m}\phi) = \int_{-\infty}^{\infty} \phi^*(x)\phi(x-n+m)dx =$$

$$\int_0^1 \phi(x - n + m)dx = \delta_{nm} \tag{90}$$

This means the functions

$$\phi_m(x) := (T^n\phi)(x) = \phi(x - n) \tag{91}$$

are orthonormal. There are an infinite number of these functions as $n : -\infty \to \infty$ in integer s.pdf.

Define $\mathcal{V}_0$ to be the subspace of the space of square integrable functions of the form

$$f(x) = \sum_{n=-\infty}^{\infty} f_n\phi_n(x) = \sum_{n=-\infty}^{\infty} f_n(T^n\phi)(x) \tag{92}$$

where the square integrability requires that the coefficients satisfy

$$\sum_{n=-\infty}^{\infty} |f_n|^2 < \infty. \tag{93}$$

This is the space of square integrable functions that are piecewise constant on each unit interval. Note that while there are an infinite number of functions in $\mathcal{V}_0$, it is a small subspace of the space of square integrable functions.

In addition to translations, we define a linear operator D corresponding to scale transformations:

$$(D\psi)(x) = \frac{1}{\sqrt{2}}\psi(x/2). \tag{94}$$

When this is applied to the scaling function it gives

$$(D\phi)(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{\sqrt{2}} & 0 < x \leq 2 \\ 0 & x > 2 \end{cases} \tag{95}$$

This function has the basic box structure, except it is twice as wide as the original scaling function and shorter by a factor of $\sqrt{2}$. Note that the normalization ensures

$$(D\psi, D\chi) = \int_{-\infty}^{\infty} \frac{1}{2}\psi^*(x/2)\chi(x/2)dx = \tag{96}$$

setting $y = x/2$,

$$\int_{-\infty}^{\infty} \frac{2}{2} \psi^*(2)\chi(y)dy = (\psi, \chi) \tag{97}$$

or

$$(D\psi, D\chi) = (\psi, \chi). \tag{98}$$

To compute the adjoint of D note that

$$(\psi, D^\dagger \chi) = (D\psi, \chi) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2}} \psi^*(x/2)\chi(x)dx. \tag{99}$$

Setting $y = x/2$

$$\int_{-\infty}^{\infty} \psi^*(y)\sqrt{2}\chi(2y)dy \tag{100}$$

implies that

$$(D^\dagger \chi)(x) = \sqrt{2}\chi(2x). \tag{101}$$

This shows that $D^\dagger = D^{-1}$ with D also unitary.

Define the functions

$$\phi_{mn}(x) = (D^m \phi_n)(x) = 2^{-m/2}\phi(2^{-m}x - n) = 2^{-m/2}\phi(2^{-m}(x - 2^m n)). \tag{102}$$

It follows that for fixed m

$$(\phi_{mn}, \phi_{mk}) = (D^m \phi_n, D^m \phi_k) = (\phi_n, D^{m-m}\phi_k) = (\phi_n, \phi_k) = \delta_{nk} \tag{103}$$

This shows that the functions $\phi_{mn}(x)$ for fixed m are orthonormal. We define the subspace $\mathcal{V}_m$ of the square integrable functions to be those functions of the form:

$$f(x) = \sum_{n=-\infty}^{\infty} f_n \phi_{mn}(x) = \sum_{n=-\infty}^{\infty} f_n (D^m T^n \phi)(x) \tag{104}$$

where the square integrability requires that the coefficients satisfy

$$\sum_{n=-\infty}^{\infty} |f_n|^2 < \infty. \tag{105}$$

Normally the scaling function is the solution of a scaling equation. The scaling equation relates the $(D\phi)(x)$ to translates of the original scaling function. The general form of this relation is

$$(D\phi)(x) = \sum h_l T^l \phi(x) \tag{106}$$

where h_l are fixed constants.

In general this equation cannot be solved analytically. In the Haar case we can write down the solution realizing that the scaled box is stretched over two adjacent boxes with a suitable reduction in height:

$$D\phi(x) = \frac{1}{\sqrt{2}}\phi(x/2) = \frac{1}{\sqrt{2}}\phi(x) + \frac{1}{\sqrt{2}}T\phi(x) =$$

$$\frac{1}{\sqrt{2}}\phi(x) + \frac{1}{\sqrt{2}}\phi(x-1). \tag{107}$$

Here $h_0 = h_1 = 1/\sqrt{2}$. This equation is special to the Haar wavelets. The best way to think of this is that the scaling function $\phi(x)$ is the solution of the scaling equation up to normalization.

To proceed a few additional relations are useful. First note

$$DT\psi(x) = D\psi(x-1) = \frac{1}{\sqrt{2}}\psi(x/2-1) = \frac{1}{\sqrt{2}}\psi(\frac{x-2}{2}) = T^2 D\psi(x) \tag{108}$$

which leads to the operator relation

$$DT = T^2 D. \tag{109}$$

It follows from this equation that

$$D\phi_n(x) = DT^n\phi(x) = T^{2n}D\phi(x) = T^{2n}(h_0\phi(x) + h_1 T\phi(x)) \tag{110}$$

This shows that all of the basis elements in $\mathcal{V}_1$ can be expressed in terms of basis elements in $\mathcal{V}_0$.

Specifically if $\psi(x) \in \mathcal{V}_1$ then

$$\psi(x) = \sum_{-\infty}^{\infty} \psi_n \phi_{1n}(x) = \tag{111}$$

$$\sum_{-\infty}^{\infty} \psi_n D\phi_n(x) = \sum_{-\infty}^{\infty} (\psi_n h_0 \phi_{2n}(x) + \psi_n h_1 \phi_{2n+1}(x)) = \tag{112}$$

$$\sum_{-\infty}^{\infty} c_n \phi_n(x) \tag{113}$$

where

$$c_{2n} = \psi_n h_0 \qquad c_{2n+1} = \psi_n h_1. \tag{114}$$

It is easy to show that

$$\sum_{-\infty}^{\infty} |c_n|^2 = \sum_{-\infty}^{\infty} |\psi_n|^2 \tag{115}$$

What we have shown, as a consequence of the scaling equation, is the inclusion

$$\mathcal{V}_0 \supset \mathcal{V}_1. \tag{116}$$

Similarly, it is not difficult to show the inclusions

$$\cdots \mathcal{V}_{-k} \supset \mathcal{V}_{-k+1} \supset \cdots \supset \mathcal{V}_0 \supset \cdots \mathcal{V}_k \supset \mathcal{V}_k \cdots \tag{117}$$

In this example there are all spaces of piecewise constant, square integrable functions that are constant on segments that differ by powers of 2.

First note that as $k \to -\infty$ the approximation to $f(x)$ given by

$$f_k(x) = \sum_{n=-\infty}^{\infty} f_{kn} \phi_{kn}(x) \tag{118}$$

with

$$f_{kn} = \int_{-\infty}^{\infty} \phi^*_{kn}(x) f(x) dx \tag{119}$$

is bounded by the upper and lower Riemann sums for s.pdf of width 2^{-k}. (to deal with the infinite interval it is best to first consider functions that vanish outside of finite intervals and take limits). Since the upper and lower Riemann sums converge to the same integral (when the function is integrable) it follows that

$$\int_{-\infty}^{\infty} |f_k(x) - f(x)|^2 dx < .pdfilon \tag{120}$$

for sufficiently large $-k$.

Similarly, as $k \to +\infty$ the width of $\phi_{kn}(x)$ grows like 2^k while the height falls off like $2^{-k/2}$. Again, if the function vanishes outside of a bounded interval then for sufficiently large k there is only one (or two) $\phi_{kn}(x)$ that are non-vanishing where the function is non-vanishing. It follows that

$$f_k(x) \sim 2^{-k/2} \phi_{kn_0}(x) \int_{-\infty}^{\infty} f(x) dx \tag{121}$$

The integral of the square of this function $\sim 2^{-k} \to 0$ as $k \to \infty$.

Note that

$$\int_{-\infty}^{\infty} f_k(x)dx \to \int_{-\infty}^{\infty} f(x)dx \tag{122}$$

as $k \to \infty$. This shows that this integral is finite in L^1 but 0 in L^2. Thus the nature of the convergence is important. It also explains the dilemma with expanding functions with non zero mean in terms of functions with 0 mean.

If we define the projection operators

$$P_k f(x) = \sum_{n=-\infty}^{\infty} f_{kn}\phi_{kn}(x) \tag{123}$$

where

$$f_{kn} = \int_{-\infty}^{\infty} \phi^*_{kn}(x)f(x)dx. \tag{124}$$

The above conditions can be stated in terms of these projectors:

$$\lim_{k\to-\infty} P_k = I \tag{125}$$

$$\lim_{k\to+\infty} P_k = 0 \tag{126}$$

We are now ready to construct wavelets. First recall the condition

$$\mathcal{V}_0 \supset \mathcal{V}_1 \tag{127}$$

Let $\mathcal{W}_1$ be the space of vectors in the space $\mathcal{V}_0$ that are orthogonal to the vectors in $\mathcal{V}_1$. We can write

$$\mathcal{V}_0 = \mathcal{V}_1 \oplus \mathcal{W}_1 \tag{128}$$

This notation means that any vector in $\mathcal{V}_0$ can be expressed as a sum of two vectors - one that is in $\mathcal{V}_1$ and one that is orthogonal to every vector in $\mathcal{V}_1$.

Note that the scaling equation implies that every vector in $\mathcal{V}_1$ can be expressed as a linear combination of vectors in $\mathcal{V}_0$ using

$$D\phi_n(x) = h_0\phi_{2n}(x) + h_1\phi_{2n+1}(x) \tag{129}$$

Clearly the functions that are orthogonal to these in $\mathcal{V}_1$ on the same interval can be expressed in terms of the functions

$$M_{1n}(x) := DM_n(x) = h_0\phi_{2n}(x) - h_1\phi_{2n+1}(x) = \frac{1}{\sqrt{2}}(\phi_{2n}(x) - \phi_{2n+1}(x)) \tag{130}$$

These are also the elements of $\mathcal{V}_0$ that satisfy

$$(DM_{1n}, D\phi_l) = 0. \tag{131}$$

Thus we have that $\mathcal{W}_1$ is that space of square integrable functions of the form

$$f(x) = \sum_{n=-\infty}^{\infty} f_n M_{1n}(x) \tag{132}$$

with

$$f(x) = \sum_{n=-\infty}^{\infty} |f_n|^2 \tag{133}$$

where we have used

$$(M_{1n}, M_{1k}) = \delta_{nk} \tag{134}$$

which follows from the definitions.

Similarly we can express $\mathcal{V}_l = \mathcal{V}_{l+1} \oplus \mathcal{W}_{l+1}$ for all values of l. For the special case $l = -1$ we define the mother wavelet as

$$M(x) = D^{-1}(h_0\phi(x) - h_1 T\phi(x)) = \tag{135}$$

$$h_0\sqrt{2}\phi(2t) - h_1\sqrt{2}\phi(2(t-1) = (\phi(2t) - \phi(2(t-1)) \tag{136}$$

which is manifestly orthogonal to the scaling function. Translates of this function define a basis for $\mathcal{W}_0$

$$M_n(x) = T^n M(x) = T^n D^{-1}(h_0\phi(x) - h_1 T\phi(x)) = \tag{137}$$

$$D^{-1}(h_0\phi_{2n}(x) - h_1\phi_{2n+1}(x)) \tag{138}$$

It is simple to show that

$$(M_n, M_k) = \delta_{nk} \tag{139}$$

If we decompose every space we have for any k

$$\mathcal{V}_{-k} = \mathcal{W}_{-k+1} \oplus \mathcal{V}_{-k+1} = \tag{140}$$

$$\mathcal{W}_{-k+1} \oplus \mathcal{W}_{-k+2} \oplus \mathcal{V}_{-k+2} = \tag{141}$$

$$\mathcal{W}_{-k+1} \oplus \mathcal{W}_{-k+2} \oplus \cdots \oplus \mathcal{W}_l \oplus \mathcal{V}_l \tag{142}$$

Note that unlike the $\mathcal{V}$ spaces, the $\mathcal{W}_k$ spaces are all mutually orthogonal, since if $k > l \rightarrow \mathcal{W}_k \subset \mathcal{V}_l$ which is orthogonal to $\mathcal{W}_l$ by definition.

If $f(x)$ is any square integrable function the conditions

$$\lim_{k\to-\infty} P_k = I \tag{143}$$

$$\lim_{k\to+\infty} P_k = 0 \tag{144}$$

mean that for sufficiently large k and l that $f(x)$ can be well approximated by a function in

$$\mathcal{W}_{-k+1} \oplus \mathcal{W}_{-k+2} \oplus \cdots \oplus \mathcal{W}_l \tag{145}$$

This means that the function can be approximated by a linear combination of basis functions (wavelets) from each of the spaces $\mathcal{W}_r$.

Basis functions for $\mathcal{W}_m$ are given by

$$M_{mn}(x) = D^m T^n M(x) = D^{m-1}(h_0\phi_{2n}(x) - h_1\phi_{2n+1}(x)) \tag{146}$$

That these are a basis with the required properties is easily shown by showing that these functions are orthogonal to $\mathcal{V}_m$ and can be used to recover the remaining vectors in $\mathcal{V}_{m-1}$.

The functions $M_{nl}(x)$ satisfy

$$(M_{nl}, M_{n'l'}) = \delta_{nn'}\delta_{ll'} \tag{147}$$

where the $\delta_{nn'}$ follows from the orthogonality of the spaces $\mathcal{W}_n$ and $\mathcal{W}_{n'}$ for $n \neq n'$.

The $\delta_{ll'}$ follows from the unitarity of D and

$$(M, T^n M) = \delta_{n0}. \tag{148}$$

To summarize the important s.pdf one starts with a scaling equation of the form:

$$D\phi(x) = \sum h_l T^l \phi(x) \tag{149}$$

where one is normally given only the coefficients h_l. This equation is solved to find the scaling function $\phi(x)$. This, along with translations and dilitations is used to construct the spaces $\mathcal{V}_l$. The scaling equation ensures the existence of space $\mathcal{W}_k$ that can be used to build discrete orthonormal basis. The mother wavelet is expressed in terms of the scaling function and the coefficients as

$$M(x) = D^{-1}(h_0\phi(x) - h_1 T\phi(x)) \tag{150}$$

which is more complicated for a general scaling equation.

In general the coefficients h_l must satisfy constraints for the solution to the scaling equation to exist.

8 Scaling Functions II

The scaling function is the solution of the scaling equation

$$D\phi(x) = \sum_l h_l T^l \phi(x). \tag{151}$$

Using the definitions of the operators D and T this equation becomes:

$$\frac{1}{\sqrt{2}}\phi(\frac{x}{2}) = \sum h_l \phi(x-l). \tag{152}$$

It can be put in the form

$$\phi(x) = \sum_l \sqrt{2} h_l \phi(2x-l). \tag{153}$$

The sums are assumed to be from $-\infty \to \infty$. Finite sums are treated by assuming that a finite number of the h_l's are non zero.

The claim is that if this equation has a solution, it is unique up to an overall normalization factor. To investigate this claim take the Fourier transform of both sides of equation (**??**) to get

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-ikx}\phi(x)dx = \sum_l \sqrt{2} h_l \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-ikx}\phi(2x-l)dx. \tag{154}$$

Changing variables $x \to 2x - l$ gives

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-ikx}\phi(x)dx = \sum_l \frac{1}{\sqrt{2}} h_l \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i(k/2)(x+l)}\phi(x)dx. \tag{155}$$

or

$$\tilde{\phi}(k) = \tilde{\phi}(\frac{k}{2})\tilde{h}(\frac{k}{2}) \tag{156}$$

where

$$\tilde{h}(k) = \sum_l \frac{h_l}{\sqrt{2}} e^{-ikl}. \tag{157}$$

This form of the scaling equation can be iterated n times to get:

$$\tilde{\phi}(k) = \tilde{\phi}(\frac{k}{2^n}) \prod_{m=1}^{n} \tilde{h}(\frac{k}{2^m}) \tag{158}$$

This equation holds for any n. For a finite n, an approximation can be made by a finite number of iterations of the form

$$\tilde{\phi}_n(k) = \tilde{\phi}_{n-1}(\frac{k}{2})\tilde{h}(\frac{k}{2}) \tag{159}$$

for any starting function $\tilde{\phi}_0(k)$. In the limit of large n the function $\tilde{\phi}_m(k)$ should converge to the scaling function. It is clear that this result is approximately independent of the choice of starting function; it really depends on the coefficients h_l.

If the limit exists as $n \to \infty$ and the scaling function is continuous in a neighborhood of zero then

$$\tilde{\phi}(k) = \lim_{n\to\infty} \tilde{\phi}(\frac{k}{2^n}) \prod_{l=1}^{n} \tilde{h}(\frac{k}{2^l}) =$$

$$\tilde{\phi}(0) \prod_{l=1}^{\infty} \tilde{h}(\frac{k}{2^l}) \tag{160}$$

If the infinite product converges, then we have an expression for the scaling function, up to normalization, which is fixed by assigning a value to $\tilde{\phi}(0)$.

The coefficients h_l are not arbitrary. First note that setting $k = 0$ gives

$$1 = \prod_{l=0}^{\infty} \tilde{h}(0) \tag{161}$$

or

$$\tilde{h}(0) = 1 = \sum_l \frac{h_l}{\sqrt{2}} \tag{162}$$

or

$$\sum_l h_l = \sqrt{2}. \tag{163}$$

This condition is clearly satisfied by the Haar wavelets. This is a necessary condition on the scaling coefficients in order to have a solution to the scaling equation.

Another condition is that is needed to make a multiresolution analysis is the orthogonality of the unit translates, $(\phi_n, \phi_m) = \delta_{nm}$. This requires

$$2 \sum_{lk} h_l^* h_k \int_{-\infty}^{\infty} \phi^*(2x - 2n - l)\phi(2x - 2m - k)dx =$$

$$2\sum_{lk} h_l^* h_k \int_{-\infty}^{\infty} \phi^*(2x)\phi(2x - 2(m-n) - (k-l))dx =$$

$$\sum_{lk} h_l^* h_k \int_{-\infty}^{\infty} \phi^*(x)\phi(x - 2(m-n) - (k-l))dx$$

$$\sum_l h_l^* h_{l-2(m-n)} = \delta_{mn} \tag{164}$$

or equivalently

$$\sum_l h_{l-2m}^* h_l = \delta_{m0}. \tag{165}$$

This is trivially satisfied for the Haar wavelets. Note that since

$$\sum_m h_{n+2m} = \sum_m h_{n+2m+2k} \tag{166}$$

the m sum has two values according to whether n is even or odd:

$$h_e := \sum_m h_{2m} \qquad h_o := \sum_m h_{2m+1}. \tag{167}$$

This means that

$$\sum_{m,n} h_n^* h_{n+2m} = \sum_m \delta_{m0} = 1 =$$

$$h_e \sum_n h_{2n}^* + h_o \sum_n h_{2n+1}^* = 1 \tag{168}$$

or

$$h_e^* h_e + h_o^* h_o = 1 \tag{169}$$

We also have

$$h_e + h_o = \sqrt{2} \tag{170}$$

Assuming that the coefficients h_l are real these can be solved to get

$$h_e = h_o = \frac{1}{\sqrt{2}}. \tag{171}$$

These condition are useful checks and uniquely determine the Haar coefficients; but the essential equations are (**??**) and (**??**).

The other conditions that require some study are the ones relating the scaling function to the mother function. The mother function satisfies

$$M(x) = \sum_n \sqrt{2} g_n \phi(2x - n). \tag{172}$$

This and all of its translates should be orthogonal to the scaling function. In terms of the coefficients:

$$(M_m, \phi) = \sum_{n,l} h_l g_n^* (\phi_{n-2m}, \phi_l)$$

$$\sum_{n,l} h_l g_n^* \delta_{n-2m,l}$$

$$\sum_n h_{n-2m} g_n^* = 0 \tag{173}$$

for all m. We also need orthonormality of the translated mother function

$$(M_m, M_n) = \sum_{l,k} g_l g_k^* (\phi_{l-2m}, \phi_{k-2n})$$

$$\sum_k g_{k-2(n-m)} g_k^* = \delta_{mn} \tag{174}$$

or equivalently

$$(M_m, M) = \sum_k g_{k+2m} g_k^* = \delta_{m0} \tag{175}$$

If we choose $g_k := (-1)^k h_{l-k}$ where l is odd it follows that

$$\sum_k 2 g_{k-2(n-m)} g_k^* = \sum_k 2(-)^{l-k+2(n-m)} h_{k-2(n-m)} (-)^k h_{l-k} =$$

$$\sum_k 2 h_{k-2(n-m)} h_k = \delta_{mn} \tag{176}$$

where we have let $l - k \to k$ in the last term. It also follows that

$$\sum_n 2 h_{n-2m} g_n^* = \sum_n 2 h_{n-2m} (-)^n h_{l-n}^* =$$

$$\sum_n 2 h_{l-n'} (-)^{l-n'-2m} h_{n'-2m}^* = (-)^l \sum_n 2 h_{l-n'} (-)^{n'} h_{n'-2m}^* \tag{177}$$

Since l is odd, the sum is equal to its negative which shows that it vanishes. The choice of l is arbitrary - it simply involves moving the origin of the mother. Since the mother is orthogonal to the translates of all of the father wavelets, it does not matter where the origin is placed.

This shows that the coefficients h_l, satisfying

$$\sum_l h_l = \sqrt{2}. \tag{178}$$

$$\sum_l h_{l-2m} h_l^* = \delta_{m0} \tag{179}$$

$$g_k := (-1)^k h_{l-k} \qquad l \qquad \text{odd} \tag{180}$$

gives a multiresolution analysis, scaling function, and a mother function using

$$\phi(x) = \frac{\tilde{\phi}(0)}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \prod_{l=1}^{\infty} \tilde{h}(\frac{k}{2^l}) dk =$$

$$\frac{\tilde{\phi}(0)}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \prod_{l=1}^{\infty} \sum_n \frac{h_n}{\sqrt{2}} e^{-ikn/2^l}$$

This is not a very useful representation for computation, however it indicates that if a scaling fucntion has a finite number of coefficients h_l that the scaling function has support on

$$[0, N(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} \cdots)] = [0, N]$$

where N is the number of scaling coefficients.

An alternative is to compute the scaling function exactly on a dense set of points. This construction also starts from the scaling equation:

$$\phi(x) = \sum_l \sqrt{2} h_l \phi(2x - l) \tag{181}$$

Let $x = n$ to get

$$\phi(n) = \sum_l \sqrt{2} h_l \phi(2n - l) \tag{182}$$

Let $k = 2n - l$ gives

$$\phi(n) = \sum_k \sqrt{2} h_{2n-k} \phi(k) \tag{183}$$

Which gives the eigenvalue equation

$$\phi(n) = \sum_m H_{nm} \phi(m) \tag{184}$$

where

$$H_{nm} = \sqrt{2} h_{2n-m} \tag{185}$$

Eigenvectors of this equation with eigenvalue 1 are solutions of the scaling function at integer points - up to normalization. Eigenvectors with eigenvalues other than 1 can be tossed out.

Rather than solve the eigenvalue problems, one of the equations can be replaced by the condition

$$\sum_n \phi(n) = 1 \tag{186}$$

which follows from the assumption that $\int M(x)dx = 0$. The support condition implies that only a finite number of the $\phi(n)$ are non-zero. Note that this condition can be imposed independent of the orthonormality condition.

For the case of the $N = 2$ Daubechies wavelets these equations are

$$\phi(0) = \sqrt{2} h_0 \phi(0)$$

$$\phi(1) = \sqrt{2}(h_0\phi(2) + h_1\phi(1) + h_2\phi(0))$$

$$\phi(2) = \sqrt{2}(h_1\phi(3) + h_2\phi(2) + h_3\phi(1))$$

$$\phi(3) = \sqrt{2} h_3 \phi(3)$$

$$1 = \phi(0) + \phi(1) + \phi(2) + \phi(3)$$

The first and forrth equation give $\phi(0) = \phi(3) = 0$. The second and third equations are eigenvalue equations

$$\begin{pmatrix} \phi(1) \\ \phi(2) \end{pmatrix} = \begin{pmatrix} \sqrt{2}h_1 & \sqrt{2}h_0 \\ \sqrt{2}h_2 & \sqrt{2}h_3 \end{pmatrix} \begin{pmatrix} \phi(1) \\ \phi(2) \end{pmatrix} \tag{187}$$

Instead of solving the eigenvalue problem for an eigenvector with eigenvalue 1, we use

$$\phi(1) + \phi(2) = 1 \tag{188}$$

with

$$\phi(1) = \sqrt{2}(h_0\phi(2) + h_1\phi(1))$$

to get

$$\phi(1) = \sqrt{2}(h_0(1 - \phi(1)) + h_1\phi(1))$$

which can be solved for

$$\phi(1) = \frac{\sqrt{2}h_0}{1+\sqrt{2}(h_0 - h_1)} \tag{189}$$

and

$$\phi(2) = \frac{1-\sqrt{2}h_1}{1+\sqrt{2}(h_0 - h_1)} \tag{190}$$

This gives exact values of the scaling function at integer points. In this case there are only two non zero terms.

Integer translates can be constructed using

$$\phi_m(n) = \phi(n-m) \tag{191}$$

The scaling equation gives

$$\phi(m/2) = \sum_l \sqrt{2}h_l\phi(m-l) = \sum_n \sqrt{2}h_{m-n}\phi(n). \tag{192}$$

Translates of the scaled scaling function are

$$\phi_n(m/2) = \phi(m/2-n) = \phi(\frac{m-2n}{2}) = \sum_k \sqrt{2}h_{m-2n-k}\phi(k) =$$

$$\sum_l \sqrt{2}h_l\phi_{2n-l}(m) \tag{193}$$

This procedure can be repeated inductively to obtain

$$\phi_l(\frac{k}{2^n}) = 2^{\frac{n}{2}} \sum_{m_1\cdots m_n} h_{m_1}\cdots h_{m_n}\phi_{2^n l-2^{n-1}m_1-2^{n-2}m_2\cdots-2m_{n-1}-m_n}(k) =$$

$$2^{\frac{n}{2}} \sum_{m_1\cdots m_n} h_{m_1}\cdots h_{m_n}\phi(k-2^n l-2^{n-1}m_1-2^{n-2}m_2\cdots-2m_{n-1}-m_n) \tag{194}$$

Note that in any of the sums the only non-zero contributions occur when the argument of $\phi(\cdot)$ is 1 or 2. This equation gives exact values of the scaling function at points $x = \frac{n}{2^k}$. These are dense and if the scaling function is continuous this method can be used to approximate the scaling function an any point.

This has the advantage that the function is computed exactly at many points - with iterative metods it is computed approximately at one point.

The general form of the equations is

$$\sum_{n=1}^{2N-1} \phi(n) = 1$$

$$\phi(0) = \sqrt{2}h_0\phi(0)$$
$$\phi(1) = \sqrt{2}(h_0\phi(2) + h_1\phi(1) + h_2\phi(0))$$
$$\phi(2) = \sqrt{2}(h_0\phi(4) + h_1\phi(3) + h_2\phi(2) + h_3\phi(1) + h_4\phi(0))$$
$$\vdots$$
$$\phi(2N-2) = \sqrt{2}(h_{2N-1}\phi(2N-3) + h_{2N-2}\phi(2N-2) + h_{2N-3}\phi(2N-1))$$
$$\phi(2N-1) = \sqrt{2}h_{2N-1}\phi(2N-1)$$

The middle equations are related to the pyramid algorithm.

9 Daubechies Wavelets

The Daubechies wavelets have two special properties. First is that there are a finite number of non-zero coefficients h_i. This gives them a compact support. The second feature is that the first N moments of the wavelets are zero.

The constraint on the moments has interesting consequences. First we note that

$$\int M(x)x^l = 0 \qquad l = 0 \cdots N-1. \tag{195}$$

From this we conclude

$$\int M_{0m}(x)x^l = \int M(x-m)x^l = \int M(y)(y+m)^l =$$

$$\sum_{k=0}^{l} \frac{l!}{k!(l-k)!} m^{l-k} \int M(x)x^k = 0. \tag{196}$$

Similarly

$$\int DM(x)x^l = \frac{1}{\sqrt{2}} \int M(x/2)x^l dx = 2^{l+1/2} \int M(y)y^l dy = 0. \tag{197}$$

It straight forward to proceed inductively to show that

$$\int M_{nk}(x)x^l = 0 \qquad l = 0 \cdots N-1. \tag{198}$$

This means that every element of the Daubechies wavelet basis is orthogonal to all polynomials of degree less than N.

If we consider instead the orthonormal basis consisting of

$$\{T^n\phi(x), D^m T^n M(x) : -m \geq 0\} \tag{199}$$

we have

$$\int \phi_m(x)x^l \neq 0 \qquad l = 0 \cdots N-1 \tag{200}$$

Although the polynomials are not square integrable; we can multiply a polynomial by a box function $b(x)$ which is 1 between x_- and x_+ and zero elsewhere. This product is square integrable and is equal to the polynomial on the interval $[x_-, x_+]$. It follows that

$$p(x)b(x) = \sum_{mn} c_{mn} M_{mn}(x) = \sum_{mn} d_n \phi_n(x) + \sum_n \sum_{m\leq 0} c_{mn} M_{mn}(x) \tag{201}$$

where

$$c_{mn} = \int_{x_-}^{x_+} M_{mn}(x)p(x)dx \tag{202}$$

$$d_n = \int_{x_-}^{x_+} \phi_n(x)p(x)dx \tag{203}$$

The moment condition means that the coefficients $c_{mn} = 0$ whenever the support of the wavelet is completely contained inside of the box. Thus in the first expression the non-zero coefficients arise from end point contributions and to many small contributions from wavelets with support that are much larger than the box.

In the second expression the wavelets with support larger than the box do not appear. The endpoint contributions only affect the answer within a distance equal to the support of the wavelet from the endpoints of the box. Inside this distance the only nonzero coefficient are due to the translates of the scaling functions. There are a finite number of these coefficients, and in this region they provide an exact representation of the polynomial. Specifically let

$$I(x) = b(x)p(x) - \sum_n d_n \phi_n(x) + \sum_n \sum_{m\leq 0} c_{mn} M_{mn}(x) \tag{204}$$

then we have

$$0 = \|I\|^2 = \int_{x_-}^{x_-+\Delta} I(x)^2 dx + \int_{x_+-\Delta}^{x_+} I(x)^2 dx +$$

$$\int_{x_-+\Delta}^{x_+-\Delta} |p(x) - \sum_n d_n \phi_n(x)|^2 dx. \tag{205}$$

Since all three terms are non-negative we conclude that

$$\int_{x_-+\Delta}^{x_+-\Delta} |p(x) - \sum_n d_n \phi_n(x)|^2 dx = 0. \tag{206}$$

Since Δ is fixed by the choice of the wavelet and $x_\pm$ is arbitrary we have

$$\int_a^b |p(x) - \sum_n d_n \phi_n(x)|^2 dx = 0 \tag{207}$$

for any interval $[a, b]$. Since $p(x)$ and $\phi(x)$ are continuous (we did not prove this for $\phi(x)$ - but that is the claim in the literature) and the sum of translates is finite it follows that

$$p(x) = \sum_n d_n \phi_n(x) \tag{208}$$

pointwise on every finite interval. This establishes the desired result.

Note that expansion in the wavelet basis gives all coefficients zero. This is not a contradiction because none of the polynomials are square integrable. This is reminiscent of the initial problem that we had in representing functions that did not average to zero. The key point is that once one puts a box around a function, wavelets with very large support (large m) lead to many small contributions.

What is interesting is that even though the Daubechies wavelets do not have order N smoothness, there are linear combinations of the scaling function that can exactly represent polynomials locally.

These properties are the key to the normalization coefficients for the Γ and the coefficients needed to compute the wavelets by the Strang method. The key formulas are

$$x^l = \sum_m c_{lm} \phi(x - m) \tag{209}$$

where

$$c_{lm} = \int \phi(x - m) x^l dx = \sum_{k=0}^{l} \frac{l!}{(l-k)!k!} m^{l-k} \int \phi(y) y^k dx =$$

$$\sum_{k=0}^{l} \frac{l!}{(l-k)!k!} m^{l-k} c_{k0}. \tag{210}$$

Starting with $l = 0$ these equations give

$$c_{0m} = c_{00} \tag{211}$$

$$c_{1m} = mc_{00} + c_{10} \tag{212}$$

This can be continued to express all of the translates in terms of c_{k0} for $k = 0, \cdots, l$.

Note that the normalization conditions gives $c_{00} = 1$

Next note that

$$1 = \sum_m c_{0m} \phi(x - m) \tag{213}$$

$$l! = \sum_m c_{lm} \frac{d^l}{dx^l} \phi(x - m). \tag{214}$$

Multiplying by $\phi(x)$ and integrating gives

$$c_{00} = \frac{1}{l!} \sum_m c_{lm} \int \phi(x) \frac{d^l}{dx^l} \phi(x - m). \tag{215}$$

10 Daubechies Scaling Coefficients

Begin by defining the two standard polynomials:

$$P(T) = \frac{1}{\sqrt{2}} \sum_{l=0}^{M} h_l T^l \tag{216}$$

and

$$Q(T) = \frac{1}{\sqrt{2}} \sum_{l=0}^{M} g_l T^l = \frac{1}{\sqrt{2}} \sum_{l=0}^{M} (-)^l h^*_{N-l} T^l \qquad N \qquad \text{odd} \tag{217}$$

First we argue that M must be odd. By contradiction assume that $M = 2K$ is even and $h_M, h_0 \neq 0$. The orthogonality condition requires

$$\sum_{l=0}^{2K} h^*_l h_{l+2K} = h^*_0 h_M = \delta_{K0} \tag{218}$$

This vanishes if $K \neq 0$, requiring either h_0 or $h_M = 0$, contradicting the assumption that M is even. It follows that M is odd or $M = 2N - 1$.

Inserting this in the expressions for the polynomials:

$$P(T) = \frac{1}{\sqrt{2}} \sum_{l=0}^{2N-1} h_l T^l \tag{219}$$

and

$$Q(T) = \frac{1}{\sqrt{2}} \sum_{l=0}^{2N-1} (-)^l h^*_{2N-1-l} T^l \tag{220}$$

It follows that if we let $m = 2N - 1 - l$

$$Q^*(-T) = \frac{1}{\sqrt{2}} \sum_{m=0}^{2N-1} h_m (T^*)^{2N-1-m} \tag{221}$$

This can be expressed as

$$Q^*(-T) = T^{2N-1} P((T^*)^{-1}) \tag{222}$$

Properties of these polynomials are used to determine the coefficients h_l. First note that for complex $z = e^{i\omega}$ on the unit circle, $z = (z^*)^{-1}$. Thus

$$Q^*(e^{i\omega}) = e^{i(2N-1)\omega} P(e^{i(\omega+\pi)}) \tag{223}$$

Note $P(1) = \sum_{l=0}^{2N-1} \frac{1}{\sqrt{2}} h_l = 1$. Next consider the orthonormality condition, for $z = e^{i2\pi\omega}$.

$$|P(z)|^2 + |P(-z)|^2 = \tag{224}$$

$$\frac{1}{2} \sum_{l,l'=0}^{2N-1} (h^*_l h_{l'} e^{i2\pi\omega(l'-l)} + (-)^{i\pi(l'-l)} h^*_l h_{l'} e^{i2\pi\omega(l'-l)}) = \tag{225}$$

$$\frac{1}{2} \sum_{l,l'=0}^{2N-1} h^*_l h_{l'} (e^{i2\pi\omega(l'-l)} + (-)^{i\pi(l'-l)} e^{i2\pi\omega(l'-l)}) = \tag{226}$$

$$\frac{1}{2} \sum_{l,l'=0}^{2N-1} h^*_l h_{l'} (e^{i\pi(l'-l)(2\omega)} + e^{i\pi(l'-l)(2\omega+1}) \tag{227}$$

shifting $l' - l = k$ gives

$$\frac{1}{2} \sum_{l}^{2N-1} \sum_{k=-l}^{2N-1-l} h^*_l h_{k+l} (e^{i\pi k(2\omega)} + e^{i\pi k 2\omega} e^{i\pi k}) \tag{228}$$

In this form it is manifestly obvious that the coefficient of the k =odd terms vanish. Thus let $k \to 2n$:

$$\sum_{l}^{2N-1} \sum_{n=-l/2}^{N-1/2-l/2} h_l^* h_{l+2n} e^{i4\pi n\omega)} \tag{229}$$

where the n sum is over successive integers between $-l/2$ and $N-1/2/-l/2$.

The condition that this is 1 for all ω gives

$$\sum_{l}^{2N-1} h_l^* h_{l+2n} = \quad constant \quad \times \delta_{n0} \tag{230}$$

or

$$\sum_{l}^{2N-1} h_l^* h_l = 1 \tag{231}$$

Thus we have

$$P(1) = 1; \qquad |P(1)|^2 + |P(-1)|^2 = 1 \tag{232}$$

Consistency of these two equations requires that $P(-1) = 0$,

The Daubechies wavelets have the property that $P(z)$ has a high order zero at $z = -1$:

$$P(z) = \left(\frac{1+z}{2} \right)^{N+1} W(z) \tag{233}$$

The normalization is chosen so the polynomial $W(z) = 1$. The problem is reduced to finding polynomials that have this property.

One technique for finding $W(z)$ utilizes trigonometric polynomials. The problem is to find polynomials $P(z)$ with the property that

$$|P(z)|^2 + |P(-z)|^2 = 1 \tag{234}$$

Let $z = e^{-i2\pi\omega}$ which gives

$$\frac{1+z}{2} = e^{-i\pi\omega} \cos(\pi\omega) \tag{235}$$

$$\frac{1-z}{2} = ie^{-i\pi\omega} \sin(\pi\omega) \tag{236}$$

In terms of these equations

$$\cos(\pi\omega) = \frac{z^{1/2} + z^{*1/2}}{2} \tag{237}$$

$$\sin(\pi\omega) = \frac{z^{*1/2} - z^{1/2}}{2i} \tag{238}$$

This gives

$$1 = \Big(\cos^2(\pi\omega) + \sin^2(\pi\omega)\Big)^{2N-1} = \tag{239}$$

$$\sum_{n=0}^{2N-1} \frac{(2N-1)!}{n!(2N-1-n)!} \cos^{2n}(\pi\omega)\sin^{4N-2-2n}(\pi\omega) = \tag{240}$$

$$\sum_{n=0}^{2N-1} \frac{(2N-1)!}{n!(2N-1-n)!} \left(\frac{z^{1/2} + z^{*1/2}}{2}\right)^{2n} \left(\frac{z^{*1/2} - z^{1/2}}{2i}\right)^{4N-2-2n} = \tag{241}$$

$$\sum_{n=0}^{2N-1} \frac{(2N-1)!}{n!(2N-1-n)!} \left(\frac{z+1}{2}\right)^{2n} \left(\frac{1-z}{2i}\right)^{4N-2-2n} z^{*(2N-1)} = \tag{242}$$

It is useful to express the first N-terms in terms of cosines and sines:

$$1 = \sum_{n=0}^{2N-1} \frac{(2N-1)!}{n!(2N-1-n)!} \cos^{2n}(\pi\omega)(1-\cos^2(\pi\omega))^{2N-1-n} = \tag{243}$$

This sum has the following properties:

- Has 2N terms.
- The last N terms in this sum have the desired zero at -1 with the correct multiplicity.
- Let P_+ be the sum of the last N terms and P_- be the sum of the first N terms.
- Both sums are non-negative.

Note that

$$P_+ = \sum_{n=N}^{2N-1} \frac{(2N-1)!}{n!(2N-1-n)!} \cos^{2n}(\pi\omega)\sin^{4N-2-2n}(\pi\omega) = \tag{244}$$

setting $m = 2N - 1 - n$ gives

$$\sum_{m=0}^{N-1} \frac{(2N-1)!}{m!(2N-1-m)!} \cos^{4N-2-2m}(\pi\omega)\sin^{2m}(\pi\omega) = \tag{245}$$

$$\cos^{2N}(\pi\omega)\sum_{m=0}^{N-1} \frac{(2N-1)!}{m!(2N-1-m)!} \left(1-\sin^2(\pi\omega)\right)^{N-1-m} \sin^{2m}(\pi\omega) \tag{246}$$

Let

$$W(x) = \sum_{m=0}^{N-1} \frac{(2N-1)!}{m!(2N-1-m)!} \left(1-x^2\right)^{N-1-m} x^{2m} \tag{247}$$

Note that

$$P_+ = \left|\frac{1+z}{2}\right|^{2N} W(\sin(\pi\omega)). \tag{248}$$

Since W is non-negative function of $\sqrt{z}$ and $\sqrt{z^*}$, it has a non-negative square root ,$R(\sqrt{z}, \sqrt{z^*})$.

Claim (proof needs to be supplied) $W = |R(z)|^2$ where $R(z)$ is a polynomial. If this is true then

$$P(z) = \left(\frac{1+z}{2}\right)^N R(z) \tag{249}$$

has all of the desired properties.

Case $N = 2$. In this case

$$W(x) = \frac{3!}{0!3!}(1-x^2)x^0 + \frac{3!}{1!2!}(1-x^2)^0 x^2 = 1 + 2x^2 = 1 - \frac{1}{2}(\sqrt{z^*} - \sqrt{z})^2 = \tag{250}$$

$$2 - \frac{1}{2}(z + z^*). \tag{251}$$

The coefficients of the polynomial R should be real for real scaling coefficients. Try Let $R(z) = a + bz$; $|R(z)|^2 = a^2 + b^2 + ab(z + z^*)$. Equating coefficients gives

$$a^2 + b^2 = 2 \qquad 2ab = -1 \tag{252}$$

These are consistent with

$$(a-b)^2 = 3; \qquad (a+b)^2 = 1 \tag{253}$$

$$a - b = \pm\sqrt{3} \qquad a + b = \pm 1 \tag{254}$$

or

$$a = \pm(\frac{1}{2} \pm \frac{\sqrt{3}}{2}) \qquad b = \pm(\frac{1}{2} \mp \frac{\sqrt{3}}{2}) \tag{255}$$

The normalization $R(1) = 1$ gives

$$R(z) = (\frac{1}{2} \pm \frac{\sqrt{3}}{2}) + (\frac{1}{2} \mp \frac{\sqrt{3}}{2})z \tag{256}$$

and finally

$$P(z) = \left(\frac{1+z}{2}\right)^2 ((\frac{1}{2} \pm \frac{\sqrt{3}}{2}) + (\frac{1}{2} \mp \frac{\sqrt{3}}{2})z) \tag{257}$$

from which on can read off the $N = 2$ Daubechies coefficients. The two sign choices are related by the symmetry $z \to z^{-1}$ followed by multiplication by a homogeneous polynomial to preserve the polynomial nature of the function.

11 Derivatives and Differential Equations

In order to use use wavelets for differential equations is necessary to be able to compute derivatives.

A typical wavelet basis consists of a the scaling function and its translates at a fixed resolution m:

$$\{\phi_{mn}(x)\}_{n=-\infty}^{\infty} : \phi_{mn}(x) = D^m T^n \phi(x) = 2^{-m/2}\phi(2^m x - n) \tag{258}$$

and the wavelets for all resolutions k less than or equal to m and their translates

$$\{M_{kn}(x)\}_{n=-\infty,k=-\infty}^{\infty,m} : M_{mn}(x) = D^k T^n M(x) = 2^{-k/2} M(2^k x - n) \tag{259}$$

Given a function $f(x)$ represented as

$$f(x) = \sum_n f^s_{mn}\phi_{mn}(x) + \sum_{kn} f^m_{kn} M_{kn}(x) \tag{260}$$

The r-th derivative of this $f(x)$ can be represented in the following forms

$$f^{(r)}(x) = \sum_n f^s_{mn}\phi^{(r)}_{mn}(x) + \sum_{kn} f^m_{kn} M^{(r)}_{kn}(x) \tag{261}$$

or

$$f^{(r)}(x) = \sum_n f^{s(r)}_{mn} \phi_{mn}(x) + \sum_{kn} f^{m(r)}_{kn} M_{kn}(x). \tag{262}$$

The coefficients in these two expansions can be related by expanding the derivatives of the basis functions in terms of the basis:

$$\phi^{(r)}_{mn}(x) = \sum_{n'} \phi_{mn'}(x) \Gamma^{1(r)}_{mn';mn} + \sum_{k'n'} M_{k'n'}(x) \Gamma^{2(r)}_{k'n';mn} \tag{263}$$

and

$$M^{(r)}_{kn}(x) = \sum_{n'} \phi_{mn'}(x) \Gamma^{3(r)}_{mn';kn} + \sum_{k'n'} M_{k'n'}(x) \Gamma^{4(r)}_{k'n';kn} \tag{264}$$

In terms of these quantities

$$f^{s(r)}_{mn} = \sum_{n'} \Gamma^{1(r)}_{mn;mn'} f^s_{mn'} + \sum_{n'k'} \Gamma^{3(r)}_{mn;k'n'} f^m_{k'n'} \tag{265}$$

$$f^{m(r)}_{kn} = \sum_{n'} \Gamma^{2(r)}_{kn;mn'} f^s_{mn'} + \sum_{n'k'} \Gamma^{4(r)}_{kn;k'n'} f^m_{k'n'} \tag{266}$$

The expansion coefficients are the overlap matrices

$$\Gamma^{1(r)}_{mn;m'n'} := (\phi_{mn}, \phi^{(r)}_{m'n'}) \tag{267}$$

$$\Gamma^{2(r)}_{kn;m'n'} := (M_{kn}, \phi^{(r)}_{m'n'}) \tag{268}$$

$$\Gamma^{3(r)}_{mn;k'n'} := (\phi_{mn}, M^{(r)}_{k'n'}) \tag{269}$$

$$\Gamma^{4(r)}_{kn;k'n'} := (M_{mn}, M^{(r)}_{k'n'}) \tag{270}$$

The scaling equation mean that all of these coefficients can be determined from a subset of the coefficients. In order to exhibit the key relations it is useful to use operators:

$$Df(x) = \frac{1}{\sqrt{2}} f(\frac{x}{2}) \tag{271}$$

$$Tf(x) = f(x-1) \tag{272}$$

$$\Delta f(x) = \frac{df}{dx}(x) \tag{273}$$

Direct computation shows

$$\Delta D = \frac{1}{2} D \Delta \tag{274}$$

$$DT = T^2 D \tag{275}$$

$$\Delta T = T\Delta \tag{276}$$

$$\Delta^\dagger = -\Delta \qquad T^\dagger = T^{-1} \qquad D^\dagger = D^{-1} \tag{277}$$

We also have the scaling equations:

$$D\phi = \sum_l h_l T^l \phi \tag{278}$$

$$DM = \sum_l g_l T^l \phi \tag{279}$$

Using the operator relations above

$$D\Delta^r \phi = 2^r \sum_l h_l T^l \Delta^r \phi \tag{280}$$

$$D\Delta^r M = 2^r \sum_l g_l T^l \Delta^r \phi \tag{281}$$

In terms of these operators

$$\Gamma^{1(r)}_{m'n';mn} := (D^{m'} T^{n'} \phi, \Delta^r D^m T^n \phi) \tag{282}$$

$$\Gamma^{2(r)}_{m'n';kn} := (D^{m'} T^{n'} M, \Delta^r D^m T^n \phi) \tag{283}$$

$$\Gamma^{3(r)}_{k'n';mn} := (D^{k'} T^{n'} \phi, \Delta^r D^m T^n M) \tag{284}$$

$$\Gamma^{4(r)}_{k'n';kn} := (D^{k'} T^{n'} M, \Delta^r D^m T^n M) \tag{285}$$

In order to evaluate these coefficients the following s.pdf are used:

1. Move all of the factors of D to a single side of the equation. Choose the side where the power of D is positive.

2. Move the D' through all derivatives.

3. Use the scaling equations to eliminate all of the D's.

4. Move all of the T' to the right side of the scalar product.

Using these all of the Γ's can be expressed in terms of

$$\Gamma^{1(r)}_{0n;00} := (T_n \phi, \Delta^r \phi) \tag{286}$$

$$\Gamma^{2(r)}_{0n;00} := (T^n M, \Delta^r \phi) \tag{287}$$

$$\Gamma^{3(r)}_{0n;00} := (T^n\phi, \Delta^r M,) \tag{288}$$

$$\Gamma^{4(r)}_{0n;00} := (T^n M, \Delta^r M) \tag{289}$$

These quantities satisfy scaling equations. To see this consider

$$\Gamma^{1(r)}_{0n;00} := (T^n\phi, \Delta^r\phi) =$$

$$(DT^n\phi, D\Delta^r\phi) = 2^r(T^{2n}D\phi, \Delta^r D\phi) =$$

$$\sum_{ll'} h^*_l h_{l'} 2^r (T^{2n+l}\phi, T^{l'}\Delta^r\phi) =$$

$$\sum_{ll'} h^*_l h_{l'} 2^r (T^{2n+l-l'}\phi, \Delta^r\phi) =$$

$$\sum_{ll'} h^*_l h_{l'} 2^r \Gamma^{1(r)}_{0,2n+l-l';00} =$$

$$\sum_k (\sum_{l'} h^*_{k-2n+l'} h_{l'}) 2^r \Gamma^{1(r)}_{0k;00}\text{`} \tag{290}$$

This gives an eigenvalue equation for the vector $\Gamma^{1(r)}_{0n;00}$. The solution is the eigenvector with eigenvalue 1. The normalization can be fixed by computing $\Gamma^{1(r)}_{00;00} := (\phi, \Delta^r\phi)$ directly

Define $H^{(r)}_{nk}$ by

$$H^{(r)}_{nk} := \sum_{l'} h^*_{k+l'-2n} h_{l'} 2^r \tag{291}$$

With this definition the eigenvalue problem takes on the form:

$$\Gamma^{1(r)}_{0n;00} = \sum_k H^{(r)}_{nk} \Gamma^{1(r)}_{0k;00} \tag{292}$$

Note that this can be written as

$$2^{-r}\Gamma^{1(r)}_{0n;00} = \sum_k H^0_{nk} \Gamma^{1(r)}_{0k;00} \tag{293}$$

This treats all allowed derivatives with a single equation - the eigenvector with eigenvalue 2^{-r} is the coefficient for the r^{th} derivative.

These quantities can be used to determine all of the other quantities using

$$\Gamma^{2(r)}_{0n;00} := (T^n M, \Delta^r\phi) =$$

$$\sum_{ll'} g_{l'}^* h_l 2^r \Gamma^{1(r)}_{0,2n+l'-l;00} =$$

$$\sum_m (\sum_l g_{l'}^*) h_{2n+l-m}) 2^r \Gamma^{1(r)}_{0'm;00} \qquad (294)$$

$$\Gamma^{3(r)}_{0n;00} := (T^n \phi, \Delta^r M) =$$

$$\sum_{ll'} h_{l'}^* g_l 2^r \Gamma^{1(r)}_{0,2n+l'-l;00} =$$

$$\sum_m (\sum_l h_l^* g_{2n+l-m}) 2^r \Gamma^{1(r)}_{0m;00} \qquad (295)$$

$$\Gamma^{4(r)}_{0n;00} := (T^n M, \Delta^r M,) =$$

$$\sum_{ll'} g_l^* g_{l'} 2^r \Gamma^{1(r)}_{0,2n+l'-l;00} =$$

$$\sum_m (\sum_l g_l^* g_{2n+l-m}) 2^r \Gamma^{1(r)}_{0m;0} \qquad (296)$$

This shows that all of the expansion coefficients for any number of derivatives can be constructed from the solutions of a single eigenvalue problem.

There are a number of stepping formulas:

$$\Gamma^{1(r)}_{0m;0n} = \Gamma^{1(r)}_{0m-n;00} \qquad (297)$$

$$\Gamma^{1(r)}_{0m+k;0,n+k} = \Gamma^{1(r)}_{0m;0n} \qquad (298)$$

$$\Gamma^{1(r)}_{kn;tm} = \sum_l h_l^* \Gamma^{1(r)}_{k-1,2n+l;tm} \qquad (299)$$

$$\Gamma^{1(r)}_{kn;tm} = 2^r \sum_l h_l \Gamma^{1(r)}_{k+1,n;t,2m+l} \qquad (300)$$

$$\Gamma^{1(r)}_{kn;tm} = \sum_l h_l \Gamma^{1(r)}_{kn;t-1,2m+l} \qquad (301)$$

$$\Gamma^{1(r)}_{kn;tm} = 2 \sum_l h_l^* \Gamma^{1(r)}_{k,2n+1;t+1,2m+l} \qquad (302)$$

which can be used to reduce the number of dilitations to zero.

12 Moments and Wavelets

Consider the equations for the moments of the scaling function and the mother function:

$$\langle x^m \rangle_\phi := \int x^m \phi(x) dx \tag{303}$$

$$\langle x^m \rangle_m := \int x^m m(x) dx \tag{304}$$

To develop relations among the moments use the unitarity of the dilitation operator:

$$\langle x^m \rangle_\phi = \int x^m \phi(x) dx =$$

$$\int D x^m D\phi(x) dx = \frac{1}{2^{m+1/2}} \int x^m D\phi(x) dx \tag{305}$$

Using the scaling equation gives

$$\frac{1}{2^{m+1/2}} \sum_l \int x^m h_l \phi(x-l) dx =$$

$$\frac{1}{2^{m+1/2}} \sum_l \int (l+y)^m h_l \phi(y) dy =$$

Nest use the binomial theorem to get:

$$\frac{1}{2^{m+1/2}} \sum_{k=0}^{m} \sum_l h_l \frac{m!}{k!(m-k)!} l^{m-k} \int y^k \phi(y) dy =$$

$$\frac{1}{2^{m+1/2}} \sum_{k=0}^{m} \sum_l h_l \frac{m!}{k!(m-k)!} l^{m-k} \langle x^k \rangle_\phi \tag{306}$$

This gives an a set eigenvalue equations for $\langle x^k \rangle_\phi$ The moments of the mother function have similar properties with h_l replaced by g_l: :

$$\langle x^m \rangle_m = \int x^m m(x) dx =$$

$$\int D x^m D m(x) dx =$$

$$\frac{1}{2^{m+1/2}}\sum_{k=0}^{m}\sum_{l} g_l \frac{m!}{k!(m-k)!} l^{m-k}\langle x^k\rangle_\phi \tag{307}$$

The nice thing about these equations is that they are lower triangular. Specifically for the $\langle x^k\rangle_\phi$ equations we have for $m = 0$:

$$\langle x^0\rangle_\phi = \frac{1}{2^{1/2}}\sum_l h_l \langle x^0\rangle_\phi \tag{308}$$

which is the constraint $\sum_l h_l = \sqrt{2}$. For $m = 1$ this equation is

$$\langle x^1\rangle_\phi = \frac{1}{2^{3/2}}\sum_l h_l \left(\frac{1!}{0!1!} l^1 \langle x^0\rangle_\phi + \frac{1!}{1!0!} l^0 \langle x^1\rangle_\phi\right) \tag{309}$$

Using the first equation gives

$$\langle x^1\rangle_\phi = \frac{1}{\sqrt{2}}\sum_l l h_l \langle x^0\rangle_\phi \tag{310}$$

This gives $\langle x^1\rangle_\phi$ in terms of $\langle x^0\rangle_\phi$

For $m = 2$ is equation becomes

$$\langle x^2\rangle_\phi =$$

$$\frac{1}{2^{5/2}}\sum_l h_l \left(\frac{2!}{0!2!} l^2 \langle x^0\rangle_\phi + \frac{2!}{1!1!} l^1 \langle x^1\rangle_\phi + \frac{2!}{2!0!} l^0 + \langle x^2\rangle_\phi\right) =$$

$$\frac{1}{2^{5/2}}\left(\sum_l h_l l^2 \langle x^0\rangle_\phi + \sqrt{2}(\sum l h_l)^2 \langle x^0\rangle_\phi + \sqrt{2}\langle x^2\rangle_\phi\right)$$

which can be solved for $\langle x^2\rangle_\phi$ in terms of $\langle x^0\rangle_\phi$

The general form of the recursion relation is

$$\langle x^m\rangle_\phi =$$

$$\frac{m!}{2^{m+1}-2}\sum_{k=0}^{m-1}\sum_l \frac{h_l l^{m-k}}{k!(m-k)!}\langle x^k\rangle_\phi \tag{311}$$

A similar recursion can be derived for moments of the mother function:

$$\langle x^m\rangle_m =$$

$$\frac{m!}{2^{m+1}} \sum_{k=0}^{m-1} \sum_l \frac{g_l l^{m-k}}{k!(m-k)!} \langle x^k \rangle_\phi \tag{312}$$

There are all fixed in terms of the value of the integral $\int \phi(x)dx$.

It follows from this relation that the condition that the first p moments of the mother function vanishing are

$$\sum_l l^k g_l = 0 \quad k = 0, 1, \cdots, p \tag{313}$$

These equations come from looking at the coefficients of the moments of the scaling function. They give all of the moments in terms of the h_l's and

$$\langle x^0 \rangle_\phi = \int \phi(x)dx \tag{314}$$

These equations, along with the normalization equation and the orthonormality equation determine the h_l's for the Daubechies wavelets. The general equations are

$$\sum_l h_l h_{l-2k} = \delta_{k0} \qquad k = 0, 1, \cdots \tag{315}$$

$$\sum_l h_l = \sqrt{2} \tag{316}$$

$$\sum_l l^k (-)^l h_{1-l} = 0 \quad k = 0, 1, \cdots, p \tag{317}$$

where 1 can be replaced by any odd integer.y

Moments of all of the other functions can be computed using

$$\langle x^m \rangle_{\phi_{rs}} = \int x^m D^r T^s \phi(x)dx =$$

$$\int (T^{-s} D^{-r} x^m) \phi(x)dx = 2^{\frac{r}{2}+mr} \int (x+s)^m \phi(x)dx =$$

$$2^{\frac{r}{2}+mr} \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} s^{m-k} \langle x^r \rangle_\phi$$

and

$$\langle x^m \rangle_{m_{rs}} = \int x^m D^r T^s m(x)dx =$$

$$\int (T^{-s} D^{-r} x^m) m(x)dx = 2^{\frac{r}{2}+mr} \int (x+s)^m m(x)dx =$$

$$2^{\frac{r}{2}+mr}\sum_{k=0}^{m}\frac{m!}{k!(m-k)!}s^{m-k}\langle x^r\rangle_m$$

We can use these to determine the normalization of the $\Gamma's$: Consider

$$\frac{d\phi}{dx}(x)=\sum_l \phi_{ml}(x)\Gamma^{1(1)}_{ml;00}+\sum_{kl} m_{kl}(x)\Gamma^{2}_{kl;00} \tag{318}$$

Multiply by x to get

$$\langle x^0\rangle_\phi=\int\phi(x)dx=-\int x\frac{d\phi}{dx}(x)dx=$$

$$\sum_l\langle x^1\rangle_{\phi_{ml}}\Gamma^{1(1)}_{ml;00}+\sum_{kl}\langle x^1\rangle_{m_{kl}}\Gamma^{2(1)}_{kl;00}$$

Given that we know the Γ's up to an overall normalization, and all of the moments in terms of $\langle x^0\rangle$, this equation fixes the normalization of the $\Gamma^{(1)}$'s in terms of $\langle x^0\rangle$.

The Γ's corresponding to the higher derivatives can be derived using

$$\int x^l\frac{d^l\phi}{dx^l}(x)dx=(-)^l l!\langle x^0\rangle_\phi=$$

$$\sum_s\langle x^l\rangle_{\phi_{ms}}\Gamma^{1(l)}_{ms;00}+\sum_{ks}\langle x^l\rangle_{m_{ks}}\Gamma^{2(l)}_{ks;00} \tag{319}$$

which expresses the Γ's for the higher derivatives in terms of higher moments.

13 Normalization issues

The orthonormality of the wavelets means that the scaling function $\phi(x)$ satisfies

$$\int\phi^2(x)dx=1. \tag{320}$$

Because this is preserved under the continuous scale transformation

$$\phi(x)\to\phi'(x):=\frac{1}{\sqrt{s}}\phi(\frac{x}{s}) \tag{321}$$

by choosing s we can simultaneously fix

$$\int\phi'(x)dx=\sqrt{s}\int\phi(x)dx \tag{322}$$

Choosing s so $\int \phi'(x)dx = 1$ fixes a starting scale. In this way we can simultaneously require

$$\int \phi^2(x)dx = \int \phi(x)dx = 1 \tag{323}$$

Given the conditions

$$\int \phi^2(x)dx = \int \phi(x)dx = 1 \tag{324}$$

we can compute all of the necessary normalization coefficients.

First we calculate the expansion of 1. In this case the assumption that the integral of the mother function gives zero implies that 1 can be expressed in terms of translates of the scaling function:

$$1 = \sum c_n^0 \phi(x-n) \tag{325}$$

Multiplying by $\phi_m(x) = \phi(x-m)$ and using the orthonormality and $\int \phi(x)dx = 1$ gives

$$1 = c_m = \int \phi_m(x)dx = \int \phi(x)dx = \beta_0 \tag{326}$$

Thus we conclude that

$$1 = \sum_n \phi_n(x) \tag{327}$$

Next we show that we can compute all of the moments

$$\beta_m := \int x^m \phi(x)dx \tag{328}$$

in terms of the h_l's and β_0. To see this note

$$\beta_m := \int x^m \phi(x)dx = \int Dx^m D\phi(x)dx =$$

$$\frac{1}{\sqrt{2}} \int \left(\frac{x}{2}\right)^m \sum_l h_l \phi(x-l)dx =$$

$$2^{-m-\frac{1}{2}} \sum_l h_l \int (x+l)^m \phi(x)dx =$$

$$2^{-m-\frac{1}{2}} \sum_l h_l \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} l^{m-k} \beta_k \tag{329}$$

Putting the $k = m$ term on the left and using $\sum_l h_l = \sqrt{2}$ gives

$$\beta_m = \frac{1}{2^m - 1} \frac{1}{\sqrt{2}} \sum_l h_l \sum_{k=0}^{m-1} \frac{m!}{k!(m-k)!} l^{m-k} \beta_k \tag{330}$$

This give an explicit expression for the m-th moment in terms of $\beta_0 = 1, \beta_1, \cdots, \beta_{m-1}$ and the scaling coefficients. This determines all of the moments.

Note that these equations follow directly from the scaling equations and the normalization condition that gives β_0. No special properties of the mother function have been used.

Given the moments we show that they can be used to compute the expansion coefficients for all monomials x^l that vanish when integrated against the mother function.

To see this note

$$x^l = \sum_k c^l_k \phi_k(x) \tag{331}$$

Using orthonormality gives

$$c^l_m = \int \phi_m(x) x^l dx = \int \phi(x)(x+m)^l dx$$

$$\sum_{k=0}^{l} \frac{l!}{k!(l-k)!} m^{l-k} \beta_k$$

which expresses these expansion coefficients in terms of the previously computed moments.

We can use these to get inhomogeneous equations for the Γs:

$$x^l = \sum_k c^l_k \phi_k(x) \tag{332}$$

If we differentiate this l times, multiply by $\phi(x)$, and integrate we get

$$l! \int \phi(x) dx = \sum_k c^l_k \int \phi(x) \frac{d^l \phi_k}{dx^l}(x-m) dx =$$

$$\sum_k c^l_k \int \phi(x+k) \frac{d^l \phi}{dx^l}(x) dx$$

which becomes

$$1 = \frac{1}{l!} \sum_k c_k^l \Gamma_{0-k;00}^l \tag{333}$$

This gives the needed inhomogeneous equation to determine the Γs for any allowed derivative.

Thus we have shown how $\beta_0 = 1$ determines all of the necessary normalizations.

14 Integral equations

Consider an integral equation of one of the two forms

$$f(x) = \int K_\lambda(x, y) f(y) dy \tag{334}$$

where λ is a parameter like an eigenvalue or

$$f(x) = g(x) + \int K(x, y) f(y) dy \tag{335}$$

In the first equation there will not be solutions unless λ takes on certain values. Each solution is associated with a specific value of λ. This can be considered like a matrix equation, where the matrix does not have an inverse for certain values of λ.

In the second equation there is no parameter. The function $g(x)$ is given this looks like a system of linear equations.

These equations can be solved using Galerkin or collocation methods. For the Galerkin method the solution can be expanded in terms of an orthonormal basis (wavelets):

$$f(x) = \sum_n \phi_n(x) f_n \tag{336}$$

where the f_n are unknown expansion coefficients. For the case of wavelets the index n is replaced by two indices.

Inserting this solution into either of the above equations gives

$$\sum_n \phi_n(x) f_n = \int K_\lambda(x, y) \sum_n \phi_n(y) f_n dy \tag{337}$$

or

$$\sum_n \phi_n(x) f_n = g(x) + \int K(x, y) \sum_n \phi_n(y) f_n dy \tag{338}$$

Assuming that both expression are well behaved it is possible to change the order of the sum and the integral to obtain:

$$\sum_n \phi_n(x) f_n = \sum_n \int K_\lambda(x,y)\phi_n(y) dy f_n \tag{339}$$

$$\sum_n \phi_n(x) f_n = g(x) + \sum_n \int K(x,y)\phi_n(y) dy f_n \tag{340}$$

This is not yet an equation. In applications the infinite sum has to be replaced by a finite sum. The equation then becomes approximate.

For the Galerkin method the equation is required to hold on the subspace generated by the first N basis functions. The equations for the coefficients are obtained by integrating against $\phi_m^*(x)$ for $m = 1, \cdots, N$. Using the orthonormality of the basis functions gives:

$$f_m = \sum_{n=1}^{N} \int \phi_m^*(x) K_\lambda(x,y)\phi_n(y) dy f_n \tag{341}$$

$$f_m = \int \phi_m^*(x) g(x) dx + \sum_{n=1}^{N} \int \phi_m^*(x) K(x,y)\phi_n(y) dy f_n \tag{342}$$

If we define

$$K_{\lambda mn} := \int \phi_m^*(x) K_\lambda(x,y)\phi_n(y) dx dy \tag{343}$$

$$K_{mn} := \int \phi_m^*(x) K(x-y)\phi_n(y) dx dy \tag{344}$$

$$g_m := \int \phi_m^*(x) g(x) dx \tag{345}$$

these equations become

$$f_m = \sum_{n=1}^{N} K_{\lambda mn} f_n \qquad m \in \{1, 2, \cdots N\} \tag{346}$$

$$f_m = g_m + \sum_{n=1}^{N} K_{mn} f_n \qquad m \in \{1, 2, \cdots N\} \tag{347}$$

These equation are linear algebraic equations for the coefficients f_n. The approximate solution can be expressed in terms of solution for these coefficients as:

$$f(x) \sim \sum_{n=1}^{N} \phi_n(x) f_n \tag{348}$$

This can be improved using the equation

$$f(x) \sim \sum_{n=1}^{N} \int K_\lambda(x,y) \sum_n \phi_n(y) dy f_n \tag{349}$$

or

$$f(x) \sim g(x0 + \sum_{n=1}^{N} \int K(x,y) \sum_n \phi_n(y) dy f_n \tag{350}$$

These equations can also be approached using the collocation method. For the collocation method the number of basis functions is also truncated. Rather than projecting on the subspace generated by the first N basis functions the equations are required to be exact at N points $\{x_i\}_{i=1}^{N}$. This gives a different set of linear algebraic equations:

$$\sum_n \phi_n(x_m) f_n = \sum_{n=1}^{N} \int K_\lambda(x_m,y)\phi_n(y) dy f_n \qquad m \in \{1, \cdots N\} \tag{351}$$

$$\sum_n \phi_n(x_m) f_n = g(x_m) + \sum_{n=1}^{N} \int K(x_m,y)\phi_n(y) dy f_n \qquad m \in \{1, \cdots N\} \tag{352}$$

In this case the equations have the structure

$$\sum_{n=1}^{N} \phi_{mn} f_n = \sum_{n=1}^{N} K_{\lambda mn} f_n \tag{353}$$

$$\sum_{n=1}^{N} \phi_{mn} f_n = g_m + \sum_{n=1}^{N} K_{mn} f_n \tag{354}$$

where

$$\phi_{mn} := \phi_n(x_m) \tag{355}$$

$$K_{\lambda mn} := \int K_\lambda(x_m,y)\phi_n(y) dy \tag{356}$$

$$K_{mn} := \int K(x_m,y)\phi_n(y) dy \tag{357}$$

$$g_m := g(x_m) \tag{358}$$

These equation give approximate values of the f_n, and an approximate solution of the form

$$f(x) \sim \sum_{n=1}^{N} \phi_n(x) f_n \tag{359}$$

This can be improved by the same interpolation method that was used in the Galerkin method.

In both cases the use of wavelet method will allow for efficient computation of the basis functions. The advantage of the integral equation method is that there there are no problems with integrating functions, like the Haar functions, that have discontinuous derivatives.

In order to obtain any additional benefit our of the wavelet basis the kernel $K(x,y)$ needs to have additional properties. In many cases of practical interest the kernel is translationally invariant. This means the

$$K(x,y) = K(x-y). \tag{360}$$

I consider the Galerkin case, where the these benefits are of most value. For a translationally invariant kernel the Galerkin method involves computing matrix elements of the general form:

$$\int \phi^*_{mn}(x) K(x-y) \phi_{kl}(y) dx dy \tag{361}$$

where here I have introduced indices on the scaling function. Similar equations are needed for the mother function. Define

$$c_n := \int \phi^*(x) K(x-y+n) \phi(y) dx dy \tag{362}$$

where $\phi(x)$ is the scaling function. Note that

$$K_{mn:kl} := \int (D^m T^n \phi)^*(x) K(x-y) (D^k T^l \phi)(y) dx dy \tag{363}$$

Using the scaling equations we show that it is possible to reduce the values of m and k respectively:

$$K_{mn:kl} := \int (D^m T^n \phi)^*(x) K(x-y) (D^k T^l \phi)(y) dx dy = \tag{364}$$

$$\sum_r h_r \int (D^m T^n \phi)^*(x) K(x-y) (D^{k-1} T^{2l+r} \phi)(y) dx dy = \sum_r h_r K_{mn:k-1,2l+r} \tag{365}$$

$$\sum_r h_r^* \int (D^{m-1}T^{2n+r}\phi)^*(x)K(x-y)(D^kT^l\phi)(y)dxdy = \sum_r h_f^* K_{m-1,2n+r:kl} \tag{366}$$

These equations show that it is possible to successively reduce the value of m and k. Once these are reduced to 0 what remains is

$$K_{0n:0l} := \int (T^n\phi)^*(x)K(x-y)(T^l\phi)(y)dxdy = \int \phi^*(x)K(x+n-y-l)\phi(y)dxdy = c_{n-l} \tag{367}$$

Similar results can be obtained for the mother functions. In that case the h_m's are replaced by the g_m's.

The problem with this is that it does not cover the case of negative values of m and k. We conclude that given a small enough base scale, translational invariance can be used to generate matrix elements of the kernel on all larger scales, in terms of translates on the base scale.

In order to deal with negative values of m and k the Kernel must have additional properties with respect to scale transformations. Unfortunately, unlike translations, most integral equation do not have scale invariant kernels. Roughly speaking, parameters with physical dimensions appearing in equations break scale invariance. The Maxwell's equation are an exception. They to not have a natural distance scale.

We say that a kernel has scale dimension s if

$$2^{-sn}K(2^n(x-y)) = K(x-y) \tag{368}$$

For example

$$K(x-y) := \frac{1}{|x-y|^{1/2}} \tag{369}$$

has scale dimension $s = -1/2$.

For these kernels direct integration shows

$$(f, KDg) := \int f^*(x)K(x-y)(Dg)(y)dxdy = \tag{370}$$

$$\int f^*(x)K(x-y)(Dg)(y)dxdy = \tag{371}$$

$$\int f^*(x)K(x-y)\frac{1}{\sqrt{2}}g(y/2)dxdy = \tag{372}$$

$$\frac{2}{\sqrt{2}}\int f^*(x)K(x-2u)g(u)dxdu = \tag{373}$$

$$\frac{4}{\sqrt{2}} \int f^*(2v) K(2(v-u)) g(u) dv du = \tag{374}$$

$$2 \int (D^\dagger f^*)(v) K(2(v-u)) g(u) dv du = \tag{375}$$

$$2^{1+s} \int (D^\dagger f^*)(v) K(v-u) g(u) dv du = \tag{376}$$

$$2^{1+s} (D^\dagger f, Kg) \tag{377}$$

Similarly it is possible to show

$$(Df, Kg) = 2^{s+1} (f, K D^\dagger g) \tag{378}$$

and

$$(Df, KDg) = 2^{s+1} (f, Kg) \tag{379}$$

$$(D^\dagger f, K D^\dagger g) = 2^{-(s+1)} (f, Kg) \tag{380}$$

These equations allow reductions of the form

$$(D^{-n} T^m \phi, K D^{-k} T^l \phi) = \tag{381}$$

assuming $n > k$ this becomes

$$2^{m(s+1)} (T^m \phi, K D^{n-k} T^l \phi) = \tag{382}$$

$$2^{m(s+1)} (\phi, K T^{2^{n-k} l - m} D^{n-k} \phi) \tag{383}$$

In this form the scaling equations can be used to eliminate powers of D. Similar relations can be derived fort he case $k > n$ and $k = n$.

We conclude that if the kernel is scale and translationally invariant, the matrix elements for the Galerkin method can be expressed in terms of matrix elements of translates of the scaling function and mother function.

15 Pyramid Method

Consider a function $\chi(x)$ that is periodic on $[0, L]$.

Divide this interval into 2^N subintervals of length $\Delta := L/2^N$

Let $\phi(x)$ be a Daubechies scaling function. Let

$$\psi(x) = \frac{1}{\sqrt{\Delta}} \phi(\frac{x}{\Delta}) \tag{384}$$

This has the following properties:

$$\int \psi(x)\psi(x-n\Delta) = \frac{1}{\Delta}\phi(\frac{x}{\Delta})\phi(\frac{x}{\Delta}-n)dx \int \phi(y)\phi(y-n)dy = \delta_{n0} \quad (385)$$

$$\sum_n \phi(x-n\Delta) = \sum_n \frac{1}{\sqrt{\Delta}}\psi(\frac{x}{\Delta}-n) = \frac{1}{\sqrt{\Delta}} \quad (386)$$

which follow from the orthogonality of the translated scaling function and the moment condition of the father wavelet (we have assumed both orthonormality of the translated wavelets and $\sum_n(\phi(x-n) = 1$ - if there are not compatible we need to readjust the constant).

We define

$$\psi_n(x) = \psi(x-n\Delta) \quad (387)$$

$$\psi_{kn}(x) = D^k\psi_n(x) = 2^{-k/2}\psi_n(\frac{x}{2^k}) \quad (388)$$

as well as the corresponding expressions for the mother function

$$u(x) = \frac{1}{\sqrt{\Delta}}m(\frac{x}{\Delta}) \quad (389)$$

$$u_n(x) = u(x-n\Delta) \quad (390)$$

$$u_{kn}(x) = D^k u_n(x) = 2^{-k/2}u(\frac{x}{2^k}) \quad (391)$$

The scaling equations for $\psi(x)$ and $u(x)$ are

$$D\psi(x) = \sum_n h_n\psi(x-n\Delta) \quad (392)$$

$$Du(x) = \sum_n g_n\psi(x-n\Delta) \quad (393)$$

Approximate $\chi(x)$ at resolution 2^N on $[0,L]$ by

$$\chi_{2^N}(x) := \sum_{n=1}^{2^N} c_n\psi(x-n\Delta) \quad (394)$$

where

$$c_n = \int \psi(x-n\Delta)\chi(x)dx \quad (395)$$

Here we assume that the integrals extend past $[0, L]$ by extending the functions so they are periodic. The coefficients c_n are an approximate representation of the function. Because

$$\chi(x) = \chi(x) \sum_n \sqrt{\Delta} \phi(x - n\Delta) \tag{396}$$

we expect that the sum of the c_n for the basis functions that are non-vanishing at x should be approximately the value of χ at x times $\sqrt{\Delta}$.

The approximation $\chi_{2^N}(x)$ is finest resolution approximation of the exact function. This is the V_0 representation of the function. This can be decomposed into a pair of vectors of length 2^{N-1} corresponding to the V_1 and W_1 subspace.

On these subspaces the expansions are

$$\chi_{2^N}(x) = \sum c_{1n} D\psi_n(x) + \sum d_{1n} D u_n(x) \tag{397}$$

Using the scaling relations gives

$$\chi_{2^N}(x) = \sum_n c_{1n} \sum_m h_m \psi(x - 2n\Delta - m\Delta) + \sum d_{1n} \sum_m g_m \psi(x - 2n\Delta - m\Delta) \tag{398}$$

Integrating against $\phi(x - k\Delta)$ gives

$$c_k = \sum_n (c_{1n} \sum_m h_{k-2n} + d_{1n} \sum_m g_{k-2n}) \tag{399}$$

This can be expressed as a matrix equation:

$$\begin{pmatrix} c_1 \\ \vdots \\ c_2^N \end{pmatrix} = \begin{pmatrix} h_{1-2} \cdots h_{1-2^{2N}} & g_{1-2} \cdots g_{1-2^{2N}} \\ \cdots & \cdots \\ h_{2^N-2} \cdots h_{2^N-2^{2N}} & g_{2^N-2} \cdots g_{2^N-2^{2N}} \end{pmatrix} \begin{pmatrix} c_{11} \\ \vdots \\ d_{12^{N-1}} \end{pmatrix} \tag{400}$$

This procedure can be repeated on the c_1 part, using h and g on next coarser scale. In this case the next matrix is smaller by a factor of four. Repeating the procedure on every level gives

$$\{c_N, d_N, d_{N-1}, d_{N-2}, \cdots d_1\} \tag{401}$$

The relations connecting these with c_0 can be put in the form of matrix. Since both bases are real and orthonormal this is necessarily are real orthogonal matrix. This is the wavelet transform.

The rows and columns are clearly manipulated by similarity transformations at each level.

16 Multigrid Methods

To understand the standard procedure consider a linear equation of the form:

$$Lu = f \tag{402}$$

where L is a linear operator and f is known. Let n denote a wavelet level with higher n denoting a finer grid.
Step 1: Start with a coarse grid - call it the $n = 1$ grid. Project L and f on the basis functions spanning this grid. This gives

$$L^1 u^1 = f^1 \tag{403}$$

In terms of the scaling functions we have

$$L^1_{mn} = \int \phi_m(x) L \phi_n(x) dx \tag{404}$$

$$f^1_m = \int \phi_m(x) f(x) dx \tag{405}$$

Solve this equation exactly for u^1. This gives a first approximation on the next level. In the case of wavelets this is the smooth part of the solution.
Step 2: The next step is to go to the next level. Define L^2 and f^2 by

$$L^2_{mn} = \int D^{-1}\phi_m(x) L D^{-1}\phi_n(x) dx \tag{406}$$

$$f^2_m = \int D^{-1}\phi_m(x) f(x) dx \tag{407}$$

We can map the course solution to the fine grid in two different ways. The first is simply to expand

$$\tilde{u}^2 = \sum_m (\tilde{u}^1, D^{-1}T^m \phi)\phi_{-1,m}(x) \tag{408}$$

The expansion coefficients $(\tilde{u}^1, D^{-1}T^m\phi)$ cam be expressed in terms of the scaling coefficients:

$$(\tilde{u}^1, D^{-1}T^m\phi) = \sum_n (D\phi, T^{m-2n}\phi) u_n = \sum_n (D\phi, T^{m-2n}\phi) u_n = \sum_n h_{m-2n} u_n \tag{409}$$

which gives

$$\tilde{u}^2 = \sum_{m,n} \phi_{-1,m}(x) h_{m-2n} \tag{410}$$

While this expresses the approximation in terms of the basis on the next level, it makes more sense to use the mother functions on the previous level, since they deal with the high frequency information. In this case we have the equivalent expression:

$$\tilde{u}^2 = \sum_m u_m \phi_{0,m}(x) + \sum_m v_m M_{0,m}(x) \tag{411}$$

where all of the coefficients $v_m = 0$.

On the next level we include the mothers. We note that the projected equation on the next level is

$$L^2 u^2 = f^2 \tag{412}$$

If we use the approximate solution $\tilde{2}$, which is exact on the previous level, we get

$$d_1^2 = L^2 \tilde{u}^2 - f^2 \tag{413}$$

which is called the defect. This measure how much the coarse grid solution fails to satisfy the equation at the next finest level.

We can also compute the error

$$v^2 = u^2 - \tilde{u}^2 \tag{414}$$

Knowing v^2 is equivalent to knowing u^2. We also have

$$L^2 v^2 = L^2(u^2 - \tilde{u}^2) = -d_1^2 \tag{415}$$

which gives an equation for the error in terms of the defect.

The idea is to map this up to the coarse level and solve for an approximate $\tilde{v}_1^2$. Projecting on the coarse level does not help, because d_2 is orthogonal to the coarse subspace; however, we can use the mothers - which is a space of the same size and project this equation on the mother subspace. This gives a correction; $\tilde{v}^2$ that is orthogonal to $\tilde{u}^2$

Define a new $\tilde{u}_2^2$ by

$$\tilde{u}_2^2 = \tilde{u}_1^2 + \tilde{v}_1^2 \tag{416}$$

This has a part on the coarse father space and a part on the coarse mother space.

Compute a new defect

$$d_2^2 = L^2 \tilde{u}_2^2 - f^2 \tag{417}$$

this has parts on the coarse father and mother subspaces.

We have the relation

$$L^2 \tilde{v}_2^2 = d_2^2 \tag{418}$$

We seek an approximation on the father space by projection - this gives a correction $\tilde{v}_2^2$ on the father space, and a new approximate solution

$$\tilde{u}_3^2 = \tilde{u}_2^2 + \tilde{v}_2^2 = \tilde{u}_1^2 + \tilde{v}_1^2 + \tilde{v}_2^2 \tag{419}$$

This process can be repeated, alternating between the mother and father space until a solution of the desired accuracy is obtained.

Step 3. The next step uses this solution to go to the next level. The above process needs to be used for each solution on the previous level.

The advantage is that if the coupling between the different scales is small the convergence should be fast.

17 Quadrature Methods

The problem is to compute the overlap coefficients

$$\nu_{jk} := (\phi_{jk}, f) = (D^j T^k \phi, f) \tag{420}$$

and

$$\mu_{jk} := (\psi_{jk}, f) = (D^j T^k \psi, f) \tag{421}$$

The scaling equations relate coefficients on one scale to the coefficients on a finer or coarser scale.

We use the basic scaling equations:

$$D\phi = \sum_l h_l T^l \phi \tag{422}$$

and

$$D\psi = \sum_l g_l T^l \phi \tag{423}$$

and the commutation relations

$$DT = T^2 D \tag{424}$$

In addition, the multiresolution analysis implies the inverse relations:

$$\phi_l = \sum_k (h_{l-2k} DT^k \phi + g_{l-2k} DT^k \psi) \tag{425}$$

(these following from completeness and orthogonality)

The scaling equations give

$$\nu_{jk} = (D^j T^k \phi, f) = (D^{j-1} T^{2k} D\phi, f) = \sum_l h_l \nu_{j-1,2k+l} = \sum_m h_{m-2k} \nu_{j-1,m} \tag{426}$$

and similarly

$$\mu_{jk} = (D^j T^k \psi, f) = (D^{j-1} T^{2k} D\psi, f) = \sum_l g_l \nu_{j-1,2k+l} = \sum_m g_{m-2k} \nu_{j-1,m} \tag{427}$$

The inverse relations give

$$\nu_{j-i,l} = \sum_k (h_{l-2k} \nu_{jk} + g_{l-2k} \mu_{jk}). \tag{428}$$

To understand the one-point quadrature rule note

$$\int f(x)\phi(x)dx = \nu \tag{429}$$

Define

$$x_1 := \int x\phi(x)dx = M_1 \tag{430}$$

Then

$$\int (a+bx)\phi(x)dx = a + bM_1 = a + bx_1 \tag{431}$$

For orthogonal wavelets note

$$k_m := \int x\phi(x)\phi(x-m) = \int (y+m)\phi(y+m)\phi(y) = \int x\phi(x+m)\phi(x) = k_{-m}$$

It follows that

$$\sum mk_m = \sum(-m)k_m = 0 \tag{432}$$

Since

$$\sum m\phi(x-m) = x - M_1 \tag{433}$$

we have

$$0 = \sum \int mx\phi(x)\phi(x-m) = \sum s\phi(x)x(x-M_1) = M_2 - M_1^2 \tag{434}$$

This means that

$$\int \phi(x)(a+bx+cx^2) = a + bM_1 + cM_2 = a + bx_1 + cx_1^2 \tag{435}$$

or that for $x_1 = M_1$ the one-point rule integrates polynomials of degree 2 exactly.

Coiflets have the property that

$$\int x^k \phi(x)dx = 0 \tag{436}$$

for $k = 1, 2, \cdots N$. In this case for $x_1 = M_1 = 0$ we have

$$\int \phi(x) \sum_{n=1}^{N} c_n x^n = \sum_{n=1}^{N} c_n x_1^n = c_0 \tag{437}$$

which integrates polynomials of degree N exactly.

For orthogonal wavelets the error will be of the order 2^{-3n} where 2^{-n} is the finest scale wavelet.

The most useful alternative is to use a multipoint formula. These are are not limited in terms of errors and still allow wavelets with small support.

Choose wavelets that have compact support on $[0, L]$. In general we need to calculate

$$I_{mn}[f] = \int \phi_{mn}(x)f(x)dx \tag{438}$$

where $\phi(x)$ is a scaling function. We also need

$$\hat{I}_{mn}[f] = \int \psi_{mn}(x)f(x)dx \tag{439}$$

where $\psi_{mn}(x)$ is a wavelet.

First observe that if we work with wavelets on the finest scale the scaling equations map to coarser scales. Thus we have

$$I_{mn}[f] = \int D^m T^n \phi(x) f(x) dx =$$

$$\sum_l h_l \int D^{m-1} T^{2n+l} \phi(x) f(x) dx =$$

$$\sum_l h_l I_{m-1,2n+l}[f]$$

and

$$\hat{I}_{mn}[f] = \int \psi_{mn}(x) f(x) dx =$$

$$\sum_l g_l I_{m-1,2n+l}[f]$$

This can be repeated by reapplying the scaling relations $m-1$ more times to express these integrals in terms of the integrals

$$I_{0l}[f]$$

which are translates of the scaling function on the finest scale. This is done using the wavelet transform.

We seek a quadrature rule of the form

$$I[f] = \sum_{n=1}^{N} w_n f(x_n) \tag{440}$$

where the points x_n are in the interval $[0, L]$ and the formula is exact for polynomials

$$I[x^m] = \sum_{n=1}^{N} w_n x_n^m \tag{441}$$

We also note that

$$I_{0,1}[f] = I[T^{-1} f] \tag{442}$$

This means that

$$I_{0,1}[x^m] = I[(x+1)^m] = \sum_{n=1}^{N} w_n (x_n + 1)^m \tag{443}$$

or more generally

$$I_{0,k}[x^m] = I[(x+k)^m] = \sum_{n=1}^{N} w_n (x_n + k)^m \tag{444}$$

which will also be exact. Thus, for a general function we have the approximations:

$$I_{0,k}[f] = \sum_{n=1}^{N} w_n f(x_n + k) \tag{445}$$

Thus, having points and weights for the scaling function provide a means to compute all overlap integrals on the finest scale. The scaling equation for the scaling function and wavelets allow one to get to the overlap integrals for the wavelet basis.

Two problems remain - they are the computation of the moments of the scaling function and the computation of the quadrature weights.

For stability it is useful to replace the system

$$I[x^m] = \sum_{n=1}^{N} w_n x_n^m \tag{446}$$

by

$$I[P_m] = \sum_{n=1}^{N} w_n P_m(x_n) \tag{447}$$

where P_n is a system of real orthogonal polynomials on the support of the scaling function. For a scaling function with support in $[0, L]$ we assume

$$\int_0^L P_m(x) P_n(x) w(x) dx = \delta_{mn} \tag{448}$$

For a polynomial naturally supported on $y \in [-1, 1]$ let

$$y = \frac{2}{L} x - 1 \tag{449}$$

If $T_n(y)$ are orthogonal on $[-1, 1]$; i.e.

$$\int_{-1}^{1} T_m(y) T_n(y) s(y) dy = \delta_{mn} \tag{450}$$

then

$$P_n(x) = T_m(y(x))\sqrt{\frac{dy}{dx}} = \sqrt{\frac{2}{L}}T_m(\frac{2}{L}x - 1) \tag{451}$$

with weight

$$w(x) = s(y(x)) = s(\frac{2}{L}x - 1) \tag{452}$$

The main reason for including the weight is that the Chebyshev polynomials have a non-trivial weight.

Because of the structure of the wavelets it is useful to pick N equally spaced quadrature points on $[0, L]$.

$$x_n = (n-1)2^s + \tau \tag{453}$$

which go from

$$x_1 = \tau \qquad \text{to} \qquad x_N = (N-1)2^s + \tau < L \tag{454}$$

While τ can be used as an adjustable parameter to increase the order of the quadrature, it seem like a better strategy is to simply increase N. In this case the problem is to solve the linear system

$$I[P_m] = \sum_{n=1}^{N} w_n P_m(x_n) \tag{455}$$

for the weights w_n. The relevant approximate quadrature is then

$$I[f] \sim \sum_{n=1}^{N} w_n f(x_n) \tag{456}$$

which is exact for polynomials of degree $\leq N-1$.

In order to solve these equations we need expression for the moments $I[P_m]$. To do this we use scaling and unitarity. We need the following two sets of coefficients

$$DP_n(x) = \frac{1}{\sqrt{2}}P_n(\frac{x}{2}) = \sum_{m=0}^{n} d_{nm}P_m(x) \tag{457}$$

$$T^{-l}P_n(x) = P_n(x+l) = \sum_{m=0}^{n} t^l_{nm}P_m(x) \tag{458}$$

We can get exact expressions for the matrices d_{mn} and t^l_{mn} using the appropriate gauss quadrature formula:

$$\int_0^L P_n(x)P_m(x)w(x)dx = \delta_{mn} = \sum_{i=1}^{K} P_n(u_i)P_m(u_i)w_{ui} \tag{459}$$

where $K > n/2, m/2$. Multiplying the above equations by $P_k(x)w(x)$ and using the quadrature rule gives:

$$d_{nk} = \sum_{i=1}^{K} w_{ui}\frac{1}{\sqrt{2}}P_n(\frac{u_i}{2})P_k(u_i) \tag{460}$$

$$t^l_{nk} = \sum_{i=1}^{K} w_{ui}P_n((u_i + l))P_k(u_i) \tag{461}$$

To compute the moments assuming $P_0(x) = c$ we have

$$I[P_0] = \int_0^L \phi(x)P_0(x)dx = c \tag{462}$$

Note

$$I[P_n] = \int_0^L \phi(x)P_n(x)dx =$$

$$\int_0^L D\phi(x)DP_n(x)dx =$$

$$\sum_l h_l \sum_m d_{nm} \int_0^L T^l\phi(x)P_m(x)dx =$$

$$\sum_l h_l \sum_m d_{nm} \int_0^L \phi(x)T^{-l}P_m(x)dx =$$

$$\sum_l h_l \sum_{mk} d_{nm}t^l_{mk}I[P_k]$$

We can separate off the $k = n$ term and write

$$I[P_n] = \frac{\sum_l h_l \sum_{mk}^{k<n} d_{nm}t^l_{mk}I[P_k]}{1 - \sum_l h_l \sum_{mk} d_{nm}t^l_{mn}} \tag{463}$$

Which can be used to recursively generate the required moments in terms of $I[P_0] = c$.

The advantage of the Chebyshev polynomials are that the points and weights are known analytically. We have the formulas

$$\int_{-1}^{l} f(x) \frac{dx}{\sqrt{1-x^2}} \approx \sum_{n=1}^{N} \frac{\pi}{N} f(\cos(\frac{(2n-1)\pi}{2N})) \tag{464}$$

which are exact for $f(x)$ a polynomial of degree $2N-1$. To get the u_i and w_{ui} these expressions have to be transformed from $[-1, 1]$ to $[0, L]$:

$$u_k = \frac{L}{2}(\cos(\frac{(2n-1)\pi}{2N} + 1) \tag{465}$$

and

$$w_{uk} = \frac{2\pi}{LN} \tag{466}$$