Chapter 2

Wavelets – Basic Theory and Construction

Phase space analysis is any useful decomposition of a function into modes that are well-localized and – at the same time – have a small spread in momentum. The properties of the Fourier transform impose limitations on such a decomposition at the very outset, but the need for phase space analysis in quantum field theory has already been made clear in Chap. 1. It is important in other areas as well – e.g., partial differential equations and signal analysis. One limitation is that a mode cannot be sharply localized in both momentum and position – i.e., no compactly supported function has a compactly supported Fourier transform. Another limitation is that – while exponential localization in both momentum, and position is possible – a small standard deviation in momentum (resp. position) creates a lower bound on the standard deviation in position (resp. momentum) by the uncertainty principle. Further limitations are inevitable if one desires the set of modes (expansion functions) to be coherent in some sense – e.g., related to one another through group operations.

Perhaps the oldest type of phase space analysis is the *discrete windowed Fourier transform* – an analysis based on expansion functions of the form

$$f_{\vec{m}\vec{n}}(\vec{x}) = e^{i2\pi\vec{m}\cdot\vec{x}} f(\vec{x} - \vec{n}) \tag{2.1}$$

with $\vec{m}$ and $\vec{n}$ ranging over cubic lattice sites in d-dimensional space with integer coordinates. These functions are coherent in the sense that they are generated by discrete translates in $2d$-dimensional phase space with f as the generating function. Naturally, the regularity and decay properties of f are an important issue. In particular, if f is the characteristic function of the unit cube $[0,1]^d$, then each expansion function is sharply localized, but discontinuous. Moreover, $\{f_{\vec{m}\vec{n}}\}$ is an orthonormal basis for the Hilbert space $L^2(\mathbb{R}^d)$ in this case; indeed, the decomposition is just a cubic partition of $\mathbb{R}^d$ with Fourier series analysis applied on each cube. The jump discontinuities are the obvious drawback, as they correspond to poor localization in momentum space. This implies that singularities in the function analyzed are not reflected by the $(\vec{m}, \vec{n})$-dependence of the expansion coefficients. Can a more regular f be found such that $\{f_{\vec{m}\vec{n}}\}$ is still an orthonormal basis? The answer is yes, but only by sacrificing localization; the trade-off is very strict, as it turns out. It has been shown by Balian [B9] and also by Low [L47] that if $\{f_{\vec{m}\vec{n}}\}$ is an orthonormal basis, then the Heisenberg uncertainty of f must be infinite! This means that no orthonormal basis of this type can be much better than the one generated by the characteristic function of the unit cube. The common practice in phase space localization of this type is to sacrifice orthogonality to obtain more regularity for f. The obvious constraint is that the calculation of expansion coefficients for a given function must still be realistic. One of the earliest proposals – independently examined by Gabor [G1] and von Neumann [unpublished] – was to choose f as a Gaussian function peaked at the center of the unit cube – a function with the best possible phase space localization. The price of this analysis was immediately evident in the over-completeness of the set $\{f_{\vec{m}\vec{n}}\}$ generated. Expansion coefficients were not unique for an arbitrary function decomposed in this way. Indeed, such an expansion was known to be numerically unstable in the sense that the sum of squares of inner products of these functions with the function to be analyzed could be arbitrarily small relative to the L^2-norm of the given function. On the other hand – since this set is over-complete anyway – one can enhance this over-completeness by generating a set with phase space translates smaller than unity in the spatial directions or smaller than 2π in the momentum directions. Expansions of this type are numerically stable, and they have been exploited in signal analysis despite the computational inconvenience due to the redundancy (see, e.g., Baastians [B1] and Janssen [J8–J10]).

This decomposition scheme is an example of a *frame* – a certain generalization of orthonormal basis where the elements are not even necessarily linearly independent. Frames were first studied in the context of nonharmonic analysis by Duffin and Schaeffer [D28] and can be defined for an abstract Hilbert space. The frame property is a generalization of Parseval's identity to a double inequality for the sum of squares of inner products, where the upper and lower bounds are proportional to the square of the norm of the test vector. From a computational point of view, frame expansions are "painless non-orthogonal expansions," and their usefulness to phase space localization in signal analysis was widely advertised by Daubechies, Grossmann, and Meyer [D4]. They even constructed frames generated by α-translates in position space and translates $< 2\pi/\alpha$ in momentum space where the generating function f is class C^∞ with compact support. Even as they established impressive results for the old-fashioned windowed Fourier transform, Daubechies, Grossmann, and Meyer were already engaged in the development of a rather different type of phase space analysis based on scaling – namely, *wavelet analysis.*

The meaning of the word "wavelet" has varying degrees of generality in the literature, but the coherence of a wavelet basis or wavelet frame usually involves scaling operations. Indeed, the orthogonality of discrete scalings is a property that shall be included in our own definition of the word, so there are results on wavelet frames that lie outside the scope of our concerns. For our purposes, a *dyadic wavelet* is a square-integrable function Ψ such that

$$\int \Psi(\vec{x})^* \Psi(2^r \vec{x} - \vec{n}) d\vec{x} = 0, \qquad \vec{n} \in \mathbb{Z}^d, \tag{2.2}$$

for all non-zero integers r. Note that this definition does not require orthogonality of lattice translates on the same scale or completeness of the basis generated. Negative values of r correspond to dyadic scalings of Ψ to large length scales, while positive values of r yield small length scales. Throughout most of this chapter we concentrate on dimension $d = 1$ (the dimension appropriate to signal analysis in any case) and extend the wavelet constructions to arbitrary dimension in the very last section. This decision is due, in part, to the ease with which the L^2-orthogonal constructions extend to the multi-dimensional case. On the other hand, we shall also construct Sobolev-orthogonal wavelets in that section. Obtaining well-behaved wavelets of this type is more involved because the coherent structure is rectangular while the Sobolev norm is elliptic.

The earliest example of an orthonormal wavelet basis was constructed by Haar [H6] decades before the subject was born. It is a dyadic scale hierarchy – including arbitrarily large and arbitrarily small length scales – of piecewise constant functions whose integrals vanish. On the same scale, the discrete translates are orthogonal because they have disjoint supports, while functions on differing scales are orthogonal because the smaller-scale function is supported on some rectangle over which the larger-scale function is constant. This type of decomposition has the same regularity and decay properties as the windowed Fourier decomposition with the characteristic function – the functions are discontinuous, but their localization is sharp. The Haar basis was originally introduced in the context of a general study of orthogonal systems, and at the time, no one seemed to have any curiosity about the existence of similar bases of more regular functions. The special features of the construction may have led some mathematicians to suspect that nothing better than the Haar basis was possible.

Since that time, ideas involving scaling transformations have slowly emerged in other areas. In statistical mechanics, for example, Wilson introduced the renormalization group [W16] as a means of isolating the long-distance behavior of the generalized Ising model. The basic transformation averages spin values over a block of sites with the cubic lattice partitioned into such blocks and the new configuration defined by scaling the coarser lattice back to the unit lattice. The Gibbs measure itself is transformed by integrating out the fluctuations and applying the additional transformation induced by the scaling. Iteration of this transformation has the effect of reducing the long-distance decay of correlation functions to behavior at large length scales. Moreover, in his effort to justify a rather drastic approximation aimed at obtaining qualitative results, Wilson assumed the existence of a wavelet basis [W17].

At roughly the same time, Glimm and Jaffe were developing phase cell methods in their effort to construct the pure scalar quantum field theory with quartic self-interaction in three space-time dimensions [G45, G57]. We have already discussed this model in Chap. 1, where we advertised the need for phase space analysis. While their decomposition of phase space into phase cells was primitive and their expansion method complex, the convergence proof contained several ingenious arguments that involved scaling.

In signal analysis, the use of scaling as an analytic tool was first proposed by Grossmann and Morlet [G96], inspired in part by the work of Morlet et al. [M48]. Specifically, they applied the *wavelet transform* to the analysis of seismic data in geophysics, using a basic signal function $f(t)$. The one-dimensional wavelet transform Ψ_ξ of an arbitrary function $\xi(t)$ is given by

$$\Psi_\xi(a,s) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} dt\, f(a^{-1}(t-s))^* \xi(t), \qquad a \neq 0, \tag{2.3}$$

with a fixed choice of $f(t)$. Subsequently, Grossmann, Morlet, and Paul studied this integral transform as a square-integrable representation of the affine group [G97, G98]. The resolution of the identity (i.e., the inversion formula) is given by

$$\xi(t) = \frac{1}{c(f)} \int_{-\infty}^{\infty} da\, |a|^{-5/2} \int_{-\infty}^{\infty} ds\, f(a^{-1}(t-s)) \Psi_\xi(a,s), \tag{2.4}$$

$$c(f) = \int_{-\infty}^{\infty} d\omega |\omega|^{-1} |\hat{f}(\omega)|^2. \tag{2.5}$$

The finiteness of the constant $c(f)$ is a condition on the basic signal $f(t)$, which is regarded as a "wavelet" in this context. Note that if $\hat{f}(\omega)$ is continuous, then $\hat{f}(0) = 0$. This weak vanishing-moment condition is satisfied by wavelets in virtually every context, independently of how they are defined.

Meanwhile, Stromberg constructed a one-dimensional orthonormal basis of wavelets more regular than the Haar wavelets [S85]. Inspired by the construction of an orthonormal basis for $L^2([0,1])$ due to Franklin [F53] – which is more regular than the Haar basis but has little wavelet coherence – Stromberg derived a piecewise linear, continuous function with exponential decay that generates an orthonormal basis for $L^2(\mathbb{R})$ through dyadic scaling and discrete scale-commensurate translation. This achievement attracted little attention at the time, as it occurred in the context of harmonic analysis, and interdisciplinary communication failed in this instance.

In mathematical physics at about the same time, Gawedzki and Kupiainen were engaged in establishing a sound mathematical basis for the renormalization group ideas of Wilson. For the transformation they rigorously determined the flow of the iteration in the neighborhood of the Gaussian fixed point for the dipole gas [G16, G25]. The type of phase space localization used for the definition of the renormalization group transformation varies with the context, but in the case where the transformation is based on block spin averaging, the constrained minimization technique for integrating out the fluctuations can be adapted as a method for constructing wavelets (see, e.g., [B12, B13]). Concurrent to the work of Gawedzki and Kupiainen was the work of Federbush on a class of infrared problems [F10]. Instead of using the renormalization group formalism, however, he developed a phase cell cluster expansion in much the same spirit as Glimm and Jaffe, except that his phase cells were expansion functions. Not only was his orthonormal basis scale-coherent, but it was a polynomial generalization of the Haar basis! His construction combined the Haar scheme with the homogeneity of monomials to yield a coherent basis of piecewise-polynomial functions with as many vanishing moments as one wished. The vanishing moments directly implied the interscale-orthogonality but were most important to the multi-scale decomposition of long-distance singularities. On the other hand, short-distance number divergences in the cluster expansion of a Euclidean field theory could not be cancelled in this phase cell formalism for a technical reason related to the jump discontinuities in these wavelets, so the rigorous treatment of ultraviolet problems in this formalism had to wait on the discovery of smoother wavelets [B12. B13]. The Stromberg construction was unknown to the mathematical physics community, but the Stromberg wavelet would have required modification, as it does not have an abundance of vanishing moments.

Interest in wavelets finally quickened when Meyer constructed an orthonormal basis of wavelets that were smooth and yet of rapid decrease [M40]. Indeed, the Fourier transform of a Meyer wavelet is class C^∞ with compact support! On the other hand, if a function has compact support in momentum space, it cannot have exponential decay in position space – only the rapid decrease, at best. Any basis of wavelets that are class C^N for conveniently chosen N and have exponential decay would be more useful to mathematical physics in general. Subsequent to the Meyer construction, Lemarié found such orthonormal bases – one for each degree N of smoothness [L22]. Mallat and Meyer almost immediately realized that the construction of a variety of wavelets could be organized into a scheme which they dubbed *multi-scale resolution analysis* [M41, M42]. Any input function for this constructive machine is called a *scaling function*, and its regularity and decay properties are shared by the wavelet produced. This point of view was central to the decomposition and reconstruction algorithm of Mallat [M8–M10], and that algorithm, in turn, eventually led to Daubechies' discovery of compactly supported wavelets with a finite but arbitrary degree of smoothness [D1].

As in the case of windowed Fourier analysis, there was immediate interest in the advantages gained by sacrificing orthogonality in favor of some weaker property. One idea – already suggested by the constrained minimization technique [B12] – was to sacrifice intrascale orthogonality while retaining interscale orthogonality. In fact, the variational approach differs from the multi-scale resolution analysis approach, since wavelets that are only interscale-orthogonal are derived first. In the Lemarié case, these wavelets are more explicit than those comprising the orthonormal basis. At the same time, the realization of the dual basis – necessary for the calculation of expansion coefficients – is equally explicit because the interscale orthogonality reduces it to a single-scale derivation (see, e.g., [B17] and [C10]). Chui and Wang subsequently constructed a basis of this type where the wavelets are compactly supported splines [C11, C12]. This is a computational advantage over Daubechies wavelets, which are not splines. On the other hand, the wavelets dual to the Chui–Wang wavelets do not have compact support.

Another industry was the construction of wavelet frames, which we shall not pursue. However, they have had great impact on signal analysis, and it is worth remarking that Daubechies, Grossmann, and Meyer were the founders [D4]. Yet another generalization of an orthonormal basis is a bi-orthonormal basis, and the multi-scale resolution analysis for constructing bi-orthonormal wavelet bases was introduced by Cohen, Daubechies, and Feauveau [C17]. In particular, they constructed a bi-orthonormal basis of compactly supported class C^N wavelets that are more amenable to decomposition and reconstruction algorithms than are the Daubechies wavelets. We shall not pursue this construction either.

The aim of this chapter is an introduction to wavelets suitable for Euclidean field theory and to rigorous results on the properties of wavelets. In the first section we discuss the limitations of the discrete windowed Fourier transform, including the Balian–Low Theorem. In the following section we introduce the multi-scale resolution analysis of Mallat and Meyer, and we apply it to the construction of Meyer wavelets and Lemarié wavelets. We devote the third section to a sketch of the Daubechies construction of compactly supported wavelets. Actually, in mathematical physics there never seems to be any need for compactly supported wavelets; exponentially localized wavelets are always good enough for our purposes. We describe the construction of Daubechies wavelets because it is the crown jewel in wavelet analysis. The succeeding section is devoted to the important issue of vanishing moments. We show how the condition of interscale orthogonality implies that the number of vanishing moments for a wavelet is comparable to its degree of smoothness. This relationship, in turn, implies that interscale-orthogonal wavelets – functions that meet our own definition of "wavelet" – cannot be exponentially localized in both position space and momentum space. In §2.5 and §2.6 we investigate further theoretical restrictions on wavelets in the form of Heisenberg inequalities. In particular, the position-momentum uncertainty of a real-valued wavelet in one dimension is $3/2$ instead of the universal $1/2$. In §2.7 we describe the constrained minimization method for constructing wavelets. This variational approach can be used to construct a variety of wavelets – including Meyer wavelets and Lemarié wavelets – but it cannot produce Daubechies wavelets. Following this point of view, §2.8 is devoted to a description of the Lemarié-type basis that is not intrascale-orthogonal. For the mother wavelet there is an explicit expression depending on the zeros of Euler–Frobenious polynomials in the unit disk. We devote §2.9 to a description of the Chui–Wang wavelets. Finally, we introduce the extension of the wavelet constructions to the multi-dimensional case in §2.10. At the same time, we construct exponentially localized Sobolev-orthonormal wavelet bases, using the variational method. The multiscale resolution analysis, which is suited for the construction of Daubechies wavelets, does not appear to be suited for the construction of Sobolev-orthogonal wavelets with good localization. On the other hand, Lemarié has shown that it can be generalized to construct a Sobolev-bi-orthonormal system of compactly supported wavelets [L26].

2.1 The Balian–Low Theorem

In the search for orthonormal bases of $L^2(\mathbb{R})$ that localize phase space, perhaps the most obvious type to try is a basis of the form

$$\mathcal{B} = \{e^{i2\pi mx} f(x-n) \colon m, n \in \mathbb{Z}\}, \tag{2.1.1}$$

where f is a square-integrable function. The example that comes to mind most readily is $f = \chi$, where χ is the characteristic function of the unit interval [0,1]. In this case, one is simply applying the orthogonal decomposition

$$L^2(\mathbb{R}) = \bigoplus_{n=-\infty}^{\infty} L^2([n, n+1]) \tag{2.1.2}$$

to the Hilbert space and then using Fourier series analysis on each interval. Unfortunately, χ is a discontinuous function, which means its Fourier transform has poor localization. The latter is regular but not integrable. Indeed,

$$\widehat{\chi}(p) = \frac{e^{-ip} - 1}{-ip}, \tag{2.1.3}$$

where $p = 0$ is a removable singularity. It is natural to wonder if an orthonormal basis of the above form is possible for a function f that has better localization in phase space. Roughly speaking, one would be partitioning phase space into "phase cells" with a rectangular lattice structure given by Fig. 2.1.1. The basis decomposition of an arbitrary function would yield its spread over those phase cells, where the area of each phase cell is 2π. Obviously, sharp localization of a basis function in its associated phase cell is impossible, as no function can have compact support in both position space and momentum space. However, a Gaussian function has excellent decay in both position and momentum. Can such a function generate any kind of basis with these phase space translates? Obviously, a set of the form

$$\mathcal{N} = \{(2a/\pi)^{1/4} e^{-a(x-x_0-n)^2} e^{i2\pi m x} \colon \ m, n \in \mathbb{Z}\} \tag{2.1.4}$$

cannot be orthonormal, but it does span $L^2(\mathbb{R})$; indeed it is over-complete. This set has been called the *von Neumann basis* in the physics community and the *Gabor basis* in the mathematics community. In spite of its drawbacks, this "basis" has actually played a role in the understanding of phase space localization. In particular, the study of this set led to some intuition that resulted in the *Balian–Low Theorem*. This result states that if a square-integrable function f generates an orthonormal basis of the form (2.1.1), then either $f'(x)$ or $xf(x)$ fails to be square-integrable. The original proof was actually quite ingenious: it was a topological argument involving the winding number of the phase of an expression. There is now a more elementary proof available, which we present here.

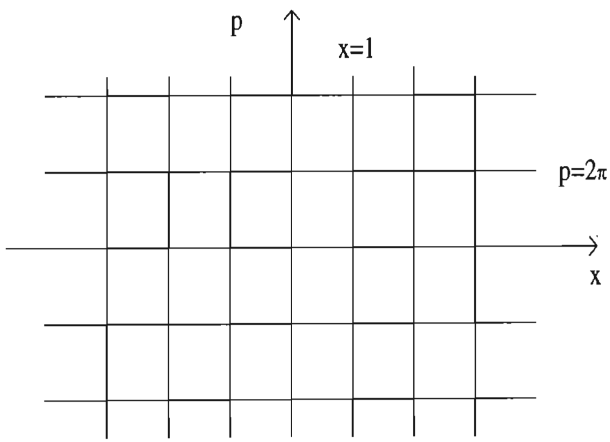

Figure 2.1.1:

Suppose that Pf and Xf both lie in $L^2(\mathbb{R})$. Let T_{mn} denote the phase space translation operator defined by

$$(T_{mn}\varphi)(x) = e^{i2\pi mx}\varphi(x-n). \tag{2.1.5}$$

By hypothesis, $\{T_{mn}f\colon\ m,n \in \mathbb{Z}\}$ is an orthonormal basis, so by Parseval's identity,

$$(Pf, Xf) = \sum_{m,n}(Pf, T_{mn}f)(T_{mn}f, Xf). \tag{2.1.6}$$

On the other hand,

$$\begin{aligned}(f,[P,T_{mn}]f) &= 2\pi m(f,T_{mn}f) = 0, &(2.1.7)\\ (f[T_{mn},X]f) &= -n(f,T_{mn}f) = 0, &(2.1.8)\end{aligned}$$

so the expansion may also be written as

$$(Pf, Xf) = \sum_{m,n}(T_{-m,-n}f, Pf)(Xf, T_{-m,-n}f), \tag{2.1.9}$$

which – again by Parseval's identity – implies

$$(Pf, Xf) = (Xf, Pf). \tag{2.1.10}$$

Since $[X,P] = i$, we have the desired contradiction.

Actually, one can get around this difficulty if some of the discrete translational symmetry in phase space is sacrificed. Orthonormal bases of the form

$$\begin{aligned}\mathcal{B} \;=\; &\{\sin(2\pi mx)f(x-n)\colon\ n \in \mathbb{Z}, m \in \mathbb{Z}^+\} \cup\\ &\{\cos(2\pi mx)f(x-n)\colon\ n \in \mathbb{Z}, m \in \mathbb{Z}^+ \cup \{0\}\}\end{aligned} \tag{2.1.11}$$

have been constructed where f and $\hat{f}$ both have exponential decay. Such a basis is a *Daubechies-Jaffard-Journé basis*. Obviously, a given basis element is roughly localized on a phase cell defined by the intersection of a set of the form

$$\{(x,p)\colon\ n \le x \le n+1\}$$

with a set of the form

$$\{(x,p)\colon\ 2\pi m - \pi \le p \le 2\pi m + \pi \text{ or } -2\pi m - \pi \le p \le -2\pi m + \pi\}.$$

If $m \neq 0$, the cell consists of two disjoint rectangles, each with area 2π, and there are two basis elements associated to this disconnected cell. If $m = 0$, the cell is a single rectangle, and it has only one basis element. The two separated peaks in momentum space are what make such a tremendous difference in the phase space localization. It has been shown that if one insists on an orthonormal basis of the form

$$\mathcal{B} = \{f_m(x-n)\colon\ m,n \in \mathbb{Z}\} \tag{2.1.12}$$

with $\hat{f}_m$ localized roughly at $2\pi m$ (but not necessarily a momentum-space translate of $\hat{f}_0$ by $2\pi m$ units), then one can beat the conclusion of the Balian–Low Theorem, but only marginally. Xf and Pf can both be square-integrable, but either $|X|^{1+\epsilon}f$ or $|P|^{1+\epsilon}f$ will fail to be.

There has long been an industry in phase space localization where one works with a set of the form

$$\mathcal{B} = \{e^{i2\pi\alpha mx} f(x-\beta n)\colon\ m,n\in\mathbb{Z}\} \tag{2.1.13}$$

with $\alpha\beta < 1$. The inner products of an arbitrary square-integrable function with these functions constitute the discrete version of a *windowed Fourier transform* of that function, and there is a trade-off implicit in the choice $\alpha\beta < 1$. The expansion functions can have better phase space localization, but they cannot be orthonormal. The latter point can be established with an elementary calculation. Indeed, suppose a set of this form is an orthonormal basis. Then for an arbitrary square-integrable function φ,

$$\begin{aligned} \|\varphi\|_2^2 &= \sum_{m,n}\left|\int_{-\infty}^{\infty}\varphi(x)e^{-i2\pi m\alpha x}f(x-\beta n)^* dx\right|^2 \\ &= \alpha^{-1}\sum_{\ell,n}\int_{-\infty}^{\infty}\varphi(x)\varphi(x-\alpha^{-1}\ell)^* f(x-\beta n)^* \\ &\qquad f(x-\beta n-\alpha^{-1}\ell)dx \end{aligned} \tag{2.1.14}$$

by Parseval's identity and Poisson summation. Notice that if the support of φ lies in an interval of length α^{-1}, then only the $\ell = 0$ term remains:

$$\|\varphi\|_2^2 = \alpha^{-1}\sum_n\int_{-\infty}^{\infty}|\varphi(x)|^2|f(x-\beta n)|^2 dx, \quad \text{diam supp } \varphi \le \alpha^{-1} \tag{2.1.15}$$

Since φ is arbitrary otherwise, it follows that

$$\sum_n |f(x-\beta n)|^2 = \alpha. \tag{2.1.16}$$

Applying the integration $\int_0^\beta dx$ to this equation, we obtain

$$\int_{-\infty}^{\infty}|f(x)|^2 dx = \alpha\beta. \tag{2.1.17}$$

But $\|f\|_2 = 1$, so we have established the claim that the inequality $\alpha\beta < 1$ rules out an orthonormal basis. On the other hand, the lack of orthogonality does not necessarily pose computational problems in the phase space analysis of an arbitrary function. The windowed Fourier analysis is done with sets that happen to be *frames* – expansion functions for which the orthogonality property is replaced by a similarity property. We give the definition in the abstract setting.

A given set $\{g_k\}$ of vectors in a Hilbert space is said to be a *frame* if $\|g_k\| = 1$ and there are constants $a_0, a_1 > 0$ such that

$$a_0\|\varphi\|^2 \le \sum_k |(\varphi, g_k)|^2 \le a_1\|\varphi\|^2 \tag{2.1.18}$$

for all vectors φ. The frame is *tight* if one can choose $a_1 = a_0$, in which case we have the identity

$$\sum_k |(\varphi, g_k)|^2 = a_0 \|\varphi\|^2 \tag{2.1.19}$$

In the case $a_0 = 1$, such a frame becomes an orthonormal basis, but in any case, frames have virtually all the nice properties of an orthonormal basis. If $\{g_k\}$ is a frame, then it automatically has a *dual frame* $\{\tilde{g}_k\}$ – not to be confused with a dual basis. In general,

$$(g_j, \tilde{g}_k) \neq \delta_{jk}, \tag{2.1.20}$$

but one now has a pair of Parseval-type identities:

$$\sum_k (\varphi, \tilde{g}_k)(g_k, \varphi') = (\varphi, \varphi'), \tag{2.1.21.0}$$

$$\sum_k (\varphi, g_k)(\tilde{g}_k, \varphi') = (\varphi, \varphi'). \tag{2.1.21.1}$$

Existence is an abstract result, and does not determine whether $\{\tilde{g}_k\}$ has any of the coherent structure that $\{g_k\}$ happens to have.

The point here is that one can construct frames for $L^2(\mathbb{R})$ of the form (2.1.13), and that if $\alpha\beta < 1$, it is possible for f to be class C^∞ with compact support! Even tight frames with this much phase space localization are possible if $\alpha\beta < 1$. As far as the value of the constant a_0 is concerned, the reasoning given by (2.1.14)–(2.1.17) above implies

$$\sum_{m,n} \left| \int \varphi(x) e^{-i2\pi\alpha m x} f(x - \beta n)^* dx \right|^2 = \frac{1}{\alpha\beta} \|\varphi\|_2^2 \tag{2.1.22}$$

for a tight frame. In particular, we may infer that $\alpha\beta = 1$ is not only a necessary condition for orthogonality, but also sufficient in the case where the frame is already tight.

We have briefly discussed the discrete windowed Fourier transform because it is a familiar approach to phase space localization. It is inappropriate for our purposes, however. The problems of constructive quantum field theory fall into two categories – infrared (long-distance) problems and ultraviolet (short-distance) problems. To isolate either the infrared or the ultraviolet behavior of a model, it is more natural to consider different scales of fluctuations. For this reason, we shall use a very different type of phase space decomposition known as a *wavelet decomposition*.

For a fixed positive integer N, a square-integrable function Ψ is an *N-adic wavelet* if and only if

$$\int \Psi(x)^* \Psi(N^r x - n) dx = 0, \qquad n \in \mathbb{Z}, r \in \mathbb{Z}\backslash\{0\}. \tag{2.1.23}$$

We intend to study orthonormal bases of the form

$$\mathcal{B} = \{2^{r/2}\Psi(2^r \cdot - n) \colon n, r \in \mathbb{Z}\}, \tag{2.1.24}$$

where $2^{r/2}$ is the obvious L^2-normalizing factor. Notice, however, that our definition of a wavelet requires neither orthogonality of discrete translates at the $r = 0$ level nor completeness of the set generated – only orthogonality to N-adic scalings and their discrete scale-commensurate translates. For convenience, we sometimes use the notation

$$\Psi_r(x) = 2^{r/2}\Psi(2^r x). \tag{2.1.25}$$

The discrete translates for $\Psi_r(x)$ are therefore $\Psi_r(x - 2^{-r}n)$ with $n \in \mathbb{Z}$. According to the convention we have already adopted, 2^{-r} is the length scale at the rth level, so the positive (resp. negative) values of r correspond to the small (resp. large) length scales.

In the case of a basis of the form (2.1.24), Ψ is referred to as the *mother wavelet.* If Ψ is well-localized in phase space, then with

$$\widehat{\Psi}_r(p) = 2^{-r/2}\widehat{\Psi}(2^{-r}p), \tag{2.1.26}$$

the wavelet basis is represented by the partition of phase space given by Fig. 2.1.2. Many of the wavelets constructed will be real-valued, in which case their phase cells must be symmetric about $p = 0$. On the other hand, the disjoint packing of scaled phase cells separates them from $p = 0$, and so the cells are disconnected as in the case of the Daubechies–Jaffard–Journé basis. Perhaps the most important observation to be made about this admittedly crude picture is the suggestion that

$$\widehat{\Psi}(0) = 0. \tag{2.1.27}$$

This is a vanishing moment property, and all of our wavelets will have it. Indeed, the number of vanishing moments a wavelet has will be important.

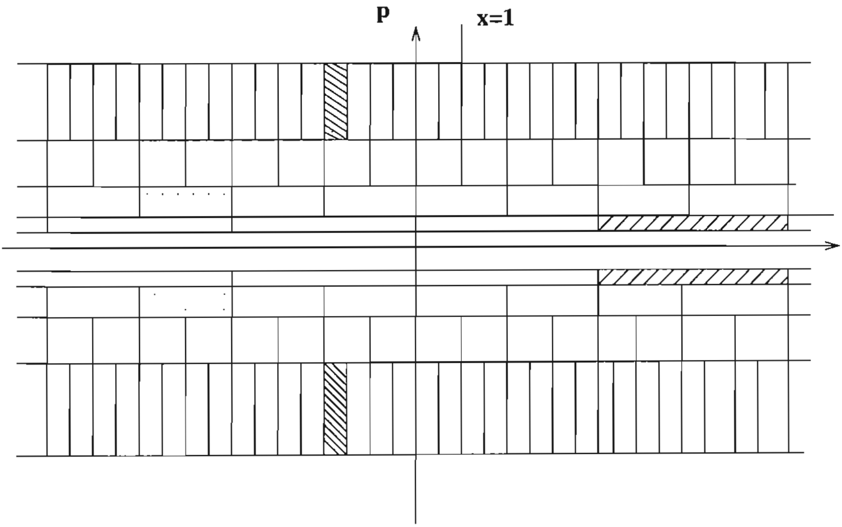

Figure 2.1.2:

An interesting variety of wavelet bases have been constructed only recently, but the oldest example has been around for some time. It is the *Haar basis* (i.e., the "square-wave" basis), and it is defined by

$$\Psi(x) = \begin{cases} 1, & 0 \leq x < \frac{1}{2}, \\ -1, & \frac{1}{2} \leq x \leq 1. \end{cases} \tag{2.1.28}$$

It is easy to convince oneself that

$$\mathcal{B} = \{\Psi_r(\cdot - 2^{-r}n) \colon n, r \in \mathbb{Z}\} \tag{2.1.29}$$

is indeed an orthonormal basis for $L^2(\mathbb{R})$. Also notice that

$$\widehat{\Psi}(0) = \int \Psi(x)dx = 0. \tag{2.1.30}$$

The usefulness of this basis is limited by the poor localization in momentum space due to the jump discontinuities in position space.

References

1. M. Baastians, "Gabor's Signal Expansion and Degrees of Freedom of a Signal," Proc. IEEE **68** (1980), 538–539.

2. R. Balian, "Un Principe d' Incertitude Fort en Théorie du Signal on en Mécanique Quantique," C.R. Acad. Sci. Paris **292**, Série 2 (1981).

3. G. Battle, "Heisenberg Proof of the Balian-Low Theorem," Lett. Math. Phys. **15** (1988), 175–177.

4. J. Bourgain, "A Remark on the Uncertainty Principle for Hilbertian Basis," J. Func. Anal. **79** (1988), 136–143.

5. I. Daubechies, "The Wavelet Transform, Time-Frequency Localization, and Signal Analysis," IEEE Trans. Inform. Theory **36-5** (1990), 961–1005.

6. I. Daubechies, S. Jaffard, and J. Journé, "A Simple Wilson Orthonormal Basis With Exponential Decay," SIAM J. Math. Anal. **22-2** (1991), 555–573.

7. I. Daubechies, A. Grossmann, and Y. Meyer, "Painless Nonorthogonal Expansions," J. Math. Phys. **27-5** (1986), 1271–1283.

8. R. Duffin and A. Schaeffer, "A Class of Nonharmonic Fourier Series," Trans. Amer. Math. Soc. **72** (1952), 341–366.

9. D. Gabor, "Theory of Communications," J. Inst. Elect. Engng. **93** (1946), 429–457.

10. A. Haar, "Zur Theorie der Orthogonalen Funktionensysteme," Math. Ann. **69** (1910), 331–371.

11. A. Janssen, "Gabor Representation of Generalized Functions," J. Math. Appl. **80** (1981), 377–394.

12. A. Janssen, "Bargmann Transform, Zak Transform, and Coherent States," J. Math. Phys. **23** (1982), 720–731.

13. A. Janssen, "The Zak Transform: A Signal Transform for Sampled Time-Continuous Signals," Philips J. Res. **42** (1987).

14. F. Low, "Complete Sets of Wave-Packets," in *A Passion for Physics–Essays in Honor of Geoffrey Chew*, World Scientific, Singapore, 1985, pp. 17–22.

2.2 Lemarié and Meyer Wavelets

We now describe a systematic way to construct an orthonormal wavelet basis known as the *Mallat–Meyer multi-scale resolution analysis*. One starts with a *scaling function* η defined by the following properties:

(a) Let $\mathcal{H}_r$ be the closed span of $\{\eta_r(\cdot - 2^{-r}n)\colon n \in \mathbb{Z}\}$. Then $\bigcap_r \mathcal{H}_r$ is the null space and $\operatorname{span}(\bigcup_r \mathcal{H}_r)$ is dense in $L^2(\mathbb{R})$.

(b) There is a summable sequence $\{c_m\}_{m\in\mathbb{Z}}$ such that

$$\eta(x) = \sum_m c_m \eta(2x - m). \tag{2.2.1}$$

This condition implies $\mathcal{H}_r \subset \mathcal{H}_{r+1}$ and $\operatorname{span}\big(\bigcup_r \mathcal{H}_r\big) = \bigcup_r \mathcal{H}_r$.

(c) $\hat{\eta}$ is continuous and $\sum_{\ell\in\mathbb{Z}} |\hat{\eta}(p + 2\pi\ell)|^2$ has no zeros.

It is easy to verify that the characteristic function χ of the unit interval [0,1] is a scaling function. Indeed,

$$\sum_{\ell\in\mathbb{Z}} |\widehat{\chi}(p + 2\pi\ell)|^2 = 1, \tag{2.2.2}$$

and $c_m = \delta_{0m} + \delta_{1m}$ in this case.

Once the notion of a scaling function has been isolated, the construction of an orthonormal wavelet basis becomes easy. First, define the function φ by

$$\widehat{\varphi}(p) = \left(\sum_{\ell\in\mathbb{Z}} |\hat{\eta}(p + 2\pi\ell)|^2\right)^{-1/2} \hat{\eta}(p), \tag{2.2.3}$$

and observe that φ satisfies condition (a). Obviously,

$$\sum_{\ell\in\mathbb{Z}} |\widehat{\varphi}(p + 2\pi\ell)|^2 = 1, \tag{2.2.4}$$

and therefore

$$\int \varphi(x - m)^* \varphi(x - n)dx = \delta_{mn}, \qquad m, n \in \mathbb{Z}. \tag{2.2.5}$$

Moreover, if we define

$$h(p) = \sum_m c_m e^{-i\frac{1}{2}mp}, \tag{2.2.6}$$

then the scale relation (2.2.1) becomes

$$\hat{\eta}(p) = \frac{1}{2} h(p) \hat{\eta}\left(\frac{1}{2}p\right). \tag{2.2.7}$$

Since $h(p)$ is 4π-periodic, so is the expression

$$\tilde{h}(p) = \left(\frac{\sum_\ell |\hat{\eta}(\frac{1}{2}p + 2\pi\ell)|^2}{\sum_\ell |\hat{\eta}(p + 2\pi\ell)|^2} \right)^{1/2} h(p). \tag{2.2.8}$$

The point is that

$$\widehat{\varphi}(p) = \frac{1}{2} \tilde{h}(p) \widehat{\varphi}\left(\frac{1}{2}p\right), \tag{2.2.9}$$

so if we set

$$\tilde{h}(p) = \sum_m \tilde{c}_m e^{-i\frac{1}{2}mp}, \tag{2.2.10}$$

we have a scale relation for φ:

$$\varphi(x) = \sum_m \tilde{c}_m \varphi(2x - m). \tag{2.2.11}$$

A scaling function φ has been constructed that has the additional property (2.2.5). Finally, the wavelet is defined by

$$\Psi(x) = 2 \sum_m (-1)^m \tilde{c}^{\,*}_{1-m} \varphi(2x - m). \tag{2.2.12}$$

The orthogonality between scales is induced by the identity

$$\sum_m (-1)^m \tilde{c}_m \tilde{c}_{1-m} = 0, \tag{2.2.13}$$

and the completeness of the orthonormal set $\{\Psi_r(\cdot - 2^{-r}n)\colon\ n, r \in \mathbb{Z}\}$ follows from property (a) of φ together with the observation that

$$\mathcal{H}_{r_0} \subset \overline{\operatorname{span}\{\Psi_r(\cdot - 2^{-r}n)\colon\ n, r \in \mathbb{Z}, r \le r_0 - 1\}}. \tag{2.2.14}$$

Indeed, if we set

$$\mathcal{K}_r = \overline{\operatorname{span}\{\Psi_r(\cdot - 2^r n)\colon\ n \in \mathbb{Z}\}}, \tag{2.2.15}$$

then $\mathcal{K}_r$ is orthogonal to $\mathcal{H}_r$ and

$$\mathcal{H}_{r+1} = \mathcal{H}_r + \mathcal{K}_r. \tag{2.2.16}$$

The iteration of this decomposition implies (2.2.14) because $\bigcap_r \mathcal{H}_r$ is the null space.

This machinery reduces the problem of constructing an orthonormal wavelet basis with given phase space localization properties to one of finding a scaling function with the desired properties. For our first class of examples we take the standard Nth-degree spline with N ranging over all non-negative integers. This means we define η as the $(N+1)$-fold convolution of χ with itself. Equivalently,

$$\hat{\eta}(p) = \widehat{\chi}(p)^{N+1}, \tag{2.2.17}$$

and it is easy enough to verify that

$$\hat{\eta}(p) = \frac{1}{2}(1 + e^{-i\frac{1}{2}p})^{N+1}\hat{\eta}\left(\frac{1}{2}p\right) \tag{2.2.18}$$

Hence

$$\eta(x) = \sum_{m=0}^{N+1} \binom{N+1}{m} \eta(2x - m), \tag{2.2.19}$$

and η also has property (c) because

$$\sum_\ell |\hat{\eta}(p + 2\pi\ell)|^2 = |e^{-ip} - 1|^{2N+2} \sum_\ell \frac{1}{(p + 2\pi\ell)^{2N+2}}, \tag{2.2.20}$$

$$\lim_{p\to 2\pi m} \frac{|e^{-ip} - 1|^{2N+2}}{(p + 2\pi\ell)^{2N+2}} = \delta_{\ell m}. \tag{2.2.21}$$

As far as condition (a) is concerned, it suffices to show that for every orthogonal projection F of $L^2(\mathbb{R})$ defined by the characteristic function of some interval, the functions

$$F\eta_r(\cdot - 2^{-r}n), \qquad r, n \in \mathbb{Z},$$

span the subspace ran F. Suppose there is an interval for which this is not true. Then there is an element $\xi \in$ ran F such that

$$\xi \perp F\eta_r(\cdot - 2^{-r}n), \qquad r, n \in \mathbb{Z}.$$

Since $\xi(x)$ has compact support, $\hat{\xi}(p)$ is smooth. Moreover,

$$\xi \perp F\eta_r(\cdot - 2^{-r}n) \Rightarrow \xi \perp \eta_r(\cdot - 2^{-r}n),$$

so $\hat{\xi}(p)$ is orthogonal to

$$e^{-i2^{-r}np}\hat{\eta}_r(p), \qquad r, n \in \mathbb{Z}.$$

By Fourier analysis on each scale,

$$2^{r/2} \sum_\ell \hat{\xi}(p + 2^{-r+1}\pi\ell)\hat{\eta}_r(p + 2^{-r+1}\pi\ell)^* = 0 \tag{2.2.22}$$

for all $r \in \mathbb{Z}$, where $2^{r/2}$ is used as a normalizing factor in the $r = -\infty$ limit. Now, we also know that

$$\lim_{|p|\to\infty} \hat{\xi}(p) = 0 \tag{2.2.23}$$

by the Riemann-Lebesgue Lemma, and

$$\begin{aligned} 2^{r/2}|\hat{\eta}_r(p)| &= 2^{rN+r}\frac{|e^{-i2^{-r}p} - 1|^{N+1}}{|p|^{N+1}} \\ &\leq 2^{rN+N+r+1}|p|^{-N-1}. \end{aligned} \tag{2.2.24}$$

Since this bound is uniform in r for $r \leq 0$, we may infer dominated convergence for the ℓ-summation as $r \to -\infty$. Only the $\ell = 0$ term survives this limit, so

$$\hat{\xi}(p) \lim_{r\to-\infty} 2^{r/2}\hat{\eta}_r(p) = 0, \tag{2.2.25}$$

but the limit is unity. Thus $\hat{\xi}(p) \equiv 0$, and so η has property (a).

The wavelet basis that we can derive from this scaling function is the *Lemarié basis*. Although η has compact support, the Lemarié wavelets are only exponentially localized. The point is that the expression

$$\left(\sum_{\ell} |\hat{\eta}(p + 2\pi\ell)|^2\right)^{-1/2}$$

is only real-analytic in p – not entire. Therefore, the scaling function φ intermediate to the derivation has only exponential falloff. As far as smoothness is concerned, the wavelet Ψ is in the same class as η – namely class $C^{N-\varepsilon}$. This is most transparent if we write the formula (2.2.12) in momentum space:

$$\widehat{\Psi}(p) = -e^{i\frac{1}{2}p}\tilde{h}(2\pi - p)^*\widehat{\varphi}\left(\frac{1}{2}p\right). \tag{2.2.26}$$

$\tilde{h}$ is a 4π-periodic function, so $\widehat{\Psi}(p)$ has the same bound on large-p decay as $\widehat{\varphi}(p)$ and $\hat{\eta}(p)$ have.

How many vanishing moments does a Lemarié wavelet have? The key quantity is

$$\begin{aligned} h(p) &= \sum_{m=0}^{N+1} \binom{N+1}{m} e^{-i\frac{1}{2}mp} \\ &= (1 + e^{-i\frac{1}{2}p})^{N+1}, \end{aligned} \tag{2.2.27}$$

and the point is that

$$\tilde{h}(2\pi - p)^* = \left(\frac{\sum_{\ell} |\hat{\eta}(2\pi\ell + \pi - p)|^2}{\sum_{\ell} |\hat{\eta}(2\pi\ell - p)|^2}\right)^{1/2} h(2\pi - p)^*, \tag{2.2.28}$$

$$\begin{aligned} h(2\pi - p)^* &= (1 - e^{-i\frac{1}{2}p})^{N+1} \\ &= O(|p|^{N+1}), \end{aligned} \tag{2.2.29}$$

which implies

$$\int \Psi(x)x^n dx = 0, \qquad 0 \le n \le N. \tag{2.2.30}$$

The number of vanishing moments for a Lemarié wavelet is essentially equal to the degree of smoothness.

We now turn to the construction of a *Meyer basis* – a wavelet basis where the Fourier transform of each basis function is a compactly supported C^∞ function. Again, we start with a scaling function η, but now we choose η by specifying $\hat{\eta}$ as a positive C^∞ function such that

$$\hat{\eta}(p) = \begin{cases} 1, & |p| \le \pi, \\ 0, & |p| \ge \pi + \delta, \end{cases} \tag{2.2.31}$$

where we shall adjust $\delta > 0$ to be as small as we need. The construction of such a function is an elementary exercise; our main task is to show that this is indeed a scaling function. Property (c) is obvious; to find the sequence $\{c_m\}_{m\in\mathbb{Z}}$ for property (b), we observe that if we choose $\delta < \frac{1}{3}\pi$,

$$\hat{\eta}\left(\frac{1}{2}p\right)\hat{\eta}(p + 4\pi\ell) = \delta_{\ell 0}\hat{\eta}(p). \tag{2.2.32}$$

If we define

$$h(p) = 2\sum_{\ell}\hat{\eta}(p + 4\pi\ell), \tag{2.2.33}$$

then

$$\hat{\eta}(p) = \frac{1}{2}h(p)\hat{\eta}\left(\frac{1}{2}p\right), \tag{2.2.34}$$

so $\eta(x)$ satisfies the scaling relation (2.2.1) with $\{c_m\}_{m\in\mathbb{Z}}$ given by (2.2.6) and our current choice of $h(p)$.

There remains the problem of establishing property (a). To this end, we may choose an arbitrary $f \in L^2(\mathbb{R})$ such that $\hat{f}$ has compact support. Thus

$$\operatorname{supp}\hat{f} \subset \{-2^\nu(\pi - \delta) \le p \le 2^\nu(\pi - \delta)\} \tag{2.2.35}$$

for some integer ν. Let

$$g(p) = \sum_{\ell}\hat{f}(p + 2^{\nu+1}\pi\ell) \tag{2.2.36}$$

and consider the Fourier series of this $2^{\nu+1}\pi$-periodic function:

$$g(p) = \sum_m a_m e^{i2^{-\nu}mp}. \tag{2.2.37}$$

Since

$$\hat{\eta}(2^{-\nu}p)\hat{f}(p + 2^{\nu+1}\pi\ell) = \delta_{\ell 0}\hat{f}(p), \tag{2.2.38}$$

it follows that

$$\begin{aligned} \hat{f}(p) &= g(p)\hat{\eta}(2^{-\nu}p) \\ &= \sum_m a_m e^{-i2^{-\nu}mp}\hat{\eta}(2^{-\nu}p). \end{aligned} \tag{2.2.39}$$

Hence

$$f(x) = 2^{\nu} \sum_{m} a_m \eta(2^{\nu} x + m), \tag{2.2.40}$$

and so we have the density property.

The regularity and decay properties that have been selected for η in this case are preserved by the derivation because

$$\left(\sum_{\ell} |\hat{\eta}(p + 2\pi\ell)|^2 \right)^{-1/2}$$

is certainly C^{∞} as well. With $\widehat{\varphi}$ as a C^{∞} function with compact support, $\widehat{\Psi}$ automatically shares these properties through (2.2.26). However, $\tilde{h}(2\pi - p)$ has a more interesting property than smoothness. There is a neighborhood of $p = 0$ on which this expression vanishes. Since $\widehat{\Psi}(p)$ shares this property, it follows that all moments vanish for a Meyer wavelet – i.e.,

$$\int \Psi(x) x^n dx = 0 \tag{2.2.41}$$

for all integers $n \geq 0$.

References

1. I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," Commun. Pure Appl. Math. **41**, No. 7 (1988), 909–996.

2. P. Lemarié, "Ondelettes à Localisation Exponentielle," J. Math. Pures Appl. **67** (1988), 227–236.

3. P. Lemaré-Rieusset, "Sur l' Existence des Analyses Multi-Résolutions en Théorie des Ondelettes," Rev. Mat. Iberoamer. **8** (1992), 457–474.

4. P. Lemarié-Rieusset and J. Kahane, *Fourier Series and Wavelets,* Gordon and Breach, Luxembourg, 1995.

5. P. Lemarié and Y. Meyer, "Ondelettes et Bases Hilbertiennes," Rev. Mat. Iberoamer. **2** (1986), 1–18.

6. S. Mallat, "Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(\mathbb{R})$," Trans. Amer. Math. Soc. **315**, No. 1 (1989), 69–87.

7. Y. Meyer, "Principe d' Incertitude, Bases Hilbertiennes et Algèbres d' Opérateurs, " Séminaire Bourbaki **662** (1986), 209–223.

8. Y. Meyer, *Ondelettes at Opérateurs I: Ondelettes*, Hermann Press, Paris, 1990.

9. Y. Meyer, *Ondelettes, Fonctions Splines, et Analyses Graduées*, Cahiers des Mathématiques de la Decision No. 8703, Univ. Paris–Dauphine, 1987.

2.3 Daubechies Wavelets

Having constructed exponentially localized wavelets for every desired degree of smoothness, one may ask whether bases of compactly supported wavelets exist. The Haar basis is obviously an orthonormal basis of compactly supported wavelets, but these basis functions are discontinuous. A major development of the subject was the construction of wavelet bases that are smoother analogs of the Haar basis. The mother wavelet of such a basis is a *Daubechies wavelet*, and the discussion of these wavelets merits a section of its own.

As in the construction of Meyer and Lemarié wavelets carried out in the previous section, Daubechies wavelets are constructed within the framework of the Mallat–Meyer multi-scale resolution analysis, but there is an important obstacle to be reckoned with. It is not enough for the initial scaling function η to have compact support unless its integer translates are already orthonormal. After all, the initial scaling function in the Lemarié case has compact support, but the orthogonalization

$$\widehat{\varphi}(p) = \left(\sum_{\ell} |\hat{\eta}(p + 2\pi\ell)|^2 \right)^{-1/2} \hat{\eta}(p) \tag{2.3.1}$$

degrades this property to exponential falloff because the inverse square root in the momentum expression is real analytic but not entire. The scaling function φ must be derived in some other way.

On the other hand, once we have a compactly supported scaling function φ whose integer translates are orthonormal, the multi-scale resolution analysis yields a compactly supported wavelet. Indeed, if φ satisfies the scale relation

$$\varphi(x) = \sum_m c_m \varphi(2x - m) \tag{2.3.2}$$

and has compact support, then only a finite number of the coefficients are non-zero. Since the corresponding wavelet is given by

$$\Psi(x) = 2 \sum_m (-1)^m c^*_{1-m} \varphi(2x - m), \tag{2.3.3}$$

it must have compact support as well. Moreover, Ψ will possess the same degree of smoothness as φ.

To construct φ, we find it useful to analyze the properties that φ must have. In momentum space, the scale relation becomes

$$\widehat{\varphi}(p) = \frac{1}{2} h(p) \widehat{\varphi}\left(\frac{1}{2}p\right), \tag{2.3.4}$$

$$h(p) = \sum_m c_m e^{-i\frac{1}{2}mp}. \tag{2.3.5}$$

Iteration of this relation yields

$$\widehat{\varphi}(p) = \widehat{\varphi}(2^{-N}p) \prod_{k=0}^{N-1} \left(\frac{1}{2} h(2^{-k}p) \right), \tag{2.3.6}$$

and the $N = \infty$ limit yields at least a formalism for the recovery of the scaling function φ from the coefficients c_m:

$$\widehat{\varphi}(p) = \widehat{\varphi}(0) \prod_{k=0}^{\infty} \left(\frac{1}{2} h(2^{-k}p) \right) \tag{2.3.7}$$

Aside from the property that only a finite number of the c_m are non-zero, what other properties must the sequence $\{c_m\}$ have? The orthonormality property of φ induces a condition on the sequence via the scale relation. We have

$$\begin{aligned}
\delta_{n0} &= \int \varphi(x)^* \varphi(x-n) dx \\
&= \sum_{m,\bar{m}} c_m^* c_{\bar{m}} \int \varphi(2x-m)^* \varphi(2x-2n-\bar{m}) dx \\
&= \frac{1}{2} \sum_{m,\bar{m}} c_m^* c_{\bar{m}} \int \varphi(x'-m)^* \varphi(x'-2n-\bar{m}) dx' \\
&= \frac{1}{2} \sum_m c_m^* c_{m-2n}.
\end{aligned} \tag{2.3.8}$$

This may be written as a condition on $h(p)$. Indeed,

$$\begin{aligned}
\sum_m c_m^* c_{m-2n} &= \frac{1}{(4\pi)^2} \int_{-2\pi}^{2\pi} dp \int_{-2\pi}^{2\pi} dp'\, h(p)^* h(p') \sum_m e^{-i\frac{1}{2}mp + i\frac{1}{2}mp' - inp'} \\
&= \frac{1}{(4\pi)^2} \int_{-2\pi}^{2\pi} dp \int_{-2\pi}^{2\pi} dp'\, h(p)^* h(p') e^{-inp'} \sum_\ell 2\pi\delta\left(\frac{1}{2}p' - \frac{1}{2}p + 2\pi\ell\right) \\
&= \frac{1}{4\pi} \int_{-2\pi}^{2\pi} dp\, |h(p)|^2 e^{-inp} \\
&= \frac{1}{4\pi} \int_0^{2\pi} e^{-inp} (|h(p)|^2 + |h(p+2\pi)|^2),
\end{aligned} \tag{2.3.9}$$

so it follows from (2.3.8) that

$$\frac{1}{4}|h(p)|^2 + \frac{1}{4}|h(p+2\pi)|^2 = 1. \tag{2.3.10}$$

In summary, we are looking for 4π-periodic functions $h(p)$ which are trigonometric polynomials, which satisfy Eq. (2.3.10), and for which the infinite product in (2.3.7) converges uniformly on compact subsets of the complex plane. The point is that if $\widehat{\varphi}(p)$ extends to an entire function whose growth in the imaginary direction is bounded exponentially, then by a contour-shifting argument in momentum space, $\varphi(x)$ has compact support.

At this stage, it is instructive to consider the Haar basis from this point of view. In this case, $\varphi = \eta = \chi$ and

$$h(p) = 1 + e^{-i\frac{1}{2}p}, \tag{2.3.11}$$

where (2.3.10) is just the trigonometric identity

$$\cos^2 \frac{1}{4}p + \sin^2 \frac{1}{4}p = 1. \tag{2.3.12}$$

The infinite product identity is

$$\prod_{k=0}^{\infty} \left(\frac{1 + e^{i2^{-k-1}p}}{2} \right) = \frac{e^{-ip} - 1}{-ip}, \tag{2.3.13}$$

and the r.h.s. is just the Fourier transform of the characteristic function χ. Actually, this formula plays a key role in the construction of Daubechies scaling functions. The strategy is to write

$$h(p) = (1 + e^{-i\frac{1}{2}p})^{N_0} Q(e^{i\frac{1}{2}p}), \tag{2.3.14}$$

where N_0 is to be chosen by the desired order of differentiability, and examine the conditions which $Q(z)$ must satisfy. Since

$$\prod_{k=0}^{\infty} \left(\frac{1 + e^{-i2^{-k-1}p}}{2} \right)^{N_0} = \widehat{\chi}(p)^{N_0}, \tag{2.3.15}$$

Eq. (2.3.7) reduces to

$$\widehat{\varphi}(p) = \widehat{\varphi}(0)\widehat{\chi}(p)^{N_0} \prod_{k=0}^{\infty} (2^{N_0 - 1} Q(e^{i2^{-k-1}p})), \tag{2.3.16}$$

so the infinite product in Q must be bounded by a certain N_0-dependent power of p if φ is to have the desired regularity. Moreover, note that Eq. (2.3.10) becomes

$$\frac{1}{4} \cos^{2N_0} \left(\frac{1}{4}p \right) |Q(e^{i\frac{1}{2}p})|^2 + \frac{1}{4} \sin^{2N_0} \left(\frac{1}{4}p \right) |Q(-e^{i\frac{1}{2}p})|^2 = 4^{-N_0}. \tag{2.3.17}$$

We impose the additional condition that the polynomial $Q(z)$ has only real coefficients. Thus we have the factorization

$$Q(z) = \prod_{\mu} (z^2 + a_\mu z + b_\mu) \tag{2.3.18}$$

with a_μ and b_μ real, so for $|z| = 1$,

$$\begin{aligned} |Q(z)|^2 &= \prod_{\mu} |z^2 + a_\mu z + b_\mu|^2 \\ &= \prod_{\mu} (1 + a_\mu^2 + b_\mu^2 + a_\mu(b_\mu + 1)(z + \bar{z}) + b_\mu(z^2 + \bar{z}^2)) \\ &= \prod_{\mu} (1 + a_\mu^2 + b_\mu^2 - 2b_\mu + a_\mu(b_\mu + 1)(z + \bar{z}) + b_\mu(z + \bar{z})^2). \end{aligned} \tag{2.3.19}$$

This makes $|Q(e^{i\frac{1}{2}p})|^2$ a polynomial in $\cos(\frac{1}{2}p)$ and therefore in the variable

$$y = \sin^2\left(\frac{1}{4}p\right) = \frac{1}{2} - \frac{1}{2}\cos\left(\frac{1}{2}p\right). \tag{2.3.20}$$

Accordingly, we set

$$|Q(e^{i\frac{1}{2}p})|^2 = P(y), \qquad 0 \leq y \leq 1, \tag{2.3.21}$$

and observe that

$$|Q(-e^{i\frac{1}{2}p})|^2 = P(1-y). \tag{2.3.22}$$

Equation (2.3.17) becomes

$$\frac{1}{4}(1-y)^{N_0}P(y) + \frac{1}{4}y^{N_0}P(1-y) = 4^{-N_0}. \tag{2.3.23}$$

Once a polynomial $P(y)$ is found to satisfy the reduced conditions, one can recover $Q(z)$ by virtue of the *Polya–Szego Lemma*. One version of this elementary lemma states that if $P(y)$ is *any* polynomial for which $P(y) \geq 0$ on the interval $0 \leq y \leq 1$, then there exists a polynomial $Q(z)$ with *real coefficients* such that

$$|Q(e^{i\xi})|^2 = P\left(\sin^2\frac{1}{2}\xi\right) \tag{2.3.24}$$

We omit the proof.

It is not difficult to find a polynomial solution of Eq. (2.3.23). Verification that

$$P(y) = 4^{-N_0+1}\sum_{\nu=0}^{N_0-1}\binom{N_0-1+\nu}{\nu}y^\nu \tag{2.3.25}$$

is a solution is based on the combinatoric identities

$$\sum_{\mu=0}^{M}\binom{n+\mu}{\mu} = \binom{n+M+1}{M}, \tag{2.3.26}$$

$$\sum_{\mu=0}^{M}\binom{n+\mu}{\mu}[y^{n+1}(1-y)^\mu + y^\mu(1-y)^{n+1}] = 1, \tag{2.3.27}$$

which can be established by induction arguments. The remaining problem is to control the product

$$F(p) = \prod_{k=0}^{\infty}(2^{N_0-1}Q(e^{i2^{-k-1}p})), \tag{2.3.28}$$

and we write the polynomial $Q(z)$ in the form

$$Q(z) = 2^{-N_0+1}\sum_m \gamma_m z^m. \tag{2.3.29}$$

Since $Q(1) = 2^{-N_0+1}$ as a consequence of $h(0) = 2$, we have

$$\sum_m \gamma_m = 1, \tag{2.3.30}$$

$$F(w) = \prod_{k=0}^{\infty} \left(\sum_m \gamma_m e^{i2^{-k-1}mw} \right), \tag{2.3.31}$$

and this product converges uniformly on compact subsets in the complex variable w because

$$\begin{aligned}
\sum_{k=0}^{\infty} \left| \sum_m \gamma_m e^{i2^{-k-1}mw} - 1 \right| &= \sum_{k=0}^{\infty} \left| \sum_m \gamma_m (e^{i2^{-k-1}mw} - 1) \right| \\
&\le \sum_m |\gamma_m| \sum_{k=0}^{\infty} |e^{i2^{-k-1}mw} - 1| \\
&\le \sum_m |\gamma_m| \sum_{k=0}^{\infty} 2^{-k-1} m|w| e^{2^{-k-1}m|\mathrm{Im}\ w|} \\
&\le \sum_m m|\gamma_m|\, |w| e^{\frac{1}{2}m|\mathrm{Im}\ w|}.
\end{aligned} \tag{2.3.32}$$

Thus $F(w)$ is an entire function; moreover, the growth of $F(w)$ in Im w is exponentially bounded. Indeed, for $|w| \ge 1$,

$$\begin{aligned}
&\sum_{k>\log_2|w|} \left| \sum_m \gamma_m e^{i2^{-k-1}mw} - 1 \right| \\
\le\ & \sum_m |\gamma_m| \sum_{k>\log_2|w|} 2^{-k-1} m|w| e^{2^{-k-1}m|\mathrm{Im}\ w|} \\
\le\ & \sum_m m|\gamma_m|\, |w| \sum_{k>\log_2|w|} 2^{-k-1} \exp(2^{-\log_2|w|-1} m|\ \mathrm{Im}\ w|) \\
\le\ & \sum_m m|\gamma_m| e^{\frac{1}{2}m},
\end{aligned} \tag{2.3.33}$$

$$\begin{aligned}
&\prod_{k\le\log_2|w|} \left| \sum_m \gamma_m e^{i2^{-k-1}mw} \right| \\
\le\ & \left(\sum_m |\gamma_m| \right)^{\log_2|w|} \exp\left(m_0 |\mathrm{Im}\ w| \sum_{k\le\log_2|w|} 2^{-k-1} \right) \\
\le\ & |w|^{\log_2 \sum_m |\gamma_m|} e^{m_0|\mathrm{Im}\ w|},
\end{aligned} \tag{2.3.34}$$

where m_0 is the largest integer m for which $\gamma_m \ne 0$. This bound on the growth of $F(w)$ together with (2.3.32) and a contour-shifting argument imply that φ has compact support, provided $F(w)$ has enough decay in Re w for bounded Im w. We focus on

$F(p)$ for simplicity, as the estimation does not significantly change with a bounded imaginary part. This issue coincides with the remaining issue of regularity of φ, since the latter is determined by how much polynomial decay in p can be extracted from $\widehat{\varphi}(p)$.

First, we can certainly refine the inequality (2.3.34) on the real axis as follows:

$$\begin{aligned}
\prod_{k\leq\log_2|p|}\left|\sum_m \gamma_m e^{i2^{-k-1}mp}\right|^2 &= \prod_{k\leq\log_2|p|}(4^{N_0-1}P(\sin^2(2^{-k-2}p))) \\
&\leq \left(\sum_{\nu=0}^{N_0-1}\binom{N_0-1+\nu}{\nu}\right)^{\log_2|p|} \\
&= \binom{2N_0-1}{N_0-1}^{\log_2|p|} \\
&\leq \left(2^{N_0-1}\left(2-\frac{\ln(N_0+1)-\ln 2}{N_0-1}\right)^{N_0-1}\right)^{\log_2|p|} \\
&\leq (2^{2N_0-2})^{\log_2|p|} \\
&= |p|^{2N_0-2}
\end{aligned} \tag{2.3.35}$$

for $|p|\geq 1$. This implies

$$|F(p)|\leq|p|^{N_0-1}, \qquad |p|\geq 1, \tag{2.3.36}$$

from which we infer

$$|\widehat{\varphi}(p)|\leq c|p|^{-1}, \qquad |p|\geq 1. \tag{2.3.37}$$

This initial estimation is clearly not sharp enough. After all, this bound is no better than the bound on $\widehat{\chi}(p)$, and χ is the scaling function for the Haar basis. With only such a bound on $\widehat{\varphi}(p)$, the whole strategy would be much ado about nothing, so we now describe how to obtain a better bound.

The obvious refinement is to take the y-dependence of $P(y)$ into account, but how to do this effectively may not be so obvious. The idea is to pair the factors when we estimate a product as in (2.3.35) above. To this end, we consider

$$\begin{aligned}
\prod_{k=0}^{2\ell+1}\left|\sum_m \gamma_m e^{i2^{-k-1}mp}\right|^2 &= \prod_{k=0}^{2\ell+1}(4^{N_0-1}P(\sin^2(2^{-k-2}p))) \\
&= \prod_{j=0}^{\ell}(4^{2N_0-2}P(\sin^2(2^{-2j-2}p))P(\sin^2(2^{-2j-3}p))) \\
&= \prod_{j=0}^{\ell}(4^{2N_0-2}P(4\cos^2(2^{-2j-3}p)\sin^2(2^{-2j-3}p)) \\
&\quad \times P(\sin^2(2^{-2j-3}p))).
\end{aligned} \tag{2.3.38}$$

Thus, each pair product has the form

$$4^{2N_0-2}P(4(1-y)y)P(y) = \sum_{\nu,\nu'=0}^{N_0-1} \binom{N_0-1+\nu}{\nu}\binom{N_0-1+\nu'}{\nu'} \times 4^{\nu}(1-y)^{\nu}y^{\nu+\nu'}. \tag{2.3.39}$$

Since

$$4^{N_0-1}P(y) \le \sum_{\nu=0}^{N_0-1} 2^{N_0+\nu-2}y^{\nu} \le 2^{N_0-2}N_0 \max\{(2y)^{N_0-1},1\}, \tag{2.3.40}$$

we have

$$4^{2N_0-2}P(4(1-y)y)P(y)$$

$$\le \begin{cases} 4^{N_0-2}N_0^2, & 0 \le y \le \frac{1}{2}-\frac{1}{4}\sqrt{2}, \quad (2.3.41.1)\\ 2^{5N_0-7}N_0^2 y^{N_0-1}(1-y)^{N_0-1}, & \frac{1}{2}-\frac{1}{4}\sqrt{2} \le y \le \frac{1}{2}, \quad (2.3.41.2)\\ 4^{3N_0-4}N_0^2 y^{2N_0-2}(1-y)^{N_0-1}, & \frac{1}{2} \le y \le \frac{1}{2}+\frac{1}{4}\sqrt{2}, \quad (2.3.41.3)\\ 2^{3N_0-5}N_0^2 y^{N_0-1}, & \frac{1}{2}+\frac{1}{4}\sqrt{2} \le y \le 1. \quad (2.3.41.4) \end{cases}$$

In each case, the range of y may be exploited to further reduce the bound. Thus,

$$4^{2N_0-2}P(4(1-y)y)P(y)$$

$$\le \begin{cases} 4^{N_0-2}N_0^2 & 0 \le y \le \frac{1}{2}-\frac{1}{4}\sqrt{2}, \quad (2.3.42.1)\\ 4^{2N_0-3}N_0^2\left(\frac{1}{2}+\frac{1}{4}\sqrt{2}\right)^{N_0-1}, & \frac{1}{2}-\frac{1}{4}\sqrt{2} \le y \le \frac{1}{2}, \quad (2.3.42.2)\\ 4^{2N_0-3}N_0^2\left(\frac{1}{2}+\frac{1}{4}\sqrt{2}\right)^{2N_0-2}, & \frac{1}{2} \le y \le \frac{1}{2}+\frac{1}{4}\sqrt{2}, \quad (2.3.42.3)\\ 2^{3N_0-5}N_0^2, & \frac{1}{2}+\frac{1}{4}\sqrt{2} \le y \le 1. \quad (2.3.42.4) \end{cases}$$

From this we easily infer that

$$4^{2N_0-2}P(4(1-y)y)P(y) \le 4^{2N_0-3}N_1^2\left(\frac{1}{2}+\frac{1}{4}\sqrt{2}\right)^{N_0-1} \tag{2.3.43}$$

for all values of y. Therefore,

$$\prod_{k=0}^{2\ell+1}\left|\sum_m \gamma_m e^{i2^{-k-1}mp}\right|^2 \le \left(4^{2N_0-3}N_0^2\left(\frac{1}{2}+\frac{1}{4}\sqrt{2}\right)^{N_0-1}\right)^{\ell+1} = (4^{N_0-2}N_0^2(2+\sqrt{2})^{N_0-1})^{\ell+1}, \tag{2.3.44}$$

so if we choose ℓ to be the largest integer such that

$$2\ell \le \log_2|p|, \tag{2.3.45}$$

we obtain

$$\prod_{k=0}^{2\ell+1}\left|\sum_m \gamma_m e^{i2^{-k-1}mp}\right|^2 \le (4^{N_0-2}N_0^2(2+\sqrt{2})^{N_0-1})^{\frac{1}{2}\log_2|p|+1}$$
$$= 4^{N_0-2}N_0^2(2+\sqrt{2})^{N_0-1}|p|^{N_0-2+\log_2 N_0+\frac{1}{2}(N_0-1)\log_2(2+\sqrt{2})}. \qquad (2.3.46)$$

In summary,

$$2\ell+2 > \log_2|p| \Rightarrow \prod_{k=2\ell+2}^{\infty}\left|\sum_m \gamma_m e^{i2^{-k-1}mp}\right| \le c \qquad (2.3.47)$$

as a consequence of the estimation, while

$$2\ell \le \log_2|p| \Rightarrow \prod_{k=0}^{2\ell+1}\left|\sum_m \gamma_m e^{i2^{-k-1}mp}\right| \le 2^{N_0-2}N_0$$
$$\times(2+\sqrt{2})^{\frac{1}{2}N_0-\frac{1}{2}}|p|^{\frac{1}{2}(N_0-1)(1+\frac{1}{2}\log_2(2+\sqrt{2}))+\frac{1}{2}\log_2 N_0-\frac{1}{2}}. \qquad (2.3.48)$$

Since ℓ has been chosen such that both conditions hold, we finally have

$$|F(p)| \le c|p|^{(N_0-1)(\frac{1}{2}+\frac{1}{4}\log_2(2+\sqrt{2}))+\frac{1}{2}\log_2 N_0-\frac{1}{2}}, \quad |p|\ge 1, \qquad (2.3.49)$$

from which it follows that

$$|\widehat{\varphi}(p)| \le c|p|^{(N_0-1)(-\frac{1}{2}+\frac{1}{4}\log_2(2+\sqrt{2}))+\frac{1}{2}\log_2 N_0-\frac{3}{2}}, \quad |p|\ge 1. \qquad (2.3.50)$$

This is a decay law because

$$-\frac{1}{2}+\frac{1}{4}\log_2(2+\sqrt{2}) < 0. \qquad (2.3.51)$$

The power of the decay is roughly proportional to N_0, so at last we have constructed an arbitrarily regular φ.

We have shown how to construct the simplest class of Daubechies wavelets. Variations on this construction can be – and have been – pursued as well. Moreover, the estimation used above for establishing regularity can be significantly refined.

References

1. I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," Commun. Pure Appl. Math. **41**, No. 7 (1988), 909–996.

2. A. Haar, "Zur Theorie der Orthogonalen Funktionsysteme," Math. Ann. **69** (1910), 331–371.

3. G. Polya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis, Vol. II*, Springer, Berlin, 1971.

2.4 Vanishing Moments

In §2.1 we produced some crude pictorial intuition for the expectation that $\widehat{\Psi}(0) = 0$ for any reasonable wavelet Ψ. We now wish to make this result precise and to generalize it to moments of arbitrary order. The constructions in the last couple of sections have a number of vanishing moments comparable to the degree of smoothness, and this is no coincidence. Also, this connection will lead to a no-go theorem on the type of phase space localization that a wavelet can have. The latter result is similar in spirit to the Balian–Low Theorem for discrete translational symmetry in phase space. The result on vanishing moments will also lead to more quantitative restrictions on phase space localization of wavelets – in the form of uncertainty relations – but we pursue this in the next section.

We first consider a hypothesis for which the intuition is transparent. Let Ψ be an N-adic wavelet such that $\widehat{\varphi}(p)$ is continuous and integrable. In particular, $\Psi(x)$ is a continuous function that vanishes at infinity. There exist integers n_0 and r_0 such that

$$\Psi(N^{-r_0}n_0) \neq 0 \tag{2.4.1}$$

because rational numbers of the form $N^{-r}n$ with $n, r \in \mathbb{Z}$ are dense in $\mathbb{R}$. Now $x_0 = N^{-r_0}n_0$ lies on every finer lattice: define

$$n(r) = N^{r-r_0}n_0, \qquad r \geq r_0, \tag{2.4.2}$$

so that

$$x_0 = N^{-r}n(r). \tag{2.4.3}$$

The orthogonality property of the N-adic wavelet implies

$$\begin{aligned} \int e^{ix_0p}\widehat{\Psi}(p)\widehat{\Psi}(N^{-r}p)^* dp &= N^r \int \Psi(x)\Psi(N^r x - n(r))^* dx \\ &= 0, \qquad r \geq \max\{1, r_0\}. \end{aligned} \tag{2.4.4}$$

Combining the integrability and continuity of $\widehat{\Psi}$ with the Dominated Convergence Theorem, we see that the momentum integral converges to $\Psi(x_0)\widehat{\Psi}(0)^*$ as $r \to \infty$. Since $\Psi(x_0) \neq 0$, it follows that $\widehat{\Psi}(0) = 0$, so the zeroth-order moment of Ψ vanishes. This consequence of integrability and continuity in momentum space is easy enough to visualize. If $\widehat{\Psi}(0)$ were non-zero, the comparison of drastically different scales would have the graphical representation given by Fig. 2.4.1, and the point is that the wavelet with larger momentum scale (smaller length scale) is almost constant over much shorter momentum intervals. The orthogonality of $e^{ipx_0}\widehat{\Psi}(p)$ to a function that is almost constant over the essential locality of $e^{ipx_0}\widehat{\Psi}(p)$ would imply that $\Psi(x_0) = 0$ in the limit as the scale difference increases, which contradicts the choice of x_0.

Actually, the version of this result that generalizes most naturally to higher-order moments is slightly different. Suppose instead that both $x\Psi(x)$ and $\Psi'(x)$ are square-integrable, which means that $\widehat{\Psi}'(p)$ and $p\widehat{\Psi}(p)$ are both square-integrable as well. This is a stronger hypothesis because it implies on one hand that $\widehat{\Psi}'(p)$ is locally integrable

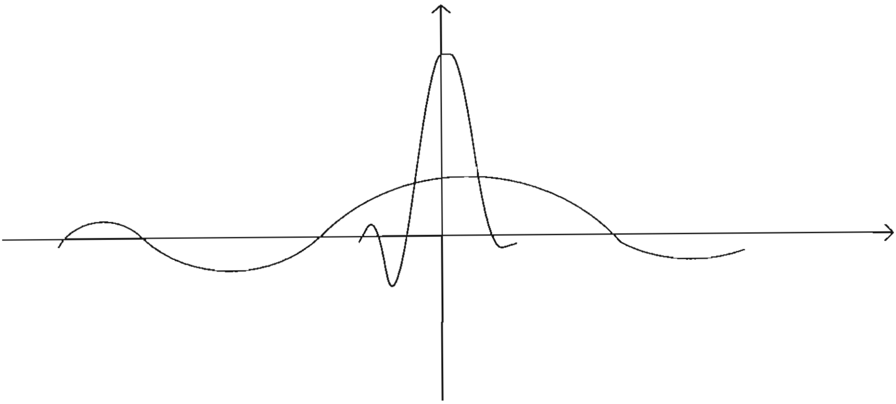

Figure 2.4.1:

and therefore $\widehat{\Psi}(p)$ is absolutely continuous, while on the other hand it implies

$$\begin{aligned}\int_{|p|\geq 1} |\widehat{\Psi}(p)|dp &\leq \left(\int_{|p|\geq 1} |p|^{-2}dp\right)^{1/2} \left(\int_{|p|\geq 1} |p|^2|\widehat{\Psi}(p)|^2 dp\right)^{1/2} \\ &\leq \sqrt{2}\left(\int |p\widehat{\Psi}(p)|^2 dp\right)^{1/2} \\ &< \infty. \end{aligned} \tag{2.4.5}$$

Now for order $\nu > 0$ we consider the hypothesis that $\Psi^{(\nu+1)}(x)$ and $x^{\nu+1}\Psi(x)$ are both square-integrable. We propose to show that

$$\int x^\mu \Psi(x)dx = 0, \qquad \mu \leq \nu, \tag{2.4.6}$$

and we do so by induction on μ. First, it is easy enough to infer that in particular, $\Psi'(x)$ and $x\Psi(x)$ are square-integrable, so this null condition certainly holds for $\mu = 0$, as we have already argued that $\widehat{\Psi}(0) = 0$ in this event. Now suppose the condition holds for $\mu \leq k-1$, where k is a positive integer $\leq \nu$, and assume

$$\int x^k \Psi(x)dx \neq 0, \tag{2.4.7}$$

which means

$$\widehat{\Psi}^{(k)}(0) \neq 0. \tag{2.4.8}$$

By the induction hypothesis,

$$\widehat{\Psi}(p) = \frac{1}{k!}\widehat{\Psi}^{(k)}(0)p^k + R_k(p), \tag{2.4.9}$$

$$R_k(p) = \frac{1}{(k+1)!}\int_0^p \widehat{\Psi}^{(k+1)}(p-q)q^k dq, \tag{2.4.10}$$

and the Taylor remainder makes sense because $\widehat{\Psi}^{(k+1)}(p)$ is square-integrable and therefore locally integrable. This time we choose integers n_0 and r_0 such that

$$\Psi^{(k)}(x_0) \neq 0, \qquad x_0 = N^{-r_0} n_0, \tag{2.4.11}$$

but we consider the same orthogonality condition as before:

$$\int e^{ix_0 p}\widehat{\Psi}(p)\widehat{\Psi}(N^{-r}p)^* dp = 0, \qquad r \geq \max\{1, r_0\}. \tag{2.4.12}$$

We have the Fourier transform identity

$$\frac{N^{-rk}}{k!}\widehat{\Psi}^{(k)}(0)^* \int e^{ix_0 p}\widehat{\Psi}(p)p^k dp = \frac{i^{-k}N^{-rk}}{k!}\widehat{\Psi}^{(k)}(0)^*\Psi^{(k)}(x_0), \tag{2.4.13}$$

while the boundedness of $\widehat{\Psi}$ – which follows from the integrability of Ψ – implies

$$\begin{aligned} |R_k(p)| &= \left|\widehat{\Psi}(p) - \frac{1}{k!}\widehat{\Psi}^{(k)}(0)p^k\right| \\ &\leq c_\Psi(1 + |p|^k). \end{aligned} \tag{2.4.14}$$

Now the strategy is to combine

$$\widehat{\Psi}(N^{-r}p) = \frac{N^{-rk}}{k!}\widehat{\Psi}^{(k)}(0)p^k + R_k(N^{-r}p) \tag{2.4.15}$$

with (2.4.12) and (2.4.13) to obtain the equation we intend to exploit:

$$\frac{i^{-k}N^{-rk}}{k!}\widehat{\Psi}^{(k)}(0)^*\Psi^{(k)}(x_0) + \int e^{ix_0 p}\widehat{\Psi}(p)R_k(N^{-r}p)^* dp = 0. \tag{2.4.16}$$

The idea is to show that the integral decreases *faster* than N^{-rk} as $r \to \infty$. This will achieve the contradiction to (2.4.8) that we desire. To this end, we decompose the integral over the regions $|p| \leq \varepsilon N^r$ and $|p| > \varepsilon N^r$. The integral over the latter region is estimated by using (2.4.14) as follows:

$$\begin{aligned} &\left|\int_{|p|>\varepsilon N^r} e^{ix_0 p}\widehat{\Psi}(p)R_k(N^{-r}p)^* dp\right| \\ \leq\ & c_\Psi \int_{|p|>\varepsilon N^r} |\widehat{\Psi}(p)|(1 + N^{-rk}|p|^k)dp \\ \leq\ & c_\Psi \left(\int_{|p|<\varepsilon N^r} |p^{k+1}\widehat{\Psi}(p)|^2 dp\right)^{1/2} \left[\left(\int_{|p|>\varepsilon N^r} |p|^{-2k-2} dp\right)^{1/2} \right. \\ & \left. + N^{-rk}\left(\int_{|p|>\varepsilon N^r} |p|^{-2} dp\right)^{1/2}\right] \end{aligned}$$

$$
\begin{aligned}
&\leq c_{\Psi}\|\Psi^{(k+1)}\|_2\left[\sqrt{\frac{2}{2k+1}}\varepsilon^{-k-\frac{1}{2}}N^{-rk-\frac{1}{2}r}\right. \\
&\qquad \left. + \sqrt{2}\varepsilon^{-\frac{1}{2}}N^{-rk-\frac{1}{2}r}\right] \\
&= O(N^{-rk-\frac{1}{2}r}). \qquad (2.4.17)
\end{aligned}
$$

To estimate the integral over the former region, we apply Schwarz estimation to the q-integration in (2.4.10) to obtain for the remainder a bound complementary to (2.4.14) – namely,

$$|R_k(p)| \leq c\|\widehat{\Psi}^{(k+1)}\|_2|p|^{k+\frac{1}{2}}. \qquad (2.4.18)$$

Thus

$$
\begin{aligned}
&\left|\int_{|p|\leq\varepsilon N^r} e^{ixop}\widehat{\Psi}(p)R_k(N^{-r}p)^* dp\right| \\
&\leq c\|\widehat{\Psi}^{(k+1)}\|_2 N^{-rk-\frac{1}{2}r}\int_{|p|\leq\varepsilon N^r}|\widehat{\Psi}(p)|\,|p|^{k+\frac{1}{2}}dp \\
&\leq c\|\widehat{\Psi}^{(k+1)}\|_2 N^{-rk-\frac{1}{2}r}\left(\int_{|p|\leq\varepsilon N^r}(1+|p|)^{-1}dp\right)^{1/2} \\
&\qquad \times\left(\int_{|p|\leq\varepsilon N^r}(1+|p|)^{2k+2}|\widehat{\Psi}(p)|^2 dp\right)^{1/2} \\
&= O(N^{-rk-\frac{1}{2}r}\sqrt{\ln(N^r)}), \qquad (2.4.19)
\end{aligned}
$$

so we have the desired type of bound on the large-r behavior of the integral. This completes the induction step in the proof (2.4.6).

An interesting consequence of this fundamental theorem is that an N-adic wavelet cannot have exponential localization in both position space and momentum space. Indeed, suppose that $\Psi(x)$ and $\widehat{\Psi}(p)$ both have exponential decay. Then $x^\mu\psi(x)$ and $\Psi^{(\mu)}(x)$ are both square-integrable for *all* integers $\mu \geq 0$, so by our result,

$$\widehat{\Psi}^{(\mu)}(0) = 0 \qquad (2.4.20)$$

for all μ as well. On the other hand, if $\widehat{\Psi}(p)$ actually vanishes to infinite order at $p=0$, then it cannot be analytic there. Finally, if $\widehat{\Psi}(p)$ is not real analytic, then $\Psi(x)$ cannot have exponential decay, and so we have a contradiction.

References

1. G. Battle, "Phase Space Localization Theorem for Ondelettes," J. Math. Phys. **30**, No. 10 (1989), 2195–2197.

2. I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," Commun. Pure Appl. Math. **41**, No. 7 (1988), 909–996.

3. P. Lemarié "Ondelettes à Localisation Exponentielle," J. Math. Pures Appl. **67** (1988), 227–236.

4. Y. Meyer, "Principe d'Incertitude, Bases Hilbertiennes et Algèbres d'Operateurs," Séminaire Bourbaki **662** (1986), 209–223.

5. M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. II, Fourier Analysis and Self-Adjointness*, Academic Press, New York, 1975.

2.5 Uncertainty Relations for $\langle P\rangle = 0$ Wavelet States

The vanishing-moment theorem for wavelets leads to inequalities of Heisenberg type (for wavelets) that are intimately related to scaling. The role of scaling in the proof of such inequalities has never been major for arbitrary states. Even in the proof of the celebrated inequality for an arbitrary state Ψ only the self-adjointness of the scaling generator S is used:

$$\begin{aligned}
\langle P^2\rangle_\Psi^{1/2}\langle X^2\rangle_\Psi^{1/2} &= \|P\Psi\|_2\|X\Psi\|_2 \\
&\geq |(P\Psi, X\Psi)| \\
&= \left|\left(\left(S+i\frac{1}{2}\right)\Psi,\Psi\right)\right| \\
&\geq \frac{1}{2}.
\end{aligned} \tag{2.5.1}$$

(Although $\hbar$ appeared as a parameter in §1.3, recall that we set $\hbar = 1$ in §1.8.) For states that are wavelets, the scaling generator will play a more quantitative role than it seems to play in this most basic estimation.

First, suppose the state Ψ happens to be the derivative of some square-integrable function, with no other assumption made. Thus

$$\Psi = P\zeta, \qquad \zeta \in L^2(\mathbb{R}), \tag{2.5.2}$$

so we have

$$\begin{aligned}
\|X\Psi\|_2 &= \|XP\zeta\|_2 \\
&= \left\|\left(S+i\frac{1}{2}\right)\zeta\right\|_2 \\
&= \left\|\left(S-i\frac{1}{2}\right)\zeta\right\|_2
\end{aligned} \tag{2.5.3}$$

Hence

$$\begin{aligned}
\langle P^2\rangle_\Psi^{1/2}\langle X^2\rangle_\Psi^{1/2} &= \|P\Psi\|_2\left\|\left(S-i\frac{1}{2}\right)\zeta\right\| \\
&\geq \left|\left(P\Psi,\left(S-i\frac{1}{2}\right)\zeta\right)\right|
\end{aligned}$$

$$\begin{aligned}
&= \left|\left(\Psi, \left(S - i\frac{3}{2}\right) P\zeta\right)\right| \\
&= \left|\left(\Psi, \left(S - i\frac{3}{2}\right) \Psi\right)\right| \\
&\geq \frac{3}{2}
\end{aligned} \tag{2.5.4}$$

because

$$[P, S] = -iP. \tag{2.5.5}$$

In this case, S plays a useful role before it is discarded.

Next, suppose Ψ is the nth derivative of a square-integrable ζ. How is the Heisenberg inequality generalized? With

$$\Psi = P^n\zeta, \qquad \zeta \in L^2(\mathbb{R}), \tag{2.5.6}$$

we propose to obtain a lower bound on

$$\langle P^{2n}\rangle_\Psi^{1/2} \langle X^{2n}\rangle_\Psi^{1/2} = \|P^n\Psi\|_2 \|X^n\Psi\|_2. \tag{2.5.7}$$

We appeal to the identity

$$X^nP^n = \prod_{k=1}^{n} \left(S + i\left(k - \frac{1}{2}\right)\right), \tag{2.5.8}$$

which is easy enough to prove by induction. Indeed, if this identity holds for a given n, then

$$\begin{aligned}
X^{n+1}P^{n+1} &= X\left(\prod_{k=1}^{n} \left(S + i\left(k - \frac{1}{2}\right)\right)\right) P \\
&= XP\prod_{k=1}^{n} \left(S + i\left(k + \frac{1}{2}\right)\right)
\end{aligned} \tag{2.5.9}$$

because

$$\left(S + i\left(k - \frac{1}{2}\right)\right) P = P\left(S + i\left(k + \frac{1}{2}\right)\right). \tag{2.5.10}$$

Combining (1.3.37) with (2.5.9) completes the induction step. Now apply this identity as follows:

$$\begin{aligned}
\|X^n\Psi\|_2^2 &= \|X^nP^n\zeta\|_2^2 \\
&= \left\|\prod_{k=1}^{n} \left(S + i\left(k - \frac{1}{2}\right)\right)\zeta\right\|_2^2 \\
&= \left(\prod_{k=1}^{n} \left(S^2 + \left(k - \frac{1}{2}\right)^2\right)\zeta, \zeta\right)
\end{aligned}$$

$$\begin{aligned}
&\geq \|\zeta\|_2^2 \prod_{k=1}^{n}\left(k-\frac{1}{2}\right)^2 + \|S\zeta\|_2^2 \sum_{\ell=1}^{n}\prod_{k\neq\ell}\left(k-\frac{1}{2}\right)^2 \\
&= \left\|\left(\gamma_n S \pm i\prod_{k=1}^{n}\left(k-\frac{1}{2}\right)\right)\zeta\right\|_2^2, \qquad (2.5.11)
\end{aligned}$$

$$\gamma_n = \left(\sum_{\ell=1}^{n}\prod_{k\neq\ell}\left(k-\frac{1}{2}\right)^2\right)^{1/2}. \qquad (2.5.12)$$

Thus,

$$\begin{aligned}
\langle P^{2n}\rangle_\Psi^{1/2}\langle X^{2n}\rangle_\Psi^{1/2} &\geq \|P^n\Psi\|_2\left\|\left(\gamma_n S \pm i\prod_{k=1}^{n}\left(k-\frac{1}{2}\right)\right)\zeta\right\|_2 \\
&\geq \left|\left(P^n\Psi, \left(\gamma_n S \pm i\prod_{k=1}^{n}\left(k-\frac{1}{2}\right)\right)\right)\zeta\right| \\
&= \left|\left(\Psi, \left(\gamma_n S - in\gamma_n \pm i\prod_{k=1}^{n}\left(k-\frac{1}{2}\right)\right)P^n\zeta\right)\right| \\
&= \left|\left(\Psi, \left(\gamma_n S - in\gamma_n \pm i\prod_{k=1}^{n}\left(k-\frac{1}{2}\right)\right)\Psi\right)\right|, \qquad (2.5.13)
\end{aligned}$$

where we have used

$$[P^n, S] = -inP. \qquad (2.5.14)$$

Clearly, the better choice in sign is $-i$, so we may conclude that

$$\langle P^{2n}\rangle_\Psi^{1/2}\langle X^{2n}\rangle_\Psi^{1/2} \geq n\gamma_n + \prod_{k=1}^{n}\left(k-\frac{1}{2}\right). \qquad (2.5.15)$$

This is the generalization of (2.5.4) that we infer from (2.5.8).

Since (2.5.6) is a translation-invariant condition, it is trivial to extend this result to the observation that (2.5.6) implies

$$\langle P^{2n}\rangle_\Psi^{1/2}\langle (X-a)^{2n}\rangle_\Psi^{1/2} \geq n\gamma_n + \prod_{k=1}^{n}\left(k-\frac{1}{2}\right), \qquad a\in\mathbb{R}. \qquad (2.5.16)$$

Now for an arbitrary observable A, define the nth-order deviation as

$$\sigma_\Psi^{(n)}(A) = \langle (A-\langle A\rangle_\Psi)^{2n}\rangle_\Psi^{1/2n} \qquad (2.5.17)$$

Then (2.5.16) amounts to the nth-order uncertainty principle

$$\sigma_\Psi^{(n)}(P)\sigma_\Psi^{(n)}(X) \geq \left[n\gamma_n + \prod_{k=1}^{n}\left(k-\frac{1}{2}\right)\right]^{1/n} \qquad (2.5.18)$$

provided that

$$\langle P\rangle_\Psi = 0. \tag{2.5.19}$$

This additional condition is satisfied by all real-valued Ψ as well as by Ψ that are either symmetric or anti-symmetric about their barycenters $\langle X\rangle_\Psi$. For $n = 1, 2, 3$ we have:

$$\sigma_\Psi^{(1)}(P)\sigma_\Psi^{(1)}(X) \geq 3/2, \tag{2.5.20.1}$$

$$\sigma_\Psi^{(2)}(P)\sigma_\Psi^{(2)}(X) \geq \sqrt{\sqrt{10}+\frac{3}{4}} \approx \frac{95}{48}, \tag{2.5.20.2}$$

$$\sigma_\Psi^{(3)}(P)\sigma_\Psi^{(3)}(X) \geq \sqrt[3]{\frac{3}{4}\sqrt{259}+\frac{15}{8}} \approx \frac{77}{32}. \tag{2.5.20.3}$$

For large n, the lower bound (2.5.18) is essentially linear in n.

What does any of this have to do with wavelets? The point is that if Ψ is a wavelet state and if $\langle P^{2n}\rangle_\Psi$ and $\langle X^{2n}\rangle_\Psi$ are finite – otherwise, any lower bound on the product would hold vacuously – then Ψ satisfies the hypothesis of the vanishing moment theorem, with the conclusion that

$$\widehat{\Psi}^{(k)}(0) = 0, \qquad k \leq n-1. \tag{2.5.21}$$

Clearly, the condition (2.5.6) is a close relative of this vanishing moment condition. Indeed, suppose the wavelet state Ψ satisfies the slightly stronger condition that

$$|X|^{n+\delta}\Psi \in L^2(\mathbb{R}) \tag{2.5.22}$$

for some $\delta > 0$. Obviously, this includes all of the wavelets we have constructed thus far: the Meyer wavelets are Schwartz functions, the Lemarié wavelets are exponentially localized, and the Daubechies wavelets are compactly supported. Now (2.5.21) implies the Taylor formula

$$\widehat{\Psi}(p) = \frac{1}{n!}\int_0^p \widehat{\Psi}^{(n)}(p-q)q^{n-1}dq, \tag{2.5.23}$$

and we wish to infer (2.5.6) – i.e., to show that $p^{-n}\widehat{\Psi}(p)$ is square-integrable. The region $|p| \leq 1$ is obviously the issue, so we need only show that

$$\widehat{\Psi}(p) = O(|p|^{n-\frac{1}{2}+\varepsilon}), \qquad |p| \leq 1, \tag{2.5.24}$$

for some $\varepsilon > 0$. To this end, we apply the Hölder inequality – with exponent $r > 2$ – to obtain

$$\begin{aligned} |\widehat{\Psi}(p)| &\leq \int_{|q|\leq|p|} dq|\widehat{\Psi}^{(n)}(p-q)|\,|q|^{n-1}dq \\ &\leq \left(\int_{|q|\leq|p|} dq|q|^{ns-s}\right)^{1/s} \|\widehat{\Psi}^{(n)}\|_r \\ &= \left(\frac{2}{ns-s+1}\right)^{1/s} |p|^{n-1+1/s}\|\widehat{\Psi}^{(n)}\|_r, \end{aligned} \tag{2.5.25}$$

$$\frac{1}{s}+\frac{1}{r} = 1. \tag{2.5.26}$$

Since $s < 2$, we have the desired bound with

$$\varepsilon = \frac{1}{s} - \frac{1}{2}, \tag{2.5.27}$$

provided $\|\widehat{\Psi}^{(n)}\|_r < \infty$ for some r sufficiently close to 2. Now by the Young inequality,

$$\|\widehat{\Psi}^{(n)}\|_r \leq c\|X^n\Psi\|_s \tag{2.5.28}$$

but with $s < 2$, we may apply the Hölder inequality again to obtain

$$\begin{aligned} \int |x^n\Psi(x)|^s dx &\leq \int (1+|x|)^{ns}|\Psi(x)|^s dx \\ &\leq \left(\int (1+|x|)^{-2\delta s/(2-s)} dx\right)^{1-\frac{s}{2}} \\ &\quad \times \left(\int (1+|x|)^{2n+2\delta}|\Psi(x)|^2 dx\right)^{s/2} \end{aligned} \tag{2.5.29}$$

The first factor is finite, provided r is chosen such that

$$\frac{2\delta s}{2-s} = \frac{2\delta r}{r-2} > 1. \tag{2.5.30}$$

Thus we have established (2.5.24) and therefore the square-integrability of $p^{-n}\widehat{\Psi}(p)$, assuming only the vanishing moment condition and the square-integrability of $(1+|x|)^{n+\delta}\Psi(x)$ for some $\delta > 0$.

In summary, if an N-adic wavelet state Ψ satisfies the decay condition (2.5.22), then

$$\langle P^{2n}\rangle_\Psi^{1/2}\langle (X-a)^{2n}\rangle_\Psi^{1/2} \geq n\gamma_n + \prod_{k=1}^{n}\left(k - \frac{1}{2}\right), \qquad a \in \mathbb{R}, \tag{2.5.31}$$

and this inequality is an uncertainty inequality if Ψ is also a $\langle P\rangle = 0$ state (real-valued, for example).

References

1. R. Balan, "An Uncertanity Inequality for Wavelet Sets," to appear in Appl. and Comput. Harmonic Anal.

2. G. Battle, "Phase Space Localization Theorem for Ondelettes," J. Math. Phys. **30**, No. 10 (1989), 2195–2197.

3. G. Battle, "Heisenberg Inequalities for Wavelet States," Appl. and Comput. Harmonic Anal. **4** (1997), 119–146.

4. A. Messiah, *Quantum Mechanics*, Vol. I, North–Holland, Amsterdam, 1965.

5. M. Reed and B. Simon, *Methods for Modern Mathematical Physics, Vol. II, Fourier Analysis and Self-Adiointness*, Academic Press, New York, 1975.

2.6 Further Constraints of Heisenberg Type

In the previous section, the inequalities we derived essentially depended on vanishing-moment conditions only. We now derive phase space localization inequalities that are more peculiar to the wavelet state condition defined in that section. For an arbitrary N-adic wavelet state Ψ, it is interesting to measure the phase space deviations about the points

$$(x,p) = (m,0), \qquad m \in \mathbb{Z}, \tag{2.6.1}$$

in phase space – i.e., to obtain lower bounds on

$$\langle (X-m)^{2n}\rangle_\Psi^{1/2n} \langle P^{2n}\rangle_\Psi^{1/2n} = \|(X-m)^n\Psi\|_2^{1/n} \|P^n\Psi\|_2^{1/n} \tag{2.6.2}$$

Since every integer translate of an N-adic wavelet state is still an N-adic wavelet state, we may set $m = 0$ without loss.

Consider the $n = 1$ case and notice the effect of the orthogonality property on the expectation of the scaling group. We have

$$\begin{aligned} \frac{d}{d\lambda}(e^{\lambda/2}\langle e^{-i\lambda S}\rangle_\Psi) &= -ie^{\lambda/2}\left\langle \left(S + \frac{1}{2}i\right) e^{-i\lambda S}\right\rangle_\Psi \\ &= -ie^{\lambda/2}\langle XPe^{-i\lambda S}\rangle_\Psi. \end{aligned} \tag{2.6.3}$$

Combining this with

$$e^{i\lambda S}Pe^{-i\lambda S} = e^{-\lambda}P \tag{2.6.4}$$

and the fundamental theorem of calculus, one obtains

$$\sqrt{N}\,\langle e^{-i(\ln N)S}\rangle_\Psi - 1 = -i\int_0^{\ln N} e^{-i\lambda/2}\langle Xe^{-i\lambda S}P\rangle_\Psi d\lambda \tag{2.6.5}$$

for an arbitrary state. For an N-adic wavelet state,

$$\langle e^{-ir(\ln N)S}\rangle_\Psi = 0, \qquad r \in \mathbb{Z}\backslash\{0\}, \tag{2.6.6}$$

so set $r = 1$. By the Schwarz inequality,

$$\begin{aligned} 1 &\le \|P\Psi\|_2\|X\Psi\|_2 \int_0^{\ln N} e^{-\lambda/2}d\lambda \\ &= 2(1-1/\sqrt{N})\|P\Psi\|_2\|X\Psi\|_2 \end{aligned} \tag{2.6.7}$$

for such a state Ψ. Thus we have the lower bound

$$\|P\Psi\|_2\|X\Psi\|_2 \ge \frac{1}{2}(1-1/\sqrt{N})^{-1}. \tag{2.6.8}$$

In contrast to the last section, we have not assumed the decay condition that $|X|^{1+\delta}\Psi$ be square-integrable for some $\delta > 0$.

The generalization of this estimation to arbitrary n is an application of Rolle's Theorem in elementary calculus together with the property

$$\mathrm{Re}\langle e^{-ir(\ln N)S}\rangle_\Psi = \mathrm{Im}\langle e^{-ir(\ln N)S}\rangle_\Psi = 0 \tag{2.6.9}$$

for all non-zero integers r. First, observe that

$$e^{-\lambda}\left(e^{\lambda}\frac{d}{d\lambda}\right)^n (e^{\lambda/2}\langle e^{-i\lambda S}\rangle_\Psi) = (-i)^n e^{(n-\frac{1}{2})\lambda}\left\langle \prod_{k=1}^{n}\left(S+i\left(k-\frac{1}{2}\right)\right)e^{-i\lambda S}\right\rangle_\Psi$$

$$= (-i)^n e^{(n-\frac{1}{2})\lambda}\langle X^n P^n e^{-i\lambda S}\rangle_\Psi. \qquad (2.6.10)$$

On the other hand, (2.6.4) obviously generalizes to

$$e^{i\lambda S}P^n e^{-i\lambda S} = e^{-n\lambda}P^n, \qquad (2.6.11)$$

so we have

$$e^{-\lambda}\left(e^{\lambda}\frac{d}{d\lambda}\right)^n (e^{\lambda/2}\langle e^{-i\lambda S}\rangle_\Psi) = (-i)^n e^{-\lambda/2}\langle X^n e^{-i\lambda S}P^n\rangle_\Psi. \qquad (2.6.12)$$

Next, observe that the operator

$$e^{-\lambda}\left(e^{\lambda}\frac{d}{d\lambda}\right)^k$$

acts separately on the real and imaginary parts of any function of λ. Thus

$$(-i)^n e^{-\lambda/2}\langle X^n e^{-i\lambda S}P^n\rangle_\Psi = u_n(\lambda)+iv_n(\lambda) \qquad (2.6.13)$$

if we define

$$u_k(\lambda) = e^{-\lambda}\left(e^{\lambda}\frac{d}{d\lambda}\right)^k u(\lambda), \qquad (2.6.14.0)$$

$$v_k(\lambda) = e^{-\lambda}\left(e^{\lambda}\frac{d}{d\lambda}\right)^k v(\lambda), \qquad (2.6.14.1)$$

$$e^{\lambda/2}\langle e^{-i\lambda S}\rangle_\Psi = u(\lambda)+iv(\lambda). \qquad (2.6.15)$$

The property (2.6.6) implies

$$u(r\ln N) = v(r\ln N) = 0 \qquad (2.6.16)$$

for all non-zero integers r. In particular,

$$\int_0^{\ln N} d\lambda_1 u_1(\lambda_1) = -u(0) = -1, \qquad (2.6.17.0)$$

$$\int_0^{\ln N} d\lambda_1 v_1(\lambda_1) = -v(0) = 0. \qquad (2.6.17.1)$$

Now observe that by Rolle's Theorem, there are sequences $(\alpha_1^{(1)},\alpha_2^{(1)},\alpha_3^{(1)},\ldots)$ and $(\beta_1^{(1)},\beta_2^{(1)},\beta_3^{(1)},\ldots)$ such that

$$r\ln N < \alpha_r^{(1)},\ \beta_r^{(1)} < (r+1)\ln N, \qquad (2.6.18)$$

$$u_1(\alpha_r^{(1)}) = 0, \tag{2.6.19.0}$$

$$v_1(\beta_r^{(1)}) = 0. \tag{2.6.19.1}$$

Thus

$$\int_{\lambda_1}^{\alpha_1^{(1)}} d\lambda_2 u_2(\lambda_2) = -e^{\lambda_1} u_1(\lambda_1), \tag{2.6.20.0}$$

$$\int_{\lambda_1}^{\beta_1^{(1)}} d\lambda_2 v_2(\lambda_2) = -e^{\lambda_1} v_1(\lambda_1). \tag{2.6.20.1}$$

Now apply Rolle's Theorem again – this time, to these zeros of $e^\lambda u_1(\lambda)$ and $e^\lambda v_1(\lambda)$ – to obtain sequences $(\alpha_1^{(2)}, \alpha_2^{(2)}, \alpha_3^{(2)}, \ldots)$ and $(\beta_1^{(2)}, \beta_2^{(2)}, \beta_3^{(2)}, \ldots)$ where

$$\alpha_r^{(1)} < \alpha_r^{(2)} < \alpha_{r+1}^{(1)}, \tag{2.6.21}$$

$$u_2(\alpha_r^{(2)}) = 0, \tag{2.6.22}$$

$$\beta_r^{(1)} < \beta_r^{(2)} < \beta_{r+1}^{(1)}, \tag{2.6.23}$$

$$v_2(\beta_r^{(2)}) = 0. \tag{2.6.24}$$

Clearly,

$$\int_{\lambda_2}^{\alpha_1^{(2)}} d\lambda_3 u_3(\lambda_3) = -e^{\lambda_2} u_2(\lambda_2), \tag{2.6.25.0}$$

$$\int_{\lambda_2}^{\beta_1^{(2)}} d\lambda_3 v_3(\lambda_3) = -e^{\lambda_2} v_2(\lambda_2). \tag{2.6.25.1}$$

Obviously, we can iterate this procedure for as long as derivatives exist. In general, we have sequences $(\alpha_1^{(k)}, \alpha_2^{(k)}, \alpha_3^{(k)}, \ldots)$ and $(\beta_1^{(k)}, \beta_2^{(k)}, \beta_3^{(k)}, \ldots)$ satisfying the conditions

$$\alpha_r^{(k-1)} < \alpha_r^{(k)} < \alpha_{r+1}^{(k-1)}, \tag{2.6.26}$$

$$u_k(\alpha_r^{(k)}) = 0, \tag{2.6.27}$$

$$\beta_r^{(k-1)} < \beta_r^{(k)} < \beta_{r+1}^{(k-1)}, \tag{2.6.28}$$

$$v_k(\beta_r^{(k)}) = 0, \tag{2.6.29}$$

and the integration formulas are

$$\int_{\lambda_k}^{\alpha_1^{(k)}} d\lambda_{k+1} u_{k+1}(\lambda_{k+1}) = -e^{\lambda_k} u_k(\lambda_k), \tag{2.6.30.0}$$

$$\int_{\lambda_k}^{\beta_1^{(k)}} d\lambda_{k+1} v_{k+1}(\lambda_{k+1}) = -e^{\lambda_k} v_k(\lambda_k). \tag{2.6.30.1}$$

Pulling this together, one obtains

$$(-1)^n = \int_0^{\ln N} d\lambda_1 e^{-\lambda_1} \int_{\lambda_1}^{\alpha_1^{(1)}} d\lambda_2 e^{-\lambda_2} \cdots \int_{\lambda_{n-1}}^{\alpha_1^{(n-1)}} d\lambda_n u_n(\lambda_n), \tag{2.6.31.0}$$

$$0=\int_0^{\ln N} d\lambda_1 e^{-\lambda_1}\int_{\lambda_1}^{\beta_1^{(1)}} d\lambda_2 e^{-\lambda_2}\cdots\int_{\lambda_{n-1}}^{\beta_1^{(n-1)}} d\lambda_n v_n(\lambda_n). \tag{2.6.31.1}$$

The second iterated integral equation is useless; the imaginary part of (2.6.13) contributes nothing to our lower bound. We focus on the first equation and notice that the smallest numbers $\alpha_1^{(k)}$ in the sequences have the ordering

$$\ln N<\alpha_1^{(1)}<\alpha_1^{(2)}<\cdots<\alpha_1^{(n-1)}. \tag{2.6.32}$$

Moreover,

$$\alpha_1^{(k)}<\alpha_2^{(k-1)}<\alpha_3^{(k-2)}<\cdots<\alpha_k^{(1)}<(k+1)\ln N, \tag{2.6.33}$$

so we have the estimate

$$1\le\int_0^{\ln N} d\lambda_1 e^{-\lambda_1}\int_{\lambda_1}^{2\ln N} d\lambda_2 e^{-\lambda_2}\int_{\lambda_2}^{3\ln N} d\lambda_3 e^{-\lambda_3}\cdots\int_{\lambda_{n-1}}^{n\ln N} d\lambda_n|u_n(\lambda_n)|. \tag{2.6.34}$$

Finally, we apply (2.6.13):

$$|u_n(\lambda)|\le e^{-\lambda/2}\|P^n\Psi\|_2\|X^n\Psi\|_2. \tag{2.6.35}$$

We obtain

$$\begin{aligned} 1 &\le \|P^n\Psi\|_2\|X^n\Psi\|_2\int_0^{\ln N} d\lambda_1 e^{-\lambda_1}\int_1^{2\ln N} d\lambda_2 e^{-\lambda_2}\cdots\int_{\lambda_{n-1}}^{n\ln N} d\lambda_n e^{-\lambda_n/2}\\ &= \|P^n\Psi\|_2\|X^n\Psi\|_2\int_{N^{-1}}^{1} d\tau_1\int_{N^{-2}}^{\tau_1} d\tau_2\cdots\int_{N^{-n}}^{\tau_{n-1}}\frac{d\tau_n}{\sqrt{\tau_n}}, \end{aligned} \tag{2.6.36}$$

and so we have the lower bound

$$\langle P^{2n}\rangle_\Psi^{1/2n}\langle X^{2n}\rangle_\Psi^{1/2n}\ge\left[\int_{N^{-1}}^{1} d\tau_1\int_{N^{-2}}^{\tau_1} d\tau_2\cdots\int_{N^{-n}}^{\tau_{n-1}}\frac{d\tau_n}{\sqrt{\tau_n}}\right]^{-1/n}. \tag{2.6.37}$$

We have derived this inequality without the assumption that $|X|^{n+\delta}\Psi$ be square-integrable for some $\delta>0$.

Notice also that this lower bound can be estimated independently of the basic scale factor N:

$$\begin{aligned} \int_{N^{-1}}^{1} d\tau_1\int_{N^{-2}}^{\tau} d\tau_2\cdots\int_{N^{-n}}^{\tau_{n-1}}\frac{d\tau_n}{\sqrt{\tau_n}} &\le \int_0^1 d\tau_1\int_0^{\tau_1} d\tau_2\cdots\int_0^{\tau_{n-1}}\frac{d\tau_n}{\sqrt{\tau_n}}\\ &= \prod_{k=1}^{n}\frac{1}{k-\frac{1}{2}}, \end{aligned} \tag{2.6.38}$$

so we also have the weaker bound

$$\langle P^{2n}\rangle_\Psi^{1/2n}\langle X^{2n}\rangle_\Psi^{1/2n}\ge\prod_{k=1}^{n}\left(k-\frac{1}{2}\right)^{1/n} \tag{2.6.39}$$

On the other hand, for the dyadic ($N = 2$) case, the bounds for $n = 1, 2, 3$ are:

$$\langle P^2\rangle_\Psi^{1/2}\langle X^2\rangle_\Psi^{1/2} \geq \frac{1}{2-\sqrt{2}} = \frac{1}{2}(2+\sqrt{2}), \tag{2.6.40.1}$$

$$\langle P^4\rangle_\Psi^{1/4}\langle X^4\rangle_\Psi^{1/4} \geq \sqrt{\frac{6}{5-2\sqrt{2}}} = \sqrt{\frac{6}{17}(5+2\sqrt{2})}, \tag{2.6.40.2}$$

$$\langle P^6\rangle_\Psi^{1/6}\langle X^6\rangle_\Psi^{1/6} \geq \sqrt[3]{\frac{120}{54-23\sqrt{2}}} = \sqrt[3]{\frac{60}{929}(54+23\sqrt{2})}. \tag{2.6.40.3}$$

This case is of interest because the wavelets that are routinely constructed are dyadic unless one desires a special property that requires N to have some other value.

How can the estimate (2.6.37) enhance the previous lower bound derived from the assumption that $|X|_\Psi^{n+\delta}$ is square-integrable? In the $n = 1$ case, one certainly has the estimate

$$\langle P^2\rangle_\Psi^{1/2}\langle X^2\rangle_\Psi^{1/2} \geq \frac{3}{2} \tag{2.6.41}$$

from the previous section. To date, no N-dependent enhancement of this lower bound has been found – only the N-dependent enhancement (2.6.8) of the universal lower bound. In this lowest-order case, we can only compare the lower bounds. Clearly, (2.6.40.1) is stronger than (2.6.41) but for $N \geq 3$, (2.6.41) is stronger than (2.6.8).

For $n > 1$ we have more interesting estimates. In these higher-order cases, we *can* derive an N-dependent enhancement of the corresponding lower bounds derived from this decay in the last section. Consider the $n = 2$ case first: with $|X|^{2+\delta}\Psi$ assumed to be square-integrable, we may infer from the previous section that

$$\Psi = P^2\zeta \tag{2.6.42}$$

for some square-integrable ζ. Thus

$$\begin{aligned}
\|X^2\Psi\|_2^2 &= \|X^2P^2\zeta\|_2^2 \\
&= \left\|\left(S+i\frac{1}{2}\right)\left(S+i\frac{3}{2}\right)\zeta\right\|_2^2 \\
&= \left(\left(S^2+\frac{1}{4}\right)\left(S^2+\frac{9}{4}\right)\zeta,\zeta\right) \\
&= (S^4\zeta,\zeta)+\frac{5}{2}(S^2\zeta,\zeta)+\frac{9}{16}(\zeta,\zeta) \\
&\geq 4\frac{1}{\lambda^4}((1-\cos\lambda S)^2\zeta,\zeta)+\frac{5}{2}(S^2\zeta,\zeta)+\frac{9}{16}(\zeta,\zeta) \\
&= 4\frac{1}{\lambda^4}\|(1-\cos\lambda S)\zeta\|_2^2+\left\|\left(\frac{1}{2}\sqrt{10}\,S-i\frac{3}{4}\right)\zeta\right\|_2^2,
\end{aligned} \tag{2.6.43}$$

where we have used the operator inequality

$$\frac{1}{2}\lambda^2S^2 \geq 1-\cos\lambda S. \tag{2.6.44}$$

Therefore,

$$\begin{aligned}
\|P^2\Psi\|_2^2\|X^2\Psi\|_2^2 &\geq 4\frac{1}{\lambda^4}|(P^2\Psi,(1-\cos\lambda S)\zeta)|^2 \\
&+ \left|\left(P^2\Psi,\left(\frac{1}{2}\sqrt{10}\,S - i\frac{3}{4}\right)\zeta\right)\right|^2 \\
&= 4\frac{1}{\lambda^4}|(\Psi,(1-\cos\lambda(S-i2))P^2\zeta)|^2 \\
&+ \left|\left(\Psi,\left(\frac{1}{2}\sqrt{10}\,S - i\sqrt{10} - i\frac{3}{4}\right)P^2\zeta\right)\right|^2 \\
&= 4\frac{1}{\lambda^4}\left|\left(\Psi,\left(1-\frac{1}{2}e^{2\lambda}e^{i\lambda S}-\frac{1}{2}e^{-2\lambda}e^{-i\lambda S}\right)\Psi\right)\right|^2 \\
&+ \frac{5}{2}|(\Psi,S\Psi)|^2 + \left(\sqrt{10}+\frac{3}{4}\right)^2 \qquad (2.6.45)
\end{aligned}$$

Now set $\lambda = \ln N$ and apply (2.6.6) to obtain

$$\|P^2\Psi\|_2^2\|X^2\Psi\|_2^2 \geq \frac{4}{(\ln N)^4} + \frac{3}{2}\sqrt{10} + \frac{169}{16}, \qquad (2.6.46)$$

so we have

$$\langle P^4\rangle_\Psi^{1/4}\langle X^4\rangle_\Psi^{1/4} \geq \left(\frac{3}{2}\sqrt{10} + \frac{169}{16} + \frac{4}{(\ln N)}\right)^{1/4}. \qquad (2.6.47)$$

Notice that without the N-dependent term, this inequality would be given by the $n = 2$ case of (5.2.15), so we have enhanced the latter inequality with a dependence on the basic scale factor N. How does this new inequality compare to the $n = 2$ case of (2.6.37)? It is stronger regardless of the value of N. Indeed

$$\begin{aligned}
\left[\int_{N^{-1}}^{1} d\tau_1 \int_{N^{-2}}^{\tau_1} \frac{d\tau_2}{\sqrt{\tau_2}}\right]^{-1} &\leq \left[\int_{1/2}^{1} d\tau_1 \int_{1/4}^{\tau_1} \frac{d\tau_2}{\sqrt{\tau_2}}\right]^{-1} \\
&= \frac{6}{17}(5+2\sqrt{2}) \\
&< \sqrt{10} + \frac{3}{4}, \qquad (2.6.48)
\end{aligned}$$

so actually, the $n = 2$ case of (5.2.15) is itself stronger for all N. The advantage of (2.6.37) lies in the absence of the decay assumption.

We now derive the N-dependent lower bound for $n > 2$ with the decay assumption that $|X|^{n+\delta}\Psi$ is square-integrable for some $\delta > 0$. Let ζ be the square-integrable function such that

$$\Psi = P^n\zeta, \qquad (2.6.49)$$

and generalize the algebra in (2.6.43) as follows:

$$\|X^n\Psi\|_2^2 = \|X^nP^n\zeta\|_2^2$$

$$
\begin{aligned}
&= \left\| \prod_{k=1}^{n} \left(S + i\left(k - \frac{1}{2}\right)\right) \zeta \right\|_2 \\
&= \left(\prod_{k=1}^{n} \left(S^2 + \left(k - \frac{1}{2}\right)^2 \right) \zeta, \zeta \right) \\
&= \sum_{\ell=0}^{n} a_\ell^{(n)} (S^{2\ell}\zeta, \zeta),
\end{aligned}
\tag{2.6.50}
$$

with the coefficients $a_\ell^{(n)} > 0$ defined by

$$
\prod_{k=1}^{n} \left(z + \left(k - \frac{1}{2}\right)^2 \right) = \sum_{\ell=0}^{n} a_\ell^{(n)} z^\ell. \tag{2.6.51}
$$

Now for $\ell = 2m$ (even ℓ) we apply (2.6.44) with $\lambda = \ln N$ to obtain

$$
(S^{4m}\zeta, \zeta) \geq 4^m \frac{1}{(\ln N)^{4m}} ((1 - \cos(\ln N)S)^{2m}\zeta, \zeta), \tag{2.6.52}
$$

while for $\ell = 2m + 1$ (odd ℓ) we apply (2.6.44) as follows:

$$
(S^{4m+2}\zeta, \zeta) \geq 4^m \frac{1}{\lambda_m^{4m}} (S^2(1 - \cos\lambda_m S)^{2m}\zeta, \zeta), \tag{2.6.53}
$$

where the value of λ_m will be chosen below.

Consider the case where n itself is odd, and pair consecutive terms in the following way:

$$
\begin{aligned}
\|X^n\Psi\|_2^2 &\geq \sum_{m=0}^{\frac{1}{2}(n-1)} 4^m \left[\frac{a_{2m}^{(n)}}{(\ln N)^{4m}} ((1 - \cos(\ln N)S)^{2m}\zeta, \zeta) \right. \\
&\quad \left. + \frac{a_{2m+1}^{(n)}}{\lambda_m^{4m}} (S^2(1 - \cos\lambda_m S)^{2m}\zeta, \zeta) \right] \\
&= \sum_{m=0}^{\frac{1}{2}(n-1)} 4^m \left\| \left(\frac{1}{\lambda_m^{2m}} \sqrt{a_{2m+1}^{(n)}} (1 - \cos\lambda_m S)^m S \right. \right. \\
&\quad \left. \left. - i \frac{1}{(\ln N)^{2m}} \sqrt{a_{2m}^{(n)}} (1 - \cos(\ln N)S)^m \right) \zeta \right\|_2^2.
\end{aligned}
\tag{2.6.54}
$$

Hence

$$
\begin{aligned}
&\|P^n\Psi\|_2^2 \|X^n\Psi\|_2^2 \\
&\geq \sum_{m=0}^{\frac{1}{2}(n-1)} 4^m \left| \left(P^n\Psi, \frac{1}{\lambda_m^{2m}} \sqrt{a_{2m+1}^{(n)}} (1 - \cos\lambda_m S)^m S\zeta \right. \right.
\end{aligned}
$$

$$- i\frac{1}{(\ln N)^{2m}}\sqrt{a_{2m}^{(n)}}(1 - \cos(\ln N)S)^m \zeta \Big) \Big|^2$$

$$= \sum_{m=0}^{\frac{1}{2}(n-1)} 4^m \Big| \Big(\Psi, \frac{1}{\lambda_m^{2m}}\sqrt{a_{2m+1}^{(n)}}(1 - \cos\lambda_m(S - in))^m$$

$$\times (S - in)P^n\zeta - i\frac{1}{(\ln N)^{2m}}\sqrt{a_{2m}^{(n)}}(1 - \cos(\ln N)(S - in))^m P^n\zeta \Big) \Big|^2$$

$$= \sum_{m=0}^{\frac{1}{2}(n-1)} 4^m \Big| \Big(\Psi, \frac{1}{\lambda_m^{2m}}\sqrt{a_{2m+1}^{(n)}} \Big(1 - \frac{1}{2}e^{i\lambda_m(S-in)}$$

$$- \frac{1}{2}e^{-i\lambda_m(S-in)}\Big)^m (S - in)\Psi - i\frac{1}{(\ln N)^{2m}}\sqrt{a_{2m}^{(n)}}$$

$$\times \Big(1 - \frac{1}{2}N^n e^{i(\ln N)S} - \frac{1}{2}N^{-n}e^{-i(\ln N)S}\Big)^m \Psi \Big) \Big|^2. \tag{2.6.55}$$

Now notice that the expansion

$$\Big(1 - \frac{1}{2}e^{i\lambda(S-in)} - \frac{1}{2}e^{-i\lambda(S-in)}\Big)^m = \sum_{\mu=0}^{m} \binom{m}{\mu} \Big(-\frac{1}{2}\Big)^\mu \sum_{q=0}^{\mu} \binom{\mu}{q} \times e^{i(\mu-2q)\lambda(S-in)} \tag{2.6.56}$$

together with (2.6.6) implies

$$\Big(\Psi, \Big(1 - \frac{1}{2}N^n e^{i(\ln N)S} - \frac{1}{2}N^{-n}e^{-i(\ln N)S}\Big)^m \Psi\Big) = \sum_{q=0}^{[m/2]} 4^{-q}\binom{m}{2q}\binom{2q}{q}. \tag{2.6.57}$$

Define

$$w_m(\lambda) = \sum_{\mu=0}^{m} \binom{m}{\mu}\Big(-\frac{1}{2}\Big)^\mu \sum_{\substack{q=0 \\ 2q\neq\mu}}^{\mu} \binom{\mu}{q} \frac{i}{\mu - 2q}(\Psi, e^{i(\mu-2q)\lambda(S-in)}\Psi) \tag{2.6.58}$$

and observe that our inequality may be written in the form

$$\|P^n\Psi\|_2^2\|X^n\Psi\|_2^2 \geq \sum_{m=0}^{\frac{1}{2}(n-1)} 4^m \Big| \frac{1}{\lambda_m^{2m}}\sqrt{a_{2m+1}^{(n)}}w_m'(\lambda_m)$$

$$+ \frac{1}{\lambda_m^{2m}}\sqrt{a_{2m+1}^{(n)}} \sum_{q=0}^{[m/2]} 4^{-q}\binom{m}{2q}\binom{2q}{q}(\Psi, (S - in)\Psi)$$

$$- i\frac{1}{(\ln N)^{2m}}\sqrt{a_{2m}^{(n)}} \sum_{q=0}^{[m/2]} 4^{-q}\binom{m}{2q}\binom{2q}{q} \Big|^2 \tag{2.6.59}$$

The significance of $w_m(\lambda)$ is the wavelet property

$$w_m(\ln N) = w_m(2\ln N) = 0. \tag{2.6.60}$$

Applying Rolle's Theorem to Im $w_m(\lambda)$, we may choose λ_m such that

$$\ln N < \lambda_m < 2\ln N \tag{2.6.61}$$
$$\text{Im } w_m'(\lambda_m) = 0. \tag{2.6.62}$$

For such a choice of λ_m, we obtain an explicit lower bound by discarding the real part of the modulus, which includes the contribution by the real part of $w_m'(\lambda_m)$. Thus

$$\begin{aligned}
\|P^n\Psi\|_2^2\|X^n\Psi\|_2^2 &\geq \sum_{m=0}^{\frac{1}{2}(n-1)} 4^m \left(\frac{n}{\lambda_m^{2m}}\sqrt{a_{2m+1}^{(n)}} + \frac{1}{(\ln N)^{2m}}\sqrt{a_{2m}^{(n)}}\right)^2 \\
&\quad \times \left(\sum_{q=0}^{[m/2]} 4^{-q}\binom{m}{2q}\binom{2q}{q}\right)^2 \\
&\geq \sum_{m=0}^{\frac{1}{2}(n-1)} \frac{1}{(\ln N)^{4m}} \left(\frac{n}{2^m}\sqrt{a_{2m+1}^{(n)}} + 2^m\sqrt{a_{2m}^{(n)}}\right)^2 \\
&\quad \times \left(\sum_{q=0}^{[m/2]} 4^{-q}\binom{m}{2q}\binom{2q}{q}\right)^2
\end{aligned} \tag{2.6.63}$$

In the case where n is even, the nth term is not paired with anything, and our estimation yields:

$$\begin{aligned}
&\|P^n\Psi\|_2^2\|X^n\Psi\|_2^2 \\
&\geq \sum_{m=0}^{\frac{1}{2}n-1} \frac{1}{(\ln N)^{4m}} \left(\frac{n}{2^m}\sqrt{a_{2m+1}^{(n)}} + 2^m\sqrt{a_{2m}^{(n)}}\right)^2 \left(\sum_{q=0}^{[m/2]} 4^{-q}\binom{m}{2q}\binom{2q}{q}\right)^2 \\
&\quad + \frac{2^n}{(\ln N)^{2n}} \left(\sum_{q=0}^{[n/4]} 4^{-q}\binom{\frac{1}{2}n}{2q}\binom{2q}{q}\right)^2
\end{aligned} \tag{2.6.64}$$

In both cases, we recover the lower bound (5.2.15) in the $N = \infty$ limit.

In the $n = 3$ case, (2.6.63) is applicable and the m-summation consists of an $m = 0$ term and an $m = 1$ term. We have

$$\begin{aligned}
\|P^3\Psi\|_2^2\|X^3\Psi\|_2^2 &\geq \left(3\sqrt{a_1^{(3)}} + \sqrt{a_0^{(3)}}\right)^2 + \frac{1}{(\ln N)^4} \\
&\quad \times \left(\frac{3}{2}\sqrt{a_3^{(3)}} + 2\sqrt{a_2^{(3)}}\right)^2,
\end{aligned} \tag{2.6.65}$$

as the q-summation consists of only the $q = 0$ term. It is easily seen that

$$\left.\begin{aligned} a_0^{(3)} &= \frac{225}{64}, \\ a_1^{(3)} &= \frac{259}{16}, \\ a_2^{(3)} &= \frac{35}{4}, \\ a_3^{(3)} &= 1, \end{aligned}\right\} \left(z+\frac{1}{4}\right)\left(z+\frac{9}{4}\right)\left(z+\frac{25}{4}\right) = \sum_{\ell=0}^{3} a_\ell^{(3)} z^\ell, \tag{2.6.66}$$

and therefore

$$\langle P^6\rangle_\Psi^{1/6}\langle X^6\rangle_\Psi^{1/6} \geq \left[\frac{45}{16}\sqrt{259} + \frac{9549}{64} + \frac{1}{(\ln N)^4}\left(3\sqrt{35} + \frac{149}{4}\right)\right]^{1/6}. \tag{2.6.67}$$

Since $\ln 2 < \sqrt{2}$, the dyadic case enhances the $n = 3$ case of (5.2.15) considerably.

In the $n = 4$ case, (2.6.64) is applicable, but the m-summation still consists of only an $m = 0$ term and an $m = 1$ term. We have

$$\begin{aligned} \|P^4\Psi\|_2^2\|X^4\Psi\|_2^2 \quad \geq \quad & \left(4\sqrt{a_1^{(4)}} + \sqrt{a_0^{(4)}}\right)^2 + \frac{1}{(\ln N)^4} \\ & \times \left(2\sqrt{a_3^{(4)}} + 2\sqrt{a_2^{(4)}}\right)^2 + \frac{16}{(\ln N)^8}\left(1 + \frac{1}{4}\binom{2}{1}\right), \end{aligned} \tag{2.6.68}$$

where the q-summation in the last contribution consists of both a $q = 0$ term and a $q = 1$ term. The coefficients $a_\ell^{(4)}$ are:

$$\left.\begin{aligned} a_0^{(4)} &= \frac{11025}{256}, \\ a_1^{(4)} &= \frac{3229}{16}, \\ a_2^{(4)} &= \frac{987}{8}, \\ a_3^{(4)} &= 21, \\ a_4^{(4)} &= 1, \end{aligned}\right\} \left(z+\frac{1}{4}\right)\left(z+\frac{9}{4}\right)\left(z+\frac{25}{4}\right)\left(z+\frac{49}{4}\right) = \sum_{\ell=0}^{4} a_\ell^{(4)} z^\ell, \tag{2.6.69}$$

and so we have:

$$\begin{aligned} &\langle P^8\rangle_\Psi^{1/8}\langle X^8\rangle_\Psi^{1/8} \\ &\geq \left[\frac{105}{8}\sqrt{3229} + \frac{51769}{16} + \frac{1}{(\ln N)^4}\left(\frac{21}{2}\sqrt{94} + \frac{1155}{8}\right) + 24\frac{1}{(\ln N)^8}\right]^{1/8} \end{aligned} \tag{2.6.70}$$

Notice that in the dyadic case, this lower bound is drastically larger than the lower bound provided by (5.2.15).

References

1. G. Battle, "Heisenberg Inequalities for Wavelet States," Appl. and Comput. Harmonic Anal. **4** (1997), 119–146.

2. A. Messiah, *Quantum Mechanics*, Vol. I, North–Holland, Amsterdam, 1965.

2.7 Variational Construction

We have already described the Meyer–Mallat construction of various wavelets, and their approach is the most widely accepted one. We now describe a variational method, which is actually more natural to the renormalization group formalism of Chap. 4. It is also the natural way to construct multi-dimensional Sobolev wavelets in §2.10. On the other hand, this method of construction cannot produce Daubechies wavelets.

Our starting point is a scaling function η whose integer translates are not necessarily orthogonal. The first and major step in the variational approach is to construct a function $\psi \in L^2(\mathbb{R})$ such that

$$\psi(2^r x - n) \perp \psi(x), \qquad n \in \mathbb{Z}, r \in \mathbb{Z}\backslash\{0\}, \tag{2.7.1}$$

in that Hilbert space – i.e., a dyadic wavelet as defined in the introductory discussion of this chapter. If the set generated spans the Hilbert space, then a mother wavelet Ψ for an orthonormal wavelet basis is given by

$$\widehat{\Psi}(p) = \left(\sum_{\ell \in \mathbb{Z}} |\widehat{\psi}(p + 2\pi\ell)|^2 \right)^{-1/2} \widehat{\psi}(p), \tag{2.7.2}$$

i.e., the standard $\mathbb{Z}$-translation-covariant orthogonalization of the set of unit-scale basis functions. In position space, this is nothing more than the application of the inverse square root of the overlap matrix to the integer translates of ψ, and the same operation was applied directly to the integer translates of the scaling function η as a first step in the Meyer–Mallat construction.

The function ψ to be constructed is first *defined* as the solution of a constrained minimization problem: ψ minimizes the L^2-norm $\|\zeta\|_2$ with respect to the constraints

$$\int \eta(2x - m)\zeta(x)dx = (-1)^m c_{1-m}, \tag{2.7.3}$$

where $\{c_m\}$ is the sequence of coefficients in the scaling relation (2.2.1). The driving principle is quite simple. The linear constraints define a closed *affine* subspace of $L^2(\mathbb{R})$ – closed, because the linear functionals are bounded – so ψ exists and is unique; it is the point in the affine subspace nearest to the origin. Therefore ψ is orthogonal to every vector in the *linear* subspace associated with the affine subspace – i.e., if

$$\int \eta(2x - m)\zeta(x)dx = 0 \tag{2.7.4}$$

for all m, then $\psi \perp \zeta$ in the Hilbert space. How is this related to the scales? The point is that for $r > 0$,

$$\begin{aligned}
&\int \eta(2x-m)\psi(2^r x-n)dx \\
= \;& 2^{-r}\int \eta(2^{-r+1}x - m + 2^{-r+1}n)\psi(x)dx \\
= \;& 2^{-r}\sum_{m_1} c_{m_1}\sum_{m_2} c_{m_2}\cdots\sum_{m_r} c_{m_r}\int \eta(2x - 2^r m + 2n \\
& -2^{r-1}m_1 - 2^{r-2}m_2 - \cdots - m_r)\psi(x)dx \\
= \;& 2^{-r}\sum_{m_1} c_{m_1}\sum_{m_2} c_{m_2}\cdots\sum_{m_r} c_{m_r}(-1)^{m_r} \\
& \times c_{1-2^r m+2n-2^{r-1}m_1-2^{r-2}m_2-\cdots-m_r},
\end{aligned} \tag{2.7.5}$$

and the transformation $m_1 \mapsto m_1, m_2 \mapsto m_2, \ldots, m_r \mapsto 1-2^r m+2n-2^{r-1}m_1-\cdots-m_r$ of summation indices has the effect of reversing the sign of this multiple sum. Therefore, the sum vanishes, and so

$$\int \eta(2x-m)\psi(2^r x-n)dx = 0, \qquad m,n \in \mathbb{Z}, r > 0. \tag{2.7.6}$$

This establishes the interscale orthogonality property by the argument above.

Now since ψ minimizes a quadratic form – namely $\|\zeta\|_2^2 = (\zeta,\zeta)$ – with respect to linear constraints of the form

$$(\zeta, g_m) = d_m, \tag{2.7.7}$$

ψ has the formal representation

$$\psi = \lim_{\alpha\to\infty} \alpha \sum_n d_n(1+\alpha K)^{-1}g_n, \tag{2.7.8}$$

$$K\zeta = \sum_m (\zeta, g_m)g_m. \tag{2.7.9}$$

On the other hand, K can be explicitly realized in momentum space if we set

$$g_m(x) = 2\eta(2x-m). \tag{2.7.10}$$

Indeed, by the Poisson summation formula, we get

$$\begin{aligned}
\widehat{K\zeta}(p) &= \frac{1}{2\pi}\sum_m \widehat{g_m}(p)\int \hat{\zeta}(q)\widehat{g_m}(q)^* dq \\
&= \frac{1}{2\pi}\sum_m \hat{\eta}\left(\frac{1}{2}p\right)e^{-i\frac{1}{2}mp}\int \hat{\zeta}(q)\hat{\eta}\left(\frac{1}{2}q\right)^* e^{i\frac{1}{2}mq}dq \\
&= \hat{\eta}\left(\frac{1}{2}p\right)\sum_\ell \int \hat{\zeta}(q)\hat{\eta}\left(\frac{1}{2}q\right)^* \delta\left(\frac{1}{2}p - \frac{1}{2}q + 2\pi\ell\right)dq \\
&= \hat{\eta}\left(\frac{1}{2}p\right)\sum_\ell \hat{\zeta}(p+4\pi\ell)\hat{\eta}\left(\frac{1}{2}p+2\pi\ell\right)^*
\end{aligned} \tag{2.7.11}$$

Thus, an arbitrary power of the operator yields

$$\widehat{K^n\zeta}(p) = \widehat{K\zeta}(p)\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell\right)\right|^2\right)^{n-1}, \tag{2.7.12}$$

and so it is trivial to verify that

$$(1+\widehat{\alpha K)^{-1}}\zeta(p) = \hat{\zeta}(p) - \alpha\left(1+\alpha\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell\right)\right|^2\right)^{-1}\widehat{K\zeta}(p). \tag{2.7.13}$$

Since

$$\begin{aligned}\widehat{Kg_n}(p) &= \hat{\eta}\left(\frac{1}{2}p\right)\sum_\ell \widehat{g_n}(p+4\pi\ell)\hat{\eta}\left(\frac{1}{2}p+2\pi\ell\right)^* \\ &= e^{-i\frac{1}{2}np}\hat{\eta}\left(\frac{1}{2}p\right)\sum_\ell\left|\hat{\eta}\left(\frac{1}{2}p+2\pi\ell\right)\right|^2,\end{aligned} \tag{2.7.14}$$

it follows that

$$\begin{aligned}(1+\widehat{\alpha K)^{-1}}g_n(p) &= e^{-i\frac{1}{2}np}\hat{\eta}\left(\frac{1}{2}p\right) - \alpha\left(1+\alpha\sum_\ell\left|\hat{\eta}\left(\frac{1}{2}p+2\pi\ell\right)\right|^2\right)^{-1} \\ &\quad\times e^{-i\frac{1}{2}np}\hat{\eta}\left(\frac{1}{2}p\right)\sum_\ell\left|\hat{\eta}\left(\frac{1}{2}p+2\pi\ell\right)\right|^2 \\ &= \left(1+\alpha\sum_\ell\left|\hat{\eta}\left(\frac{1}{2}p+2\pi\ell\right)\right|^2\right)^{-1} e^{-i\frac{1}{2}np}\hat{\eta}\left(\frac{1}{2}p\right),\end{aligned} \tag{2.7.15}$$

and therefore

$$\widehat{\psi}(p) = \sum_n d_n\left(\sum_\ell\left|\hat{\eta}\left(\frac{1}{2}p+2\pi\ell\right)\right|^2\right)^{-1} e^{-i\frac{1}{2}np}\hat{\eta}\left(\frac{1}{2}p\right). \tag{2.7.16}$$

In this case, the linear constraints are

$$d_m = 2(-1)^m c_{1-m}, \tag{2.7.17}$$

and so we have the solution explicitly given in momentum space:

$$\widehat{\psi}(p) = -2e^{-i\frac{1}{2}p}\left(\sum_\ell\left|\hat{\eta}\left(\frac{1}{2}p+2\pi\ell\right)\right|^2\right)^{-1}\hat{\eta}\left(\frac{1}{2}p\right)h(p+2\pi)^* \tag{2.7.18}$$

with h given by (2.2.6). Property (c) in §2.2 and the summability of $\{c_m\}$ guarantee that this momentum expression is well-defined and continuous, so ψ is certainly square-integrable if η is. It is worthwhile to remark that

$$\mathcal{K}_r = \mathcal{H}_r^\perp \cap \mathcal{H}_{r+1} = \overline{\operatorname{span}\{\psi_r(\cdot - 2^{-r}n):\ n\in\mathbb{Z}\}} \tag{2.7.19}$$

with $\mathcal{H}_r$ defined via Property (a) in §2.2.

To see this, we set $r = 0$ without loss of generality (by the dyadic scale covariance of the construction) and note that it suffices to show that $\eta(2x)$ and $\eta(2x-1)$ can both be expanded in the $\psi(x-n)$ and $\eta(x-n)$. In momentum space this would mean

$$\hat{\eta}\left(\frac{1}{2}p\right) = f(p)\hat{\eta}(p) + g(p)\widehat{\psi}(p), \tag{2.7.20}$$

$$e^{-i\frac{1}{2}p}\hat{\eta}\left(\frac{1}{2}p\right) = \tilde{f}(p)\hat{\eta}(p) + \tilde{g}(p)\widehat{\psi}(p) \tag{2.7.20'}$$

for 2π-periodic functions $f, g, \tilde{f}, \tilde{g}$ with square-summable Fourier coefficients. Inserting (2.2.7) and (2.7.18) and dividing out the resulting common factor $\hat{\eta}\left(\frac{1}{2}p\right)$, these conditions reduce to

$$1 = \frac{1}{2}h(p)f(p) - 2e^{-i\frac{1}{2}p}\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell\right)\right|^2\right)^{-1} h(p+2\pi)^* g(p), \tag{2.7.21}$$

$$1 = \frac{1}{2}e^{i\frac{1}{2}p}h(p)\tilde{f}(p) - 2\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell\right)\right|^2\right)^{-1} h(p+2\pi)^* \tilde{g}(p). \tag{2.7.21'}$$

On the other hand, the 2π-periodicity of the unknown functions imposes the additional conditions

$$1 = \frac{1}{2}h(p+2\pi)f(p) + 2e^{-i\frac{1}{2}p}\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell + \pi\right)\right|^2\right)^{-1} h(p)^* g(p), \tag{2.7.22}$$

$$1 = -\frac{1}{2}e^{i\frac{1}{2}p}h(p+2\pi)\tilde{f}(p) - 2\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell + \pi\right)\right|^2\right)^{-1} h(p)^* \tilde{g}(p). \tag{2.7.22'}$$

(2.7.21) and (2.7.22) constitute a 2×2 linear system for $f(p)$ and $g(p)$, as do (2.7.21′) and (2.7.22′) for $\tilde{f}(p)$ and $\tilde{g}(p)$. Solving for $f(p)$ and $g(p)$, we obtain

$$\begin{aligned} f(p) &= 2e^{i\frac{1}{2}p}D(p)^{-1}\left[h(p)^*\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell + \pi\right)\right|^2\right)^{-1}\right. \\ &\quad \left. + \, h(p+2\pi)^*\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell\right)\right|^2\right)^{-1}\right], \end{aligned} \tag{2.7.23}$$

$$g(p) = \frac{1}{2}D(p)^{-1}(h(p) - h(p+2\pi)), \tag{2.7.24}$$

where $D(p)$ is the determinant of the coefficient matrix and is given by

$$\begin{aligned} D(p) &= e^{-i\frac{1}{2}p}\left[|h(p)|^2\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell + \pi\right)\right|^2\right)^{-1}\right. \\ &\quad \left. + \, |h(p+2\pi)|^2\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell\right)\right|^2\right)^{-1}\right]. \end{aligned} \tag{2.7.25}$$

Since $D(p+2\pi) = -D(p)$, $f(p)$ and $g(p)$ are both 2π-periodic. Similarly,

$$\begin{aligned} \tilde{f}(p) &= -2\widetilde{D}(p)^{-1}\left[h(p)^*\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p+2\pi\ell+\pi\right)\right|^2\right)^{-1}\right. \\ &\qquad \left. -h(p+2\pi)^*\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p+2\pi\ell\right)\right|^2\right)^{-1}\right], \qquad (2.7.26) \\ \tilde{g}(p) &= \frac{1}{2}e^{i\frac{1}{2}p}\widetilde{D}(p)^{-1}(h(p)+h(p+2\pi)), \qquad (2.7.27) \end{aligned}$$

where $\widetilde{D}(p)$ is the determinant for the latter system. Actually,

$$\widetilde{D}(p) = -D(p)^*, \qquad (2.7.28)$$

so $\widetilde{D}(p+2\pi) = -\widetilde{D}(p)$, and therefore $\tilde{f}(p)$ and $\tilde{g}(p)$ are 2π-periodic as well. This completes the proof of (2.7.19), which, in turn, implies

$$\{\psi_r(\cdot - 2^{-r}n) \colon\ n, r \in \mathbb{Z}\}$$

spans $L^2(\mathbb{R})$, because $\bigcap_r \mathcal{H}_r$ is assumed to be the null space.

Thus, we obtain an orthonormal basis of wavelets if we orthogonalize this interscale-orthogonal set scale by scale. Clearly, the mother wavelet Ψ for the desired basis is given by (2.7.2). This basis transformation is the last step in the variational construction and is applied to a wavelet, while the same basis transformation is the first step in the Meyer–Mallat construction and is applied to a scaling function. Given the same initial scaling function η, however, the resulting mother wavelet Ψ is the same. To show this, note that

$$\begin{aligned} \sum_\ell |\widehat{\psi}(p+2\pi\ell)|^2 &= \sum_\ell |\widehat{\psi}(p+4\pi\ell+2\pi)|^2 + \sum_\ell |\widehat{\psi}(p+4\pi\ell)|^2, \qquad (2.7.29) \\ \sum_\ell |\widehat{\psi}(p+4\pi\ell)|^2 &= 4|h(p+2\pi)|^2\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p+2\pi\ell\right)\right|^2\right)^{-1}, \qquad (2.7.30) \end{aligned}$$

so by (2.7.25), we may write

$$\sum_\ell |\widehat{\psi}(p+2\pi\ell)|^2 = 4e^{i\frac{1}{2}p}D(p). \qquad (2.7.31)$$

The inverse square root of the bracketed expression in (2.7.25) is involved in this $\mathbb{Z}$-translation-covariant orthogonalization. On the other hand, the scale relation (2.2.7) implies

$$|h(p)|^2\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p+2\pi\ell\right)\right|^2 = 4\sum_\ell |\hat{\eta}(p+4\pi\ell)|^2 \qquad (2.7.32)$$

and therefore

$$|h(p)|^2 \sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell\right)\right|^2 + |h(p+2\pi)|^2 \sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell + \pi\right)\right|^2 = 4\sum_\ell |\hat{\eta}(p+2\pi\ell)|^2 \tag{2.7.33}$$

This reduces (2.7.31) to

$$\begin{aligned}\sum_\ell |\widehat{\psi}(p+2\pi\ell)|^2 &= 16\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell + \pi\right)\right|^2\right)^{-1} \\ &\times \left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell\right)\right|^2\right)^{-1} \sum_\ell |\hat{\eta}(p+2\pi\ell)|^2. \end{aligned}\tag{2.7.34}$$

Inserting this formula and the formula (2.7.18) for $\widehat{\psi}(p)$, we realize (2.7.2) more explicitly as

$$\begin{aligned}\widehat{\Psi}(p) &= -\frac{1}{2}e^{-i\frac{1}{2}p}\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell\right)\right|^2\right)^{-1/2}\left(\sum_\ell |\hat{\eta}(p+2\pi\ell)|^2\right)^{-1/2} \\ &\times \left(\sum_\ell \left|\eta\left(\frac{1}{2}p + 2\pi\ell + \pi\right)\right|^2\right)^{1/2} h(p+2\pi)^*\hat{\eta}\left(\frac{1}{2}p\right). \end{aligned}\tag{2.7.35}$$

Now by (2.2.8) we have

$$\tilde{h}(p+2\pi) = \left(\sum_\ell |\hat{\eta}(p+2\pi\ell)|^2\right)^{-1/2}\left(\sum_\ell \left|\hat{\eta}\left(\frac{1}{2}p + 2\pi\ell + \pi\right)\right|^2\right)^{1/2} h(p+2\pi), \tag{2.7.36}$$

and so we may write

$$\widehat{\Psi}(p) = -\frac{1}{2}e^{-i\frac{1}{2}p}\tilde{h}(p+2\pi)^*\widehat{\varphi}\left(\frac{1}{2}p\right), \tag{2.7.37}$$

where we have also applied (2.2.3). Since the formula is just the momentum-space version of (2.2.12), we have identified the mother wavelet constructed here with the mother wavelet produced by the multiscale resolution analysis of Meyer and Mallat.

References

1. G. Battle, "A Block-Spin Construction of Ondelettes, Part I: Lemarié Functions," Commun. Math. Phys. **110** (1987), 601–615.

2. G. Battle, "Cardinal Spline Interpolation and the Block-Spin Construction of Wavelets," in *Wavelets – A Tutorial in Theory and Applications*, C. Chui, ed., Academic Press, San Diego, California, 1992, pp. 73–93.

3. G. Battle, "Wavelets: A Renormalization Group Point of View," in *Wavelets and Their Applications*, B. Ruskai et al. (eds.), Bartlett and Jones, Boston, 1992, pp. 323–349.

4. P. Lemarié-Rieusset and J. Kahane, *Fourier Series and Wavelets,* Gordon and Breach, Luxembourg, 1995.

5. S. Mallat, "Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(\mathbb{R})$," Trans. Amer. Math. Soc. **315**, No. 1 (1989), 69–87.

2.8 Without Intrascale Orthogonality

In the variational construction we have just described, the major step is to derive a basis $\{\psi(2^r x - n)\colon\ r, n \in \mathbb{Z}\}$ of wavelets that are orthogonal for differing length scales but not orthogonal on the same scale. From a computational point of view, this basis is actually more desirable than the orthonormal basis we obtained from it. To be sure, we must use the dual basis to compute expansion coefficients, but the derivation of the dual basis is rather painless. Given the orthogonality between length scales, the dual basis is realized scale by scale.

Our focus here is on the case of Lemarié wavelets – i.e., the case where the initial scaling function η is given by

$$\hat{\eta}(p) = \widehat{\chi}(p)^{N+1} \tag{2.8.1}$$

The degree of regularity for ψ is parametrized by N, with each value of N corresponding to an *Euler–Frobenius polynomial*. From such a polynomial one can derive an explicit position-space representation of ψ by the calculus of residues. Indeed, ψ is a linear combination of discrete translates of η with coefficients determined by the complex roots of the associated Euler–Frobenius polynomial. Recall from (2.7.16) and (2.7.17) that

$$\widehat{\psi}(p) = 2\sum_n (-1)^n c_{1-n} \left(\sum_\ell \left| \hat{\eta}\left(\frac{1}{2}p + 2\pi\ell\right)\right|^2 \right)^{-1} e^{-i\frac{1}{2}np}\hat{\eta}\left(\frac{1}{2}p\right), \tag{2.8.2}$$

where the sequence $\{c_n\}$ is given by

$$\eta(x) = \sum_n c_n \eta(2x - n). \tag{2.8.3}$$

The key observation for what follows is the identity

$$\sum_\ell \frac{1}{(p' + 2\pi\ell)^2} = \frac{1}{4}\csc^2\left(\frac{1}{2}p'\right), \tag{2.8.4}$$

which is a result of the Poisson summation formula. Indeed,

$$1 = \sum_n e^{inp'}\delta_{0n} \quad = \quad \sum_n e^{inp'} \int \chi(x)\chi(x-n)dx$$

$$
\begin{aligned}
&= \frac{1}{2\pi}\sum_n \int |\widehat{\chi}(q)|^2 e^{-in(q-p')}dq \\
&= \sum_\ell |\widehat{\chi}(p'+2\pi\ell)|^2 \\
&= |e^{ip'}-1|^2 \sum_\ell \frac{1}{(p'+2\pi\ell)^2} \\
&= 4\sin^2\left(\frac{1}{2}p'\right)\sum_\ell \frac{1}{(p'+2\pi\ell)^2}. \qquad (2.8.5)
\end{aligned}
$$

To apply (2.8.4), observe that

$$
\begin{aligned}
\sum_\ell |\hat{\eta}(p'+2\pi\ell)|^2 &= \sum_\ell |\widehat{\chi}(p'+2\pi\ell)|^{2N+2} \\
&= 4^{N+1}\sin^{2N+2}\left(\frac{1}{2}p'\right)\sum_\ell \frac{1}{(p'+2\pi\ell)^{2N+2}} \\
&= \frac{4^{N+1}}{(2N+1)!}\sin^{2N+2}\left(\frac{1}{2}p'\right)\frac{d^{2N}}{dp'^{2N}}\sum_\ell \frac{1}{(p'+2\pi\ell)^2} \\
&= \frac{4^{N}}{(2N+1)!}\sin^{2N+2}\left(\frac{1}{2}p'\right)\frac{d^{2N}}{dp'^{2N}}\csc^2\left(\frac{1}{2}p'\right). \qquad (2.8.6)
\end{aligned}
$$

On the other hand, even-order derivatives of $\csc^2\left(\frac{1}{2}p'\right)$ are polynomials in $\csc^2\left(\frac{1}{2}p'\right)$ and $\cot^2\left(\frac{1}{2}p'\right)$, homogeneous of degree equal to two plus the even order, with at least one factor of $\csc^2\left(\frac{1}{2}p'\right)$ in each term. Multiplying this by $\sin^{2N+2}\left(\frac{1}{2}p'\right)$, we obtain for (2.8.6) a polynomial in $\cos^2\left(\frac{1}{2}p'\right)$ only, and with no power higher than $\cos^{2N}\left(\frac{1}{2}p'\right)$. With

$$\cos^2\left(\frac{1}{2}p'\right) = \frac{1}{2}+\frac{1}{2}\cos p', \qquad (2.8.7)$$

we may therefore see it as an Nth-degree polynomial in $\cos p'$. Let R_N denote this last polynomial, so that

$$\sum_\ell |\hat{\eta}(p'+2\pi\ell)|^2 = R_N(\cos p'). \qquad (2.8.8)$$

Combining this with (2.8.2), setting $p' = \frac{1}{2}p$, we obtain

$$\widehat{\psi}(p) = 2\sum_n (-1)^n c_{1-n}\frac{e^{-i\frac{1}{2}np}\hat{\eta}\left(\frac{1}{2}p\right)}{R_N\left(\cos\left(\frac{1}{2}p\right)\right)}. \qquad (2.8.9)$$

On the other hand,

$$
\begin{aligned}
\sum_n (-1)^n c_{1-n}e^{-i\frac{1}{2}np} &= -e^{-i\frac{1}{2}p}\sum_n (-1)^n c_n e^{i\frac{1}{2}np} \\
&= -e^{-i\frac{1}{2}p}\sum_{n=0}^{N+1}(-1)^n\binom{N+1}{n}e^{i\frac{1}{2}np} \\
&= -e^{-i\frac{1}{2}p}(1-e^{i\frac{1}{2}p})^{N+1} \qquad (2.8.10)
\end{aligned}
$$

when the scaling function is given by (2.8.1), so

$$\widehat{\psi}(p) = -2e^{-i\frac{1}{2}p}(1 - e^{i\frac{1}{2}p})^{N+1}\frac{\widehat{\chi}\left(\frac{1}{2}p\right)^{N+1}}{R_N\left(\cos\left(\frac{1}{2}p\right)\right)}. \tag{2.8.11}$$

In position space we have a spline with half-integer knots realized as an expansion in half-integer translates of $\eta(2x)$. Indeed, if we set

$$\frac{e^{-i\frac{1}{2}p}(1 - e^{i\frac{1}{2}p})^{N+1}}{R_N\left(\cos\left(\frac{1}{2}p\right)\right)} = \sum_m a_m e^{-i\frac{1}{2}mp}, \tag{2.8.12}$$

we have

$$\psi(x) = -2\sum_m a_m \eta(2x - m), \tag{2.8.13}$$

$$\eta = (*\chi)^{N+1}, \tag{2.8.14}$$

$$a_m = \frac{1}{4\pi}\int_0^{4\pi} e^{i\frac{1}{2}(m-1)p}(1 - e^{i\frac{1}{2}p})^{N+1}\frac{1}{R_N\left(\cos\left(\frac{1}{2}p\right)\right)}dp. \tag{2.8.15}$$

The calculation of these coefficients is our major concern here.

These integrals can easily be written as contour integrals in the complex plane. If we set

$$z = e^{i\frac{1}{2}p}, \qquad 0 \le p \le 4\pi, \tag{2.8.16}$$

we have a parametrization of the unit circle $|z| = 1$ and the formula

$$a_m = \frac{1}{2\pi i}\int_{|z|=1} z^{m-2}(1 - z)^{N+1}\frac{1}{R_N\left(\frac{1}{2}z + \frac{1}{2}z^{-1}\right)}dz. \tag{2.8.17}$$

On the other hand, if we set

$$z = e^{-i\frac{1}{2}p}, \qquad 0 \le p \le 4\pi, \tag{2.8.18}$$

then we have the alternative formula

$$a_m = -\frac{1}{2\pi i}\int_{|z|=1} z^{-m}(1 - z^{-1})^{N+1}\frac{1}{R_N\left(\frac{1}{2}z + \frac{1}{2}z^{-1}\right)}dz. \tag{2.8.19}$$

Since

$$K_N(z) \equiv z^N R_N\left(\frac{1}{2}z + \frac{1}{2}z^{-1}\right) \tag{2.8.20}$$

is a $2N$th-degree polynomial with no zero at $z = 0$, we can avoid computing residues at $z = 0$ for almost all m. The two formulas become

$$\begin{aligned} a_m &= \frac{1}{2\pi i}\int_{|z|=1} z^{m+N-2}(1 - z)^{N+1}\frac{1}{K_N(z)}dz, & (2.8.21)\\ a_m &= -\frac{1}{2\pi i}\int_{|z|=1} z^{-m-1}(z - 1)^{N+1}\frac{1}{K_N(z)}dz, & (2.8.22)\end{aligned}$$

and we choose to use (2.8.21) for $m \geq 2 - N$. If $\alpha_1, \ldots, \alpha_k$ are the distinct roots of $K_N(z)$ in the interior of the unit disk, then

$$a_m = \sum_{\iota=1}^{k} \operatorname{Res}\left(z^{m+N-2}(1-z)^{N+1}\frac{1}{K_N(z)}, a_\iota\right), \qquad m \geq 2-N. \tag{2.8.23}$$

(Note that $K_N(z)$ has no zeros on the boundary $|z| = 1$ because $R_N\left(\cos\frac{1}{2}p\right)$ does not vanish anywhere.) We use (2.8.22) in the complementary case. Thus

$$\begin{aligned} a_m &= -\sum_{\iota=1}^{k} \operatorname{Res}\left(z^{-m-1}(z-1)^{N+1}\frac{1}{K_N(z)}, \alpha_\iota\right), \\ &\quad m \leq \min\{-1, 1-N\}, \end{aligned} \tag{2.8.24}$$

$$\begin{aligned} a_m &= -\operatorname{Res}\left(z^{-m-1}(z-1)^{N+1}\frac{1}{K_N(z)}, 0\right) \\ &\quad -\sum_{\iota=1}^{k} \operatorname{Res}\left(z^{-m-1}(z-1)^{N+1}\frac{1}{K_N(z)}, \alpha_\iota\right), \quad 0 \leq m \leq 1-N. \end{aligned} \tag{2.8.25}$$

Obviously, this last case is vacuus if $N > 1$ and it is the only case where there is a residue at $z = 0$.

We now derive the position space representation for the dual basis, which can be calculated on each scale because the basis is interscale-orthogonal. As always, we concentrate on the unit-scale functions – i.e., the unit translates of the mother wavelet – without loss of generality. The overlap matrix is

$$\begin{aligned} \mathcal{S}_{mn} &= \int \psi(x-m)\psi(x-n)dx \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i(m-n)p}|\widehat{\psi}(p)|^2 dp \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(m-n)p}\sum_{\ell}|\widehat{\psi}(p+2\pi\ell)|^2 dp, \end{aligned} \tag{2.8.26}$$

so the dual basis $\{u_m\}$ is given by

$$\begin{aligned} u_m(x) &= \sum_n (\mathcal{S}^{-1})_{mn}\psi(x-n) \\ &= \frac{1}{2\pi}\sum_n \psi(x-n)\int_{-\pi}^{\pi} e^{-i(m-n)p}\left(\sum_{\ell}|\widehat{\psi}(p+2\pi\ell)|^2\right)^{-1} dp\,. \end{aligned} \tag{2.8.27}$$

The Poisson summation formula

$$\sum_n e^{inp}\psi(x-n) = \sum_j e^{ix(p+2\pi j)}\widehat{\psi}(p+2\pi j) \tag{2.8.28}$$

implies

$$\begin{aligned} u_m(x) &= \frac{1}{2\pi}\sum_j \int_{-\pi}^{\pi} e^{i(x-m)(p+2\pi j)}\widehat{\psi}(p+2\pi j)\left(\sum_\ell |\widehat{\psi}(p+2\pi\ell)|^2\right)^{-1} dp \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i(x-m)p}\widehat{\psi}(p)\left(\sum_\ell |\widehat{\psi}(p+2\pi\ell)|^2\right)^{-1} dp. \end{aligned} \tag{2.8.29}$$

Thus

$$u_m(x) = u_0(x-m), \tag{2.8.30}$$

and in momentum space,

$$\hat{u}_0(p) = \left(\sum_\ell |\widehat{\psi}(p+2\pi\ell)|^2\right)^{-1}\widehat{\psi}(p). \tag{2.8.31}$$

On the other hand, (2.8.11) implies

$$\begin{aligned} \sum_\ell |\widehat{\psi}(p+2\pi\ell)|^2 &= 4\sum_\ell |1-e^{i\pi\ell}e^{i\frac{1}{2}p}|^{2N+2}\frac{|\hat{\eta}(\frac{1}{2}p+\pi\ell)|^2}{R_N(\cos(\frac{1}{2}p+\pi\ell))^2} \\ &= 4\frac{|1-e^{i\frac{1}{2}p}|^{2N+2}}{R_N(\cos(\frac{1}{2}p))^2}\sum_j \left|\hat{\eta}\left(\frac{1}{2}p+2\pi j\right)\right|^2 \\ &\quad + 4\frac{|1+e^{i\frac{1}{2}p}|^{2N+2}}{R_N(-\cos(\frac{1}{2}p))^2}\sum_j \left|\hat{\eta}\left(\frac{1}{2}p+\pi+2\pi j\right)\right|^2 \\ &= 4\frac{|1-e^{i\frac{1}{2}p}|^{2N+2}}{R_N(\cos(\frac{1}{2}p))} + 4\frac{|1+e^{i\frac{1}{2}p}|^{2N+2}}{R_N(-\cos(\frac{1}{2}p))}. \end{aligned} \tag{2.8.32}$$

Hence

$$\begin{aligned} \hat{u}_0(p) &= -\frac{1}{2}e^{-i\frac{1}{2}p}(1-e^{i\frac{1}{2}p})^{N+1}R_N\left(-\cos\left(\frac{1}{2}p\right)\right)\hat{\eta}\left(\frac{1}{2}p\right)^{N+1} \\ &\quad \times\left[R_N\left(\cos\left(\frac{1}{2}p\right)\right)|1+e^{i\frac{1}{2}p}|^{2N+2}\right. \\ &\quad \left.+ R_N\left(-\cos\left(\frac{1}{2}p\right)\right)|1-e^{i\frac{1}{2}p}|^{2N+2}\right]^{-1}. \end{aligned} \tag{2.8.33}$$

For the complex integration variable (2.8.16), we obtain

$$u_0(x) = -\sum_m b_m\eta(2x-m), \tag{2.8.34}$$

$$b_m = \frac{1}{2\pi i}\int_{|z|=1} z^{m-2}(1-z)^{N+1}R_N\left(-\frac{1}{2}z-\frac{1}{2}z^{-1}\right)$$

$$\begin{aligned}
&\times \left[R_N\left(\frac{1}{2}z+\frac{1}{2}z^{-1}\right)(1+z)^{N+1}(1+z^{-1})^{N+1}\right. \\
&+\left. R_N\left(-\frac{1}{2}z-\frac{1}{2}z^{-1}\right)(1-z)^{N+1}(1-z^{-1})^{N+1}\right]^{-1} dz \\
=\ &\frac{1}{2\pi i}\int_{|z|=1} z^{N+m-1}(1-z)^{N+1}K_N(-z)(-1)^N \\
&\times [K_N(z)(1+z)^{2N+2}-(1-z)^{2N+2}K_N(-z)]^{-1}dz, \qquad (2.8.35)
\end{aligned}$$

while the alternate substitution (2.8.18) yields

$$\begin{aligned}
b_m\ =\ &-\frac{1}{2\pi i}\int_{|z|=1} z^{-m}(1-z^{-1})^{N+1}R_N\left(-\frac{1}{2}z-\frac{1}{2}z^{-1}\right) \\
&\times \left[R_N\left(\frac{1}{2}z+\frac{1}{2}z^{-1}\right)(1+z)^{N+1}(1+z^{-1})^{N+1}\right. \\
&+\left. R_N\left(-\frac{1}{2}z-\frac{1}{2}z^{-1}\right)(1-z)^{N+1}(1-z^{-1})^{N+1}\right]^{-1} dz \\
=\ &-\frac{1}{2\pi i}\int_{|z|=1} z^{-m}(z-1)^{N+1}K_N(-z)(-1)^N \\
&\times [K_N(z)(1+z)^{2N+2}-(1-z)^{2N+2}K_N(-z)]^{-1}dz. \qquad (2.8.36)
\end{aligned}$$

As involved as these contour integrals appear to be, the residue calculations are very closely related to (2.8.23)–(2.8.25), as we now show.

First note that the variable relation $z=e^{ip'}$ obviously yields the relations

$$\csc^2\left(\frac{1}{2}p'\right) = -\frac{4z}{(z-1)^2}, \qquad (2.8.37)$$

$$\frac{d}{dp'}f(e^{ip'}) = iz\frac{df}{dz}\bigg|_{z=e^{ip'}}, \qquad (2.8.38)$$

from which it follows that

$$\begin{aligned}
R_N(\cos p')\ =\ &\frac{4^N}{(2N+1)!}\sin^{2N+2}\left(\frac{1}{2}p'\right) \\
&\times\frac{d^{2N}}{dp'^{2N}}\csc^2\left(\frac{1}{2}p'\right) = \frac{4^N}{(2N+1)!}\frac{(z-1)^{2N+2}}{(-4z)^{N+1}} \\
&\times\left(iz\frac{d}{dz}\right)^{2N}\left[\frac{-4z}{(z-1)^2}\right] \\
=\ &\frac{1}{(2N+1)!}\frac{(z-1)^{2N+2}}{z^{N+1}}\left(z\frac{d}{dz}\right)^{2N}\left[\frac{z}{(z-1)^2}\right]. \qquad (2.8.39)
\end{aligned}$$

Second, we insert (2.8.20) to get

$$K_N(z) = \frac{1}{(2N+1)!}\frac{(z-1)^{2N+2}}{z}\left(z\frac{d}{dz}\right)^{2N}\left[\frac{z}{(z-1)^2}\right], \qquad (2.8.40)$$

which immediately implies

$$\begin{aligned} & K_N(z)(1+z)^{2N+2} - K_N(-z)(1-z)^{2N+2} \\ = & \frac{(z^2-1)^{2N+2}}{z(2N+1)!}\left(z\frac{d}{dz}\right)^{2N}\left[\frac{z}{(z-1)^2} - \frac{z}{(z+1)^2}\right]. \end{aligned} \tag{2.8.41}$$

Next, we observe that

$$\frac{z}{(z-1)^2} - \frac{z}{(z+1)^2} = \frac{4z^2}{(z^2-1)^2}, \tag{2.8.42}$$

and the variable relation $w = z^2$ yields

$$z\frac{d}{dz}f(z^2) = 2w\frac{df}{dw}\bigg|_{w=z^2} \tag{2.8.43}$$

Finally,

$$K_N(z)(1+z)^{2N+2} - K_N(-z)(1-z)^{2N+2} = 4^{N+1}K_N(z^2)z, \tag{2.8.44}$$

and we may now insert this in (2.8.35) and (2.8.36) to obtain

$$b_m = \frac{4^{-N-1}}{2\pi i}\int_{|z|=1} z^{m+N-2}(1-z)^{N+1}K_N(-z)\frac{(-1)^N}{K_N(z^2)}dz, \tag{2.8.45}$$

$$b_m = -\frac{4^{-N-1}}{2\pi i}\int_{|z|=1} z^{-m-1}(z-1)^{N+1}K_N(-z)\frac{(-1)^N}{K_N(z^2)}dz, \tag{2.8.46}$$

respectively.

Both ψ and u_0 are exponentially localized, and we are now in a position to determine the rates of exponential decay. Clearly, it suffices to determine the exponential decay rates of the sequences $\{a_m\}$ and $\{b_m\}$, since η has compact support. Without loss of generality, we consider $m \geq 2-N$ or $m \leq min\{-1, 1-N\}$; in the former case, a_m is given by (2.8.23), and therefore $a_m = O(\rho^m)$ for large positive m, where

$$\rho = \max_{\iota}|\alpha_\iota| < 1. \tag{2.8.47}$$

In the latter case, a_m is given by (2.8.24), and so $a_m = O(\rho^{-m})$ for large negative m. This means the rate of exponential decay in either direction is the positive number $-\ln\rho$. Now consider the same two cases for the sequence $\{b_m\}$; in the former case, we apply (2.8.45) to obtain

$$b_m = -4^{-N-1}\sum_{\iota,\pm}\text{Res}\left(z^{m+N-2}(z-1)^{N+1}\frac{K_N(-z)}{K_N(z^2)}, \pm\sqrt{\alpha_\iota}\right). \tag{2.8.48}$$

Thus $b_m = O(\rho^{m/2})$ for large positive m. In the latter case, we apply (2.8.46) to obtain

$$b_m = 4^{-N-1}\sum_{\iota,\pm}\text{Res}\left(z^{-m-1}(1-z)^{N+1}\frac{K_N(-z)}{K_N(z^2)}, \pm\sqrt{\alpha_\iota}\right), \tag{2.8.49}$$

which implies $b_m = O(\rho^{-m/2})$ for large negative m. Therefore, the rate of exponential decay in either direction is $-\frac{1}{2}\ln\rho$.

This incidentally proves that the exponential decay rate of $u_0(x)$ is half that of $\psi(x)$, regardless of the value of N.

References

1. G. Battle, "Cardinal Spline Interpolation and the Block-Spin Construction of Wavelets," in *Wavelets - A Tutorial in Theory and Applications*, C. Chui, ed., Academic Press, San Diego, California, 1992, pp. 73–93.

2. C. Chui and J. Wang, "A Cardinal Spline Approach to Wavelets," Proc. Amer. Math. Soc. **113** (1991), 785–793.

3. P. Lemarié-Rieusset, "Une Remarque sur les Propriétés Multi-Résolutions des Fonctions Splines," C.R. Acad. Sci. Paris **317** (1993), 1115–1117.

4. P. Lemarié-Rieusset and J. Kahane, *Fourier Series and Wavelets,* Gordon and Breach, Luxembourg, 1995.

2.9 Chui–Wang Wavelets

There is a class of interscale-orthogonal wavelet bases whose mother wavelets are compactly supported splines. For such a basis, the mother wavelet of the dual basis is not compactly supported, but the interscale orthogonality of the compactly supported splines in the former basis is remarkable. The mother wavelet is a *Chui–Wang wavelet*, and the point of view of the construction lies somewhere between the Meyer–Mallat approach and the variational approach. We begin with a brief discussion of the wavelets constructed in the previous section to describe this intermediate point of view.

Let η_M be the scaling function given by

$$\hat{\eta}_M(p) = \widehat{\chi}(p)^M \tag{2.9.1}$$

and consider the L^2-subspace $\mathcal{H}^{(M)}$ spanned by the integer translates of η_M. In approximation theory the *fundamental Mth-order cardinal interpolating spline* is the solution to minimizing $\|\zeta\|_2$ in $\mathcal{H}^{(M)}$ with respect to the constraints

$$\zeta(m) = \delta_{m0}, \qquad m \in \mathbb{Z}. \tag{2.9.2}$$

Standard notation for the solution of this problem is $L_M(x)$, and we denote the expansion of this fundamental spline by

$$L_M(x) = \sum_n b_n^{(M)} \eta_M(x-n). \tag{2.9.3}$$

Obviously, this is not an orthogonal expansion, but the coefficients are uniquely determined. Indeed, they can be explicitly realized in terms of the zeros of the $2N$th-degree Euler–Frobenius polynomial in the unit disk of the complex plane. The crucial observation here is that the derivative function

$$L_{2J}^{(J)}(2x-1) = 2^{-J}\frac{d^J}{dx^J}L_{2J}(2x-1)$$

is actually the mother wavelet of an interscale-orthogonal basis. Indeed, if we integrate by parts we have

$$\begin{aligned} & \int L_{2J}^{(J)}(2x-1)\eta_J(2^{r+1}x-m)dx \\ = \quad & (-1)^J 2^{rJ} \int L_{2J}(2x-1)\eta_J^{(J)}(2^{r+1}x-m)dx \\ = \quad & 2^{rJ} \int L_{2J}(2x-1)\sum_{j=0}^{J}(-1)^j\binom{J}{j}\delta(2^{r+1}x-m-j)dx \\ = \quad & 2^{rJ-r-1}\sum_{j=0}^{J}(-1)^j\binom{J}{j}L_{2J}(2^{-r}(m+j)-1). \end{aligned} \tag{2.9.4}$$

By (2.9.2) these evaluations of L_{2J} all vanish if $r<0$, so

$$L_{2J}^{(J)}(2x-1) \perp \operatorname{span}\{\eta_J(2^{r+1}x-n)\colon\ n\in\mathbb{Z}\}, \qquad r<0.$$

On the other hand,

$$\begin{aligned} & L_{2J}^{(J)}(2^{r+1}x-2m-1) \\ = \quad & \sum_n b_n^{(2J)}\eta_{2J}^{(J)}(2^{r+1}x-2m-n-1) \\ = \quad & \sum_n b_n^{(2J)}\sum_{j=0}^{J}(-1)^{J-j}\binom{J}{j}\eta_J(2^{r+1}x-2m-n-j-1). \end{aligned} \tag{2.9.5}$$

Therefore,

$$L_{2J}^{(J)}(2x-1) \perp L_{2J}^{(J)}(2^{r+1}x-2m-1), \qquad r<0,$$

and so the interscale orthogonality is established.

This wavelet basis is precisely the interscale-orthogonal Lemarié basis constructed in the previous section. To see this, observe first that

$$\begin{aligned} & \int L_{2J}^{(J)}(2x-1)\eta_J(2x-m)dx \\ = \quad & (-1)^J \int L_{2J}(2x-1)\eta_J^{(J)}(2x-m)dx \\ = \quad & \int L_{2J}(2x-1)\sum_{j=0}^{J}(-1)^j\binom{J}{j}\delta(2x-m-j)dx \\ = \quad & \sum_{j=0}^{J}(-1)^j\binom{J}{j}L_{2J}(m+j-1) \\ = \quad & -(-1)^m\sum_{j=0}^{J}\binom{J}{j}\delta_{1-m,j}. \end{aligned} \tag{2.9.6}$$

Second, observe that this sum is just $-(-1)^m c^{(J)}_{1-m}$, where

$$c^{(J)}_m = \sum_{j=0}^{J} \binom{J}{j} \delta_{mj} \tag{2.9.7}$$

are the coefficients for the scaling relation

$$\eta_J(x) = \sum_n c^{(J)}_n \eta_J(2x-n), \tag{2.9.8}$$

so the constraints (2.7.3) are satisfied by $-L^{(N+1)}_{2N+2}(2x-1)$. On the other hand, for any L^2-function f satisfying (2.7.3), the orthogonal projection $E^{(N+1)}f$ of f onto $\mathcal{H}^{(N+1)}$ satisfies (2.7.3) as well – and it is the only element of $\mathcal{H}^{(N+1)}$ satisfying (2.7.3). Thus $E^{(N+1)}f$ is the solution to minimizing the L^2-norm with respect to the constraints (2.7.3). This implies $-L^{(N+1)}_{2N+2}(2x-1)$ is that unique element of $\mathcal{H}^{(N+1)}$ and therefore the function $\psi(x)$ that minimizes the L^2-norm with respect to (2.7.3). This function is the wavelet given in the previous section, so the claim is established. We could have made just this identification

$$\psi(x) = -L^{(N+1)}_{2N+2}(2x-1) \tag{2.9.9}$$

by simply calculating both quantities in momentum space, but the reasoning we have just given is more to the point.

We now consider the Chui–Wang wavelet, which can actually be written as

$$\psi(x) = \sum_\nu (-1)^\nu \eta_{2J}(\nu+1)\eta^{(J)}_{2J}(2x-\nu). \tag{2.9.10}$$

The verification of interscale orthogonality is the integration by parts

$$\begin{aligned}
&\int \psi(x)\eta_J(2^{r+1}x-m)dx \\
= \;& \sum_\nu (-1)^\nu \eta_{2J}(\nu+1) \int \eta^{(J)}_{2J}(2x-\nu) \\
& \times\; \eta_J(2^{r+1}x-m)dx \\
= \;& (-1)^J 2^{rJ} \sum_\nu (-1)^\nu \eta_{2J}(\nu+1) \int \eta_{2J}(2x-\nu) \\
& \times\; \eta^{(J)}_J(2^{r+1}x-m)dx \\
= \;& 2^{rJ} \sum_\nu (-1)^\nu \eta_{2J}(\nu+1) \int \eta_{2J}(2x-\nu) \\
& \times \sum_{j=0}^{J} (-1)^j \binom{J}{j} \delta(2^{r+1}x-m-j)dx \\
= \;& 2^{rJ} \sum_\nu (-1)^\nu \eta_{2J}(\nu+1) \sum_{j=0}^{J} (-1)^j \binom{J}{j} \eta_{2J}(2^{-r}(m+j)-\nu).
\end{aligned} \tag{2.9.11}$$

This numerical expression is identically zero for $r < 0$! Why? This ingenious construction exploits the familiar identity

$$\sum_{\nu}(-1)^{\nu} c_{\nu+1} c_{2n-\nu} = 0 \tag{2.9.12}$$

which follows from the observation that the sum is reversed in sign by the index change $\nu \mapsto 2n - 1 - \nu$. We have already used it in other wavelet constructions, but this application is more subtle. In this case, $c_{\mu} = \eta_{2J}(\mu)$ and $n = 2^{-r}(m + j)$, $r < 0$.

We may now argue as we did in the case of the spline $L_{2J}^{(J)}(2x - 1)$ above. Since we have just shown that

$$\int \psi(x) \eta_J(2^{r+1}x - m) dx = 0, \qquad r < 0, \tag{2.9.13}$$

we conclude that

$$\psi(x) \perp \psi(2^r x - m), \qquad r < 0,$$

from the observation that

$$\psi(2^r x - m) = \sum_{\nu}(-1)^{\nu} \eta_{2J}(\nu + 1) \sum_{j=0}^{J} (-1)^{J-j} \binom{J}{j} \eta_J(2^{r+1}x - 2m - \nu - j), \tag{2.9.14}$$

where the point of this last equation is that $\psi(2^r x - m)$ lies in the span of the $2^{-r-1}\mathbb{Z}$-translates of $\eta_J(2^{r+1}x - 2m)$.

How do we know that ψ has compact support? Since η_{2J} is obtained by convolution of χ with itself $2J$ times, we know that each half-integer translate of $\eta_{2J}^{(J)}(2x)$ is compactly supported. On the other hand, the linear combination (2.9.10) of these translates is a finite sum because

$$\eta_{2J}(\nu + 1) = 0, \qquad \nu < 0 \text{ or } \nu \geq 2J - 1. \tag{2.9.15}$$

Specifically, $\psi(x)$ is supported in the interval $0 \leq x \leq 2J - 1$. As far as regularity is concerned, ψ has the same degree of smoothness as η_J has – namely, class $C^{J-1-\varepsilon}$

We now examine the dual basis, which is itself interscale-orthogonal. To compute the unit-scale elements, we apply (2.8.30) because general formulas from the previous section apply here. The functions are unit-scale translates of u, where

$$\hat{u}(p) = \left(\sum_{\ell} |\widehat{\psi}(p + 2\pi\ell|^2 \right)^{-1} \widehat{\psi}(p). \tag{2.9.16}$$

On the other hand, the Fourier transform of (2.9.10) is just

$$\begin{aligned} \widehat{\psi}(p) &= \left(i\frac{1}{2}p\right)^J \hat{\eta}_{2J}\left(\frac{1}{2}p\right) \sum_{\nu}(-1)^{\nu} \eta_{2J}(\nu + 1) e^{-i\frac{1}{2}\nu p} \\ &= (-1)^J (e^{-i\frac{1}{2}p} - 1)^J \hat{\eta}_J\left(\frac{1}{2}p\right) \sum_{\nu}(-1)^{\nu} \eta_{2J}(\nu + 1) e^{-i\frac{1}{2}\nu p} \\ &\equiv (e^{-i\frac{1}{2}p} - 1)^J \hat{\eta}_J\left(\frac{1}{2}p\right) \lambda(p + 2\pi), \end{aligned} \tag{2.9.17}$$

where the Fourier sum

$$\lambda(p) = \sum_{\nu} \eta_{2J}(\nu+1)e^{-i\frac{1}{2}\nu p} \tag{2.9.18}$$

is 4π-periodic. Hence

$$\begin{aligned}
\sum_{\ell} |\widehat{\psi}(p+2\pi\ell)|^2 &= |e^{-i\frac{1}{2}p} - 1|^{2J} |\lambda(p+2\pi)|^2 \sum_{\ell'} \left|\hat{\eta}_J\left(\frac{1}{2}p + 2\pi\ell'\right)\right|^2 \\
&\quad + |e^{-i\frac{1}{2}p} + 1|^{2J} |\lambda(p)|^2 \sum_{\ell'} \left|\hat{\eta}_J\left(\frac{1}{2}p + \pi + 2\pi\ell'\right)\right|^2 \\
&= |e^{-i\frac{1}{2}p} - 1|^{2J} |\lambda(p+2\pi)|^2 R_{J-1}\left(\cos\left(\frac{1}{2}p\right)\right) \\
&\quad + |e^{-i\frac{1}{2}p} + 1|^{2J} |\lambda(p)|^2 R_{J-1}\left(-\cos\left(\frac{1}{2}p\right)\right),
\end{aligned} \tag{2.9.19}$$

and so

$$\begin{aligned}
u(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ipx}\widehat{\psi}(p)\left(\sum_{\ell}|\widehat{\psi}(p+2\pi\ell)|^2\right)^{-1} dp \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ipx}\hat{\eta}_J\left(\frac{1}{2}p\right)\lambda(p+2\pi)(e^{-i\frac{1}{2}p}-1)^J\Big[|e^{-i\frac{1}{2}p}-1|^{2J} \\
&\quad \times|\lambda(p+2\pi)|^2 R_{J-1}\left(\cos\left(\frac{1}{2}p\right)\right) + |e^{-i\frac{1}{2}p}+1|^{2J}|\lambda(p)|^2 \\
&\quad \times R_{J-1}\left(-\cos\left(\frac{1}{2}p\right)\right)\Big]^{-1} dp \\
&= -2\sum_m b_m \eta_J(2x-m),
\end{aligned} \tag{2.9.20}$$

where – following the general formulation in the previous section – we have

$$\begin{aligned}
b_m &= -\frac{1}{4\pi}\int_0^{4\pi} e^{i\frac{1}{2}pm}\lambda(p+2\pi)(e^{-i\frac{1}{2}p}-1)^J\Big[|e^{-i\frac{1}{2}p}-1|^{2J} \\
&\quad \times|\lambda(p+2\pi)|^2 R_{J-1}\left(\cos\left(\frac{1}{2}p\right)\right) + |e^{-i\frac{1}{2}p}+1|^{2J} \\
&\quad \times|\lambda(p)|^2 R_{J-1}\left(-\cos\left(\frac{1}{2}p\right)\right)\Big]^{-1} dp \\
&= \frac{1}{2\pi i}\int_{|z|=1} z^{-m-1}\sum_{\nu}\eta_{2J}(\nu+1)(-z)^{\nu}(z-1)^J\Big[(z-1)^J(z^{-1}-1)^J \\
&\quad \times\sum_{\nu,\nu'}\eta_{2J}(\nu+1)\eta_{2J}(\nu'+1)(-z)^{\nu-\nu'}R_{J-1}\left(\frac{1}{2}z+\frac{1}{2}z^{-1}\right) \\
&\quad + (z+1)^J(z^{-1}+1)^J\sum_{\nu,\nu'}\eta_{2J}(\nu+1)\eta_{2J}(\nu'+1)z^{\nu-\nu'}
\end{aligned}$$

$$\times R_{J-1}\left(-\frac{1}{2}z-\frac{1}{2}z^{-1}\right)\Bigg]^{-1}dz, \tag{2.9.21}$$

if we set $z = e^{-i\frac{1}{2}p}$. On the other hand,

$$b_m = \frac{1}{2\pi i}\int_{|z|=1} z^{m-1}\sum_\nu \eta_{2J}(\nu+1)(-z)^{-\nu}(z^{-1}-1)^J\begin{bmatrix}\text{same}\\ \text{expression}\end{bmatrix}^{-1} dz \tag{2.9.22}$$

if we set $z = e^{i\frac{1}{2}p}$. The coefficients are computed by residues; to avoid high-order poles at $z=0$, it is most convenient to use (2.9.21) for $m \le 0$ and to use (2.9.22) for $m \ge 1$. By (2.8.20) we may write (2.9.21) as

$$\begin{aligned} b_m &= \frac{1}{2\pi i}\int_{|z|=1} z^{4J-m-4}\sum_\nu \eta_{2J}(\nu+1)(-z)^\nu(z-1)^J(-1)^J[(1-z)^{2J} \\ &\times \sum_{\nu,\nu'}\eta_{2J}(\nu+1)\eta_{2J}(\nu'+1)(-z)^{2J+\nu-\nu'-2}K_{J-1}(z) \\ &-(z+1)^{2J}\sum_{\nu,\nu'}\eta_{2J}(\nu+1)\eta_{2J}(\nu'+1)z^{2J+\nu-\nu'-2}K_{J-1}(-z)]^{-1}dz, \end{aligned} \tag{2.9.23}$$

and (2.9.22) as

$$\begin{aligned} b_m &= \frac{1}{2\pi i}\int_{|z|=1} z^{2J+m-2}\sum_\nu \eta_{2J}(\nu+1)(-z)^{2J-\nu-2}(1-z)^J(-1)^J \\ &\times \begin{bmatrix}\text{same}\\ \text{polynomial}\end{bmatrix}^{-1} dz. \end{aligned} \tag{2.9.24}$$

The residues are determined by the zeros of the bracketed polynomial.

Since almost none of the coefficients b_m prove to be zero, the dual wavelet $u(x)$ cannot have compact support. On the other hand,

$$|b_m| \le ce^{-c|m|}, \tag{2.9.25}$$

so $u(x)$ is exponentially localized.

References

1. G. Battle, "Cardinal Spline Interpolation and the Block Spin Construction of Wavelets," in *Wavelets – A Tutorial in Theory and Applications*, C.K. Chui, ed., Academic Press, San Diego, 1992, pp. 73–93.

2. C. Chui and J. Wang, "On Compactly Supported Spline Wavelets and a Duality Principle," Trans. Amer. Math. Soc. **330** (1992), 903–915.

3. C. Chui and J. Wang, "A Cardinal Spline Approach to Wavelets," Proc. Amer. Math. Soc. **113** (1991), 785–793.

4. C. Chui and J. Wang, "A General Framework of Compactly Supported Splines and Wavelets," J. Approx. Theory **71** (1992), 263–304.

5. P. Lemarié-Rieusset, "Une Remarque sur les Propriétés Multi-Résolutions des Fonctions Splines," C.R. Acad. Sci. Paris **317** (1993), 1115–1117.

2.10 Multi-Dimensional Wavelets

The extension of the one-dimensional constructions to an arbitrary number of dimensions is quite easy, but it is based on tensor product combinations of the mother wavelet with the unit-scale-orthogonalized scaling function φ given by

$$\widehat{\varphi}(p) = \left(\sum_{\ell} |\hat{\eta}(p + 2\pi\ell)|^2\right)^{-1/2} \hat{\eta}(p). \tag{2.10.1}$$

Too often, the remark is made that a one-dimensional wavelet basis can be extended to a multi-dimensional basis by a tensor product. Such a remark is misleading because a straight tensor product of copies of a wavelet basis yields basis functions that can have length scale 2^{100} in the north–south direction and length scale 2^{-100} in the east–west direction! The point is that a multi-dimensional wavelet basis with only *one* scale parameter is more desirable – actually, a necessity from a practical point of view.

For d dimensions, consider the special vectors $\vec{\varepsilon} \in \{0,1\}^d \backslash \{\vec{0}\}$ and define

$$\Psi^{(\vec{\varepsilon})}(\vec{x}) = \prod_{\varepsilon_\iota = 0} \varphi(x_\iota) \prod_{\varepsilon_\iota = 1} \Psi(x_\iota). \tag{2.10.2}$$

The functions

$$\Psi^{(\vec{\varepsilon})}_{r,\vec{n}}(\vec{x}) = 2^{r\frac{d}{2}} \Psi^{(\vec{\varepsilon})}(2^r\, \vec{x} - \vec{n}) \tag{2.10.3}$$

with $r \in \mathbb{Z}$, $\vec{n} \in \mathbb{Z}^d$, constitute an orthonormal basis of $L^2(\mathbb{R}^d)$. The functions $\Psi^{(\vec{\varepsilon})}$ are all mother wavelets in this construction, and there are $2^d - 1$ of them; the function $\prod\limits_{\iota} \varphi(x_\iota)$ is not included among the mother wavelets.

The moments of a function $f(\vec{x})$ are said to vanish up to order M in d dimensions if the functions $f(\vec{x}) \prod\limits_{\iota=1}^{d} x^{s_\iota}$ are integrable and have vanishing integrals for $s_\iota \geq 0$ and $\sum\limits_{\iota=1}^{d} s_\iota \leq M$. Notice that if $f(\vec{x})$ happens to have the product form

$$f(\vec{x}) = \prod_{\iota=1}^{d} f_\iota(x_\iota), \tag{2.10.4}$$

then the moments of $f(\vec{x})$ vanish up to order M if and only if the moments of f_ι vanish up to order M for *some* ι. Since there is *at least one* factor of Ψ in (2.10.2), it follows that every multi-dimensional mother wavelet $\Psi^{(\vec{\varepsilon})}$ inherits the vanishing moments property that the one-dimensional mother wavelet Ψ happens to have. (Recall that the scaling function φ – in contrast to Ψ – does not even have a vanishing 0th-order moment.)

A function $f(\vec{x})$ is said to be class C^N in d dimensions if the functions

$$\left(\prod_{\iota=1}^{d} \frac{\partial^{s_\iota}}{\partial x_\iota^{s_\iota}}\right) f(\vec{x})$$

are continuous for $s_\iota \geq 0$ and $\sum_{\iota=1}^{d} s_\iota \leq N$. Clearly, each mother wavelet $\Psi^{(\vec{\varepsilon})}$ inherits the degree of smoothness possessed by φ and Ψ.

Each $\Psi^{(\vec{\varepsilon})}$ obviously has compact support if φ and Ψ do, so the existence of multi-dimensional Daubechies wavelets is no problem. In general, this multi-dimensional extension preserves long-distance decay properties. For example, if

$$|\varphi(x)| \leq c_0 e^{-m_0|x|}, \tag{2.10.5.0}$$

$$|\Psi(x)| \leq c_0 e^{-m_0|x|}, \tag{2.10.5.1}$$

then

$$\begin{aligned} |\Psi^{(\vec{\varepsilon})}(\vec{x})| &\leq c_0^d \exp\left(-m_0 \sum_{\iota=1}^{d} |x_\iota|\right) \\ &\leq c_0^d \exp\left(-m_0 \sqrt{\sum_{\iota=1}^{d} x_\iota^2}\right), \end{aligned} \tag{2.10.6}$$

so multi-dimensional Lemarié wavelets can be built from one-dimensional Lemarié wavelets, preserving even the rate of exponential decay. As another example, consider the decay property

$$|\varphi(x)| \leq c_0(1+x^2)^{-N}, \tag{2.10.7.0}$$

$$|\Psi(x)| \leq c_0(1+x^2)^{-N}. \tag{2.10.7.1}$$

In this case,

$$\begin{aligned} |\Psi^{(\vec{\varepsilon})}(\vec{x})| &\leq c_0^d \prod_{\iota=1}^{d} (1+x_\iota^2)^{-N} \\ &\leq c_0^d \left(1 + \sum_{\iota=1}^{d} x_\iota^2\right)^{-N}, \end{aligned} \tag{2.10.8}$$

so power laws are preserved as well. Finally, consider the smoothness and decay of a Meyer wavelet – i.e., the property $\widehat{\varphi}$, $\widehat{\Psi} \in C_0^\infty(\mathbb{R})$. Since

$$\widehat{\Psi^{(\vec{\varepsilon})}}(\vec{p}) = \prod_{\varepsilon_\iota=0} \widehat{\varphi}(p_\iota) \prod_{\varepsilon_\iota=1} \widehat{\Psi}(p_\iota), \tag{2.10.9}$$

$\widehat{\Psi^{(\vec{\varepsilon})}} \in C_0^\infty(\mathbb{R}^d)$ is an obvious property of the multi-dimensional mother wavelets in this case.

Having described the construction of orthonormal bases of wavelets for $L^2(\mathbb{R}^d)$, we wish to construct such bases for the Sobolev space defined by the norm

$$\|f\|_{2,s}^2 = \frac{1}{(2\pi)^d} \left(\prod_{\iota=1}^{d} \int_{-\infty}^{\infty} dp_\iota \right) | \vec{p} |^{2s} |\hat{f}(\vec{p})|^2, \tag{2.10.10}$$

where s is any positive integer. In particular,

$$\|f\|_{2,1}^2 = \int |\nabla f|^2, \tag{2.10.11}$$

and indeed, this will be the most appropriate Hilbert space norm for the renormalization group analysis of Euclidean field theory. Now observe that it is quite easy to obtain *a* wavelet basis that is orthonormal with respect to the inner product corresponding to (2.10.10). For a given wavelet basis $\{\Psi_k\}$ that is orthonormal with respect to the inner product for $L^2(\mathbb{R}^d)$, the functions $\Psi_{k,s}$ defined by

$$\widehat{\Psi_{k,s}}(\vec{p}) = | \vec{p} |^{-s} \, \widehat{\Psi}_k(\vec{p}) \tag{2.10.12}$$

certainly form a basis with the desired orthonormality. In this notation k ranges over the triples $(\vec{\varepsilon}, r, \vec{n})$ and

$$\Psi_{k,s}(\vec{x}) = 2^{(s-\frac{d}{2})r} \Psi_s^{(\vec{\varepsilon})}(2^r \, \vec{x} - \vec{n}). \tag{2.10.13}$$

The coherent structure with respect to discrete scale-commensurate translation is preserved because the transformation is just multiplication by $| \vec{p} |^{-s}$ in momentum space, while the coherence in scale is preserved by the homogeneity of $| \vec{p} |^{-s}$. As far as the regularity of $\Psi_s^{(\vec{\varepsilon})}$ is concerned, note that if

$$\widehat{\Psi^{(\vec{\varepsilon})}}(\vec{p}) \prod_{\iota=1}^{d} p_\iota^{s_\iota}$$

is integrable for $\sum_\iota s_\iota \leq N$, then so is

$$\widehat{\Psi^{(\vec{\varepsilon})}}(\vec{p}) | \vec{p} |^{-s} \prod_{\iota=1}^{d} p_\iota^{s_\iota}$$

for $\sum_\iota s_\iota \leq N + s$. This momentum-space description of how the regularity is affected is similar in spirit to the quasi-correct statement that if $\Psi^{(\vec{\varepsilon})}$ is class C^N then $\Psi_s^{(\vec{\varepsilon})}$ is class C^{N+s} (The latter statement is false without some technical condition, but the point is that (2.10.12) can only increase smoothness.) *The issue raised by this construction lies in the effect of* (2.10.12) *on long-distance decay properties.* For dimension $d > 1$, for example, this transformation obviously destroys real analyticity in momentum space – specifically at $\vec{p}=0$ – and therefore destroys exponential decay in position space.

On the other hand, (2.10.12) preserves the properties of Meyer wavelets. Recall from the construction in §2.2 that a one-dimensional Meyer wavelet Ψ not only has a compactly supported Fourier transform, but also has the property

$$\widehat{\Psi}(p)=0, \qquad |p|\leq\delta, \tag{2.10.14}$$

for some $\delta>0$. It follows from (2.10.9) that $\widehat{\Psi^{(\vec{\varepsilon})}}(\vec{p})$ vanishes in the region

$$\bigcup_{\varepsilon_\iota=1}\{\vec{p}\colon\ |p_\iota|\leq\delta\}\supset\{\vec{p}\colon\ |\vec{p}\,|\leq\delta\}, \tag{2.10.15}$$

so each $\Psi^{(\vec{\varepsilon})}$ certainly has the property

$$\widehat{\Psi^{(\vec{\varepsilon})}}(\vec{p})=0, \qquad |\vec{p}\,|\leq\delta. \tag{2.10.16}$$

This means that multiplication by $|\vec{p}\,|^{-s}$ introduces no singularities, so $\widehat{\Psi_s^{(\vec{\varepsilon})}}$ is class C^∞ as well.

Finally, (2.10.12) preserves the vanishing-moment property for all of the wavelets that have been constructed. If for $\sum_\iota s_\iota\leq M$

$$\left(\prod_{\iota=1}^{d}\frac{\partial^{s_\iota}}{\partial p_\iota^{s_\iota}}\right)\widehat{\Psi^{(\vec{\varepsilon})}}(\vec{p})$$

is continuous and vanishes at $\vec{p}=0$, then for $\sum_\iota s_\iota\leq M-s$

$$\begin{aligned}\left(\prod_{\iota=1}^{d}\frac{\partial^{s_\iota}}{\partial p_\iota^{s_\iota}}\right)\left(|\vec{p}\,|^{-s}\widehat{\Psi^{(\vec{\varepsilon})}}(\vec{p})\right) &= \left(\prod_{\iota=1}^{d}\sum_{\sigma_\iota=0}^{s_\iota}\right)\prod_{\iota=1}^{d}\binom{s_\iota}{\sigma_\iota}\\ &\times\left(\left(\prod_{\iota=1}^{d}\frac{\partial^{\sigma_\iota}}{\partial p_\iota^{\sigma_\iota}}\right)|\vec{p}\,|^{-s}\right)\left(\left(\prod_{\iota=1}^{d}\frac{\partial^{s_\iota-\sigma_\iota}}{\partial p_\iota^{s_\iota-\sigma_\iota}}\right)\widehat{\Psi^{(\vec{\varepsilon})}}(\vec{p})\right)\end{aligned} \tag{2.10.17}$$

is continuous and vanishes at $\vec{p}=0$ because

$$\left(\prod_{\iota=1}^{d}\frac{\partial^{\sigma_\iota}}{\partial p_\iota^{\sigma_\iota}}\right)|\vec{p}\,|^{-s}\leq c|\vec{p}\,|^{-s-\sum_\iota\sigma_\iota}, \tag{2.10.18}$$

$$\lim_{\vec{p}\to0}|\vec{p}\,|^{-\omega}\left(\prod_{\iota=1}^{d}\frac{\partial^{s_\iota-\sigma_\iota}}{\partial p_\iota^{s_\iota-\sigma_\iota}}\right)\widehat{\Psi^{(\vec{\varepsilon})}}(\vec{p})=0, \quad \sum_\iota(s_\iota-\sigma_\iota)\leq M-\omega. \tag{2.10.19}$$

Technically, the momentum-space observation is not quite the same as the assertion that if the moments of $\Psi^{(\vec{\varepsilon})}$ vanish up to order M, then the moments of $\Psi_s^{(\vec{\varepsilon})}$ vanish up to order $M-s$. Indeed, this latter implication is false without some additional condition. However, we have made the point that (2.10.12) does not destroy the vanishing-moment property, but only reduces the order up to which the moments vanish.

The strategy given by (2.10.12) fails to preserve the exponential decay of a Lemarié wavelet. It is worth noting, however, that in the special case $d = 1$, one can slightly alter this strategy. If we define the Sobolev wavelet Ψ_s by

$$\widehat{\Psi}_s(p) = p^{-s}\widehat{\Psi}(p), \tag{2.10.20}$$

then any analyticity that $\widehat{\Psi}(p)$ may have is shared by $\widehat{\Psi}_s(p)$, provided $\widehat{\Psi}(p)$ vanishes at $p = 0$ up to order $\geq s$. This means the construction produces one-dimensional Sobolev wavelets that are compactly supported as well as wavelets that are exponentially localized; one need only apply (2.10.20) to a Daubechies wavelet. We now turn to the problem of constructing a multi-dimensional Sobolev-orthonormal basis of wavelets that are exponentially localized.

The idea is to use the variational approach introduced in §2.7, where η is still given by (2.8.1), but now the Sobolev norm is the Hilbert space norm to be minimized on the affine manifold defined by the linear constraints. Naturally, each of the $2^d - 1$ constructions $\psi^{\vec{\varepsilon},s}$ corresponds to a certain set of constraints. Indeed, the constraints are determined by $\vec{\varepsilon}$ in a way that reflects (2.10.2)—specifically, $\psi^{\vec{\varepsilon},s}$ minimizes $\|\zeta\|_{2,s}$ with respect to the constraints

$$\int \prod_{\iota=1}^{d} \eta(2x_\iota - m_\iota)\zeta(\vec{x})d\,\vec{x} = \prod_{\varepsilon_\iota=0} c_{m_\iota} \prod_{\varepsilon_\iota=1} (-1)^{m_\iota} c_{1-m_\iota}. \tag{2.10.21}$$

The construction guarantees interscale orthogonality of the generated basis functions $\psi^{\vec{\varepsilon},s}(2^r\,\vec{x} - \vec{n})$ with respect to the Sobolev inner product. Indeed, for arbitrary ξ in the Sobolev space such that

$$\int \prod_{\iota=1}^{d} \eta(2x_\iota - m_\iota)\xi(\vec{x})d\,\vec{x} = 0, \qquad \vec{m} \in \mathbb{Z}^d, \tag{2.10.22}$$

$\zeta = \psi^{\vec{\varepsilon},s} + t\xi$ satisfies (2.10.21), so

$$\frac{d}{dt}\|\psi^{\vec{\varepsilon},s} + t\xi\|_{2,s}^2\Big|_{t=0} = 0 \tag{2.10.23}$$

by the variational definition of $\psi^{\vec{\varepsilon},s}$. But this derivative at $t = 0$ is just twice the real part of the Sobolev inner product of ξ with $\psi^{\vec{\varepsilon},s}$. Since (2.10.23) holds for $i\xi$ as well as for ξ, the imaginary part of the Sobolev inner product vanishes as well, and therefore $\psi^{\vec{\varepsilon},s}$ is Sobolev-orthogonal to all ξ such that (2.10.22) holds. On the other hand, (2.10.22) applies to all functions $\psi^{\vec{\varepsilon},s}(2^r\,\vec{x} - \vec{n})$ for which $r > 0$. This is a consequence of iterating the scale relation

$$\prod_{\iota=1}^{d} \eta(x_\iota) = \sum_{\vec{m}'} \left(\prod_{\iota=1}^{d} c_{m'_\iota}\right)\left(\prod_{\iota=1}^{d} \eta(2x_\iota - m'_\iota)\right) \tag{2.10.24}$$

and applying the algebraic identity

$$\sum_{m'_\iota} c_{m'_\iota} \sum_{m''_\iota} c_{m''_\iota} \cdots \sum_{m^{(r)}_\iota} c_{m^{(r)}_\iota} (-1)^{m^{(r)}_\iota} c_{1-2^r m_\iota + 2n_\iota - 2^{r-1} m'_\iota - 2^{r-2} m''_\iota - \cdots - m^{(r)}_\iota} = 0 \tag{2.10.25}$$

for an ι such that $\varepsilon_\iota = 1$.

To derive from this a Sobolev-orthonormal basis of wavelets, one simply applies the inverse square root of the overlap matrix to the sequence of integer translates of the $\psi^{\vec{\varepsilon},s}$. In momentum space, the mother wavelets are given by

$$\widehat{\Psi^{\vec{\varepsilon},s}}(\vec{p}) = \left(\sum_{\vec{\ell} \in \mathbb{Z}^d} |\widehat{\psi^{\vec{\varepsilon},s}}(\vec{p} + 2\pi \vec{\ell})|^2 \right)^{-1/2} \widehat{\psi^{\vec{\varepsilon},s}}(\vec{p}). \tag{2.10.26}$$

The question of how much smoothness and long-distance decay these wavelets have must now be addressed, as it tests the value of this construction. It is resolved by analyzing the expression for $\widehat{\psi^{\vec{\varepsilon},s}}(\vec{p})$.

The momentum expression is derived in exactly the same way for $\psi^{\vec{\varepsilon},s}$ as it was derived for the one-dimensional ψ in §2.7, except the factor $|\vec{p}|^{-s}$ is involved in the Poisson summation. The formula parallel to (2.7.16) is just

$$\widehat{\psi^{\vec{\varepsilon},s}}(\vec{p}) = \sum_{\vec{n}} \left(\prod_{\varepsilon_\iota = 0} c_{n_\iota} \right) \prod_{\varepsilon_\iota = 1} (2(-1)^{n_\iota} c_{1-n_\iota}) \left(\sum_{\vec{\ell}} \left| \frac{1}{2} \vec{p} + 2\pi \vec{\ell} \right|^{-2s} \right.$$
$$\left. \times \prod_{\iota=1}^{d} \left| \hat{\eta} \left(\frac{1}{2} p_\iota - 2\pi \ell_\iota \right) \right|^2 \right)^{-1} e^{-i\frac{1}{2}\vec{n}\cdot\vec{p}} |\vec{p}|^{-2s} \prod_{\iota=1}^{d} \hat{\eta} \left(\frac{1}{2} p_\iota \right), \tag{2.10.27}$$

which can also be written in the form

$$\widehat{\psi^{\vec{\varepsilon},s}}(\vec{p}) = \prod_{\varepsilon_\iota = 0} h(p_\iota) \prod_{\varepsilon_\iota = 1} (2e^{-i\frac{1}{2}p_\iota} h(p_\iota + 2\pi)^*) \left(\sum_{\vec{\ell}} \left| \frac{1}{2} \vec{p} + 2\pi \vec{\ell} \right|^{-2s} \right.$$
$$\left. \times \prod_{\iota=1}^{d} \left| \hat{\eta} \left(\frac{1}{2} p_\iota + 2\pi \ell_\iota \right) \right|^2 \right)^{-1} |\vec{p}|^{-2s} \prod_{\iota=1}^{d} \hat{\eta} \left(\frac{1}{2} p_\iota \right), \tag{2.10.28}$$

where the choice of η for the construction of exponentially localized wavelets is given by

$$\prod_{\iota=1}^{d} \hat{\eta}(p_\iota) = \prod_{\iota=1}^{d} \widehat{\chi}(p_\iota)^{N+1}$$
$$= i^{(N+1)d} \prod_{\iota=1}^{d} (e^{-ip_\iota} - 1)^{N+1} \prod_{\iota} \frac{1}{p_\iota^{N+1}}. \tag{2.10.29}$$

Now write

$$
\begin{aligned}
&\left|\frac{1}{2}\vec{p}\right|^{2s}\sum_{\vec{\ell}}\left|\frac{1}{2}\vec{p}+2\pi\vec{\ell}\right|^{-2s}\prod_{\iota}\left|\hat{\eta}\left(\frac{1}{2}p_\iota+2\pi\ell_\iota\right)\right|^2 \\
= \quad & (1+\Gamma_s(\vec{p}))\prod_{\iota}\left|\hat{\eta}\left(\frac{1}{2}p_\iota\right)\right|^2, && (2.10.30) \\
& \Gamma_s(\vec{p}) = \sum_{\vec{\ell}\neq 0}\frac{|\frac{1}{2}\vec{p}|^{2s}}{|\frac{1}{2}\vec{p}+2\pi\vec{\ell}|^{2s}}\prod_{\iota}\left(\frac{\frac{1}{2}p_\iota}{\frac{1}{2}p_\iota+2\pi\ell_\iota}\right)^{2N+2}, && (2.10.31)
\end{aligned}
$$

so that (2.10.28) assumes the form

$$
\begin{aligned}
\widehat{\psi^{\vec{\epsilon},s}}(\vec{p}) \quad = \quad & 4^s\sum_{\iota}\hat{\eta}\left(\frac{1}{2}p_\iota\right)^{*-1}(1+\Gamma_s(\vec{p}))^{-1}\prod_{\epsilon_\iota=0}h(p_\iota) \\
& \times\prod_{\epsilon_\iota=1}(2e^{-i\frac{1}{2}p_\iota}h(p_\iota+2\pi)^*). && (2.10.32)
\end{aligned}
$$

Recall that for this choice of η,

$$h(p_\iota) = (1+e^{-i\frac{1}{2}p_\iota})^{N+1} \tag{2.10.33}$$

and therefore

$$h(p_\iota+2\pi)^* = (1-e^{i\frac{1}{2}p_\iota})^{N+1}. \tag{2.10.34}$$

The replacement of each real variable p_ι with a complex variable z_ι analytically continues all of these functions of momentum. We have the analytic continuations

$$
\begin{aligned}
& |\vec{p}|^{2s} \longmapsto \left(\sum_{\iota} z_\iota^2\right)^s, \\
& \hat{\eta}(p_\iota) \longmapsto i^{N+1}(e^{-iz_\iota}-1)^{N+1}z_\iota^{-N-1}, \\
& \hat{\eta}(p_\iota)^* \longmapsto (-i)^{N+1}(e^{iz_\iota}-1)^{N+1}z_\iota^{-N-1}, \\
& h(p_\iota) \longmapsto (1+e^{-i\frac{1}{2}z_\iota})^{N+1}, \\
& h(p_\iota+2\pi)^* \longmapsto (1-e^{i\frac{1}{2}z_\iota})^{N+1}, \\
& \Gamma_s(\vec{p}) \longmapsto \Gamma_s(\vec{z}), \\
& \Gamma_s(\vec{z}) = \sum_{\vec{\ell}\neq 0}\left(\frac{\sum_\iota\left(\frac{1}{2}z_\iota\right)^2}{\sum_\iota\left(\frac{1}{2}z_\iota+2\pi\ell_\iota\right)^2}\right)^s\prod_{\iota}\left(\frac{\frac{1}{2}z_\iota}{\frac{1}{2}z_\iota+2\pi\ell_\iota}\right)^{2N+2}, && (2.10.35)
\end{aligned}
$$

which yield the analytic continuation $\widehat{\psi^{\vec{\epsilon},s}}(\vec{p}) \longmapsto f^{\vec{\epsilon},s}(\vec{z})$ with

$$\begin{aligned} f^{\vec{\varepsilon},s}(\vec{z}) &= 4^s i^{(N+1)d} \prod_\iota \left(\frac{\frac{1}{2} z_\iota}{e^{i\frac{1}{2}z_\iota} - 1} \right)^{N+1} (1 + \Gamma_s(\vec{z}))^{-1} \\ &\quad \times \prod_{\varepsilon_\iota = 0} (1 + e^{-i\frac{1}{2}z_\iota})^{N+1} \prod_{\varepsilon_\iota = 1} (2e^{-i\frac{1}{2}z_\iota}(1 - e^{i\frac{1}{2}z_\iota})^{N+1}). \quad (2.10.36) \end{aligned}$$

Obviously,

$$\lim_{z_\iota \to 0} \frac{\frac{1}{2} z_\iota}{e^{i\frac{1}{2}z_\iota} - 1} = -i, \quad (2.10.37)$$

and since $\Gamma_s(\vec{p})$ is manifestly positive, it is easy to bound Re $\Gamma_s(\vec{z})$ away from -1 on the region where:

(a) each z_ι is bounded away from $4\pi\ell$ for all $\ell \in \mathbb{Z}\backslash\{0\}$,

(b) Im z_ι is sufficiently small for each z_ι.

These observations imply that $f^{\vec{\varepsilon},s}(\vec{z})$ is analytic everywhere in the region

$$(\mathbb{R} + i[-\delta, \delta])^d \backslash \bigcup_\iota \left(\bigcup_{\ell \neq 0} \{ \vec{z} \colon \, |z_\iota - 4\pi\ell| < \delta \} \right)$$

for some small $\delta > 0$. The crucial point about (2.10.36) is that analyticity at $\vec{z} = 0$ is manifest.

To establish exponential decay of $\psi^{\vec{\varepsilon},s}(\vec{x})$, we need to show that $f^{\vec{\varepsilon},s}(\vec{z})$ is analytic in the Cartesian product region

$$(\mathbb{R} + i[-\delta, \delta])^d$$

and has some mild polynomial decay in Re $\vec{z}$ on that region. Pick an arbitrary non-zero $\vec{m}$ and observe that

$$\begin{aligned} &\left(\sum_\iota \left(\frac{1}{2} z_\iota + 2\pi m_\iota \right)^2 \right)^s \prod_\iota \left(\frac{1}{2} z_\iota + 2\pi m_\iota \right)^{2N+2} (1 + \Gamma_s(\vec{z})) \\ &= \left(\sum_\iota \left(\frac{1}{2} z_\iota \right)^2 \right)^s \prod_\iota \left(\frac{1}{2} z_\iota \right)^{2N+2} (1 + \Gamma_s(\vec{z} + 4\pi \vec{m})), \quad (2.10.38) \end{aligned}$$

so we may rewrite (2.10.36) as

$$\begin{aligned} f^{\vec{\varepsilon},s}(\vec{z}) &= 4^s i^{(N+1)d} \prod_\iota \left(\frac{\frac{1}{2} z_\iota + 2\pi m_\iota}{e^{i\frac{1}{2}z_\iota} - 1} \right)^{2N+2} \prod_\iota \left(\frac{e^{i\frac{1}{2}z_\iota} - 1}{\frac{1}{2} z_\iota} \right)^{N+1} \\ &\quad \times \left(\sum_\iota \left(\frac{1}{2} z_\iota \right)^2 \right)^{-s} \left(\sum_\iota \left(\frac{1}{2} z_\iota + 2\pi m_\iota \right)^2 \right)^s (1 + \Gamma_s(\vec{z} + 4\pi \vec{m}))^{-1} \\ &\quad \times \prod_{\varepsilon_\iota = 0} (1 + e^{-i\frac{1}{2}z_\iota})^{N+1} \prod_{\varepsilon_\iota = 1} (2e^{i\frac{1}{2}z_\iota}(1 - e^{-i\frac{1}{2}z_\iota})^{N+1}). \quad (2.10.39) \end{aligned}$$

Since

$$\lim_{z_\iota \to -4\pi m_\iota} \frac{\frac{1}{2}z_\iota + 2\pi m_\iota}{e^{i\frac{1}{2}z_\iota} - 1} = -i, \tag{2.10.40}$$

$$\lim_{z_\iota \to -4\pi m_\iota} \frac{e^{i\frac{1}{2}z_\iota} - 1}{\frac{1}{2}z_\iota} = i\delta_{m_\iota 0}, \tag{2.10.41}$$

the same reasoning as above shows that $f^{\vec{\varepsilon},s}(\vec{z})$ is analytic everywhere in the region

$$(\mathbb{R} + i[-\delta, \delta])^d \backslash \bigcup_\iota \left(\bigcup_{\ell \neq -m_\iota} \{\vec{z}\colon |z_\iota - 4\pi\ell| < \delta\} \right).$$

The factor $\left(\sum_\iota \left(\frac{1}{2}z_\iota\right)^2\right)^{-s}$ poses no problem, because this region is bounded away from $\vec{z} = 0$. In particular, (2.10.40) and (2.10.41) establish analyticity at $\vec{z} = -4\pi\,\vec{m}$. Now our choice of non-zero $\vec{m}$ is arbitrary, so if we take the union of all the regions of analyticity generated, we finally have the desired Cartesian product of strips centered on the real axis.

Why does $f^{\vec{\varepsilon},s}(\vec{z})$ have polynomial decay in $\mathrm{Re}\,\vec{z}$ for sufficiently small $\mathrm{Im}\,\vec{z}$? To answer this, consider the analytic continuation of the formula (2.10.28):

$$\begin{aligned} f^{\vec{\varepsilon},s}(\vec{z}) &= \prod_{\varepsilon_\iota = 0} (1 + e^{-i\frac{1}{2}z_\iota})^{N+1} \prod_{\varepsilon_\iota = 1} (2e^{-i\frac{1}{2}z_\iota}(1 - e^{i\frac{1}{2}z_\iota})^{N+1}) \\ &\quad \times \left(\sum_{\vec{\ell}} \left(\sum_\iota \left(\frac{1}{2}z_\iota + 2\pi\ell_\iota \right)^2 \right)^{-s} \prod_\iota \left(\frac{e^{-i\frac{1}{2}z_\iota} - 1}{\frac{1}{2}z_\iota + 2\pi\ell_\iota} \right)^{2N+2} \right)^{-1} \\ &\quad \times \left(\sum_\iota z_\iota^2 \right)^{-s} \prod_\iota \left(\frac{e^{-i\frac{1}{2}z_\iota} - 1}{\frac{1}{2}z_\iota} \right)^{N+1}. \end{aligned} \tag{2.10.42}$$

Every factor in this expression is periodic in the real directions except the last two factors. Moreover, the periodic contribution to this product is bounded for small $\mathrm{Im}\,\vec{z}$ because

$$\sum_{\vec{\ell}} \left(\sum_\iota \left(\frac{1}{2}z_\iota + 2\pi\ell_\iota \right)^2 \right)^{-s} \prod_\iota \left(\frac{e^{-i\frac{1}{2}z_\iota} - 1}{\frac{1}{2}z_\iota + 2\pi\ell_\iota} \right)^{2N+2}$$

is bounded away from zero in such a region. As for the last two factors,

$$\left(\sum_\iota z_\iota^2 \right)^{-s}$$

clearly decays in $\mathrm{Re}\,\vec{z}$ for small $\mathrm{Im}\,\vec{z}$, while

$$\left| \frac{e^{-i\frac{1}{2}z_\iota} - 1}{\frac{1}{2}z_\iota} \right| \le c|\mathrm{Re}\ z_\iota|^{-1}, \qquad |\mathrm{Im}\ z_\iota| \le \delta. \tag{2.10.43}$$

This decay serves two purposes. On one hand, it allows the contour-shifting argument for translating the real analyticity of $\widehat{\psi^{\vec{\varepsilon},s}}(\vec{p})$ into the exponential decay of $\psi^{\vec{\varepsilon},s}(\vec{x})$. On the other hand, it establishes that the smoothness of $\psi^{\vec{\varepsilon},s}(\vec{x})$ is class C^{N+2s-1}, so the wavelet we have constructed shares the properties of a Lemarié wavelet.

We return to (2.10.26) to consider the wavelets $\Psi^{\vec{\varepsilon},s}$ which generate an orthonormal basis instead of just an interscale-orthogonal basis. The analytic continuation $\widehat{\psi^{\vec{\varepsilon},s}}(\vec{p}) \longmapsto F^{\vec{\varepsilon},s}(\vec{z})$ is given by

$$F^{\vec{\varepsilon},s}(\vec{z}) = \left(\sum_{\vec{\ell}} f^{\vec{\varepsilon},s}(\vec{z}+2\pi\vec{\ell})f^{\vec{\varepsilon},s}(-\vec{z}-2\pi\vec{\ell})\right)^{-1/2} f^{\vec{\varepsilon},s}(\vec{z}), \tag{2.10.44}$$

which follows from inspecting the analytic continuation of $\widehat{\psi^{\vec{\varepsilon},s}}(\vec{p}+2\pi\vec{\ell})^*$ The proof that $\Psi^{\vec{\varepsilon},s}(\vec{x})$ shares the regularity and decay properties of $\psi^{\vec{\varepsilon},s}(\vec{x})$ now follows from arguments very similar to those given above.

References

1. G. Battle, "A Block Spin Construction of Ondelettes, Part I: Lemarié Functions," Commun. Math. Phys. **110** (1987), 601–615.

2. G. Battle, "A Block Spin Construction of Ondelettes, Part II: The QFT Connection," Commun. Math. Phys. **114** (1988), 93–102.

3. J. Kahane and P. Lemarié, *Fourier Series and Wavelets*, Gordon and Breach, Amsterdam, 1995.