We have to prove now that (3.38) defines an orthonormal wavelet. It belongs to W_0 by Lemma 3.1.7 and (3.55) with $j = -1$. We can write any $f \in W_{-1}$ as

$$\widehat{f}(\xi) = p(2\xi)\overline{\nu}(2\xi)\widehat{\psi}(2\xi) \tag{3.58}$$

If $\sum a_k e(-k\xi)$ is the Fourier expansion of the periodic function $p(\xi)\overline{\nu}(\xi)$, taking inverse Fourier transforms

$$f(x) = \frac{1}{2}\sum_{k\in\mathbb{Z}} c_k \psi(x/2 - k). \tag{3.59}$$

Then $\{\psi_{-1k}\}_{k\in\mathbb{Z}}$ spans W_{-1} and it is orthonormal proceeding as in (3.57). By (3.55), $\{\psi_{0k}\}_{k\in\mathbb{Z}}$ is an orthonormal basis of W_0 and as explained in (3.37) it is enough to assure that ψ is an orthonormal wavelet. □

In case you are wondering, it is possible to obtain examples of wavelets not associated to any MRA [Vya09, Th.3.4] but they cannot have a continuous Fourier transform with good decay [HW96, §7.3, Cor.3.16].

Suggested Readings. The rigorous development of the theory can be read in books addressed to a mathematical audience. My favorites are [Bré02], [Pin02] and [HW96]. The latter, one of the pioneering textbooks, takes quite effort to discuss some properties related to analysis, like convergence or atomic decomposition of functions. By reasons that probably rely on tradition, my feeling is that very rarely the literature about wavelets uses the most standard normalization of the Fourier transform (1.36), [Pin02] is an exception.

3.1.3 Construction of wavelets

In principle Theorem 3.1.4 or Corollary 3.1.5 offer a recipe to create wavelets: Take the scaling function of a MRA, compute the coefficients in (3.34) and you are done; applying the formulas the outcome must be an orthonormal wavelet. The problem is that it is not clear how difficult is to create a MRA. Let us say that one starts with a function ϕ such that $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ is orthonormal and define V_0 to fulfill the first property in (3.33). The second property can be also taken as a definition but the third property seems very hard to check. Why should $\bigcup V_j$ generate $L^2(\mathbb{R})$?

Let us adopt an optimistic attitude thinking about the case of the Haar wavelet. The scaling function was the characteristic function of $[0, 1)$ and it was as easy as saying that any L^2 function is a limit of step functions. If instead of the characteristic function of $[0, 1)$ we have something like a tent function, forgetting about the first property we have that the step functions are replaced by piecewise linear functions once we have adjusted the width of the tent function to avoid empty room. This also works and in general it seems that anything that can resemble a "lump" after scaling should work. It is actually true, under natural technical conditions, if finer scales are really finer and you avoid zero average functions the third property in (3.33) is assured.

Theorem 3.1.9. *Let $\phi \in L^2(\mathbb{R})$ be such that $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$, $\phi(x) = \sum_{k\in\mathbb{Z}} a_k \phi(2x-k)$ for some $\{a_k\}_{k\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$ and $\widehat{\phi}$ continuous at 0 with $\widehat{\phi}(0) \neq 0$. Then the spaces V_j spanned by $\{\phi(2^j \cdot -k)\}_{k\in\mathbb{Z}}$ define a MRA with scaling function ϕ.*

As a byproduct of the proof it follows that the only possibility to fulfill the hypotheses in this result is

$$|\widehat{\phi}(0)| = 1. \tag{3.60}$$

The second condition assures that we are not leaving "empty room" when considering a finer scale and the third is the zero average ban.

Proof. Let us consider the projection operator onto V_j

$$\pi_j f = \sum_{k\in\mathbb{Z}} \langle f, \phi_{jk}\rangle \phi_{jk} \qquad \text{with} \quad \phi_{jk}(x) = 2^{j/2}\phi(2^j x - k). \tag{3.61}$$

If we prove $\|\pi_j f\|_2 \to 0$ when $j \to -\infty$ for every $f \in L^2(\mathbb{R})$ we will have deduce $\bigcap V_j = \{0\}$ because $f \in \bigcap V_j$ would satisfy $\pi_j f = f$ for every j. Projections are bounded operators then, by density we can assume $f \in C_0^\infty$. If $[-M, M]$ contains the support of f

$$\|\pi_j f\|_2^2 = \sum_{k\in\mathbb{Z}} \big|\langle f, \phi_{jk}\rangle\big|^2 \le \|f\|_2^2 \sum_{k\in\mathbb{Z}} \int_{-M}^{M} |\phi_{jk}|^2 = \|f\|_2^2 \int_{U_j} |\phi|^2 \tag{3.62}$$

where $U_j = \bigcup_{k\in\mathbb{Z}} \big(-k - 2^j M, -k + 2^j M\big)$. As $\bigcap U_j$ has vanishing measure, we have $\|\pi_j f\|_2 \to 0$ by Lebesgue's dominated convergence theorem.

If $\overline{\bigcup V_j} \neq L^2(\mathbb{R})$ then taking $f \neq 0$ in the orthogonal complement of the union we would have $\|f - \pi_j f\|_2 = \|f\|_2$ for every j, then it is enough to prove that under our hypotheses

$$\lim_{j\to+\infty} \|f - \pi_j f\|_2 \neq \|f\|_2 \qquad \text{for any } f \in L^2(\mathbb{R}) - \{0\}. \tag{3.63}$$

Again by density we can assume $\widehat{f} \in C_0^\infty$. If $[-M, M]$ contains the support of $\widehat{f}$

$$\|f - \pi_j f\|_2^2 = \|f\|^2 - \|\pi_j f\|_2^2 = \|f\|^2 - \sum_{k\in\mathbb{Z}} \big|\langle f, \phi_{jk}\rangle\big|^2 = \|f\|^2 - \sum_{k\in\mathbb{Z}} \Big|\int_{-M}^{M} \widehat{f}\,\overline{\widehat{\phi}}_{jk}\Big|^2 \tag{3.64}$$

by Parseval identity. For j large enough we can replace M by 2^{j-1}, then the integral equals

$$\int_{-2^{j-1}}^{2^{j-1}} \widehat{f}(\xi) 2^{-j/2} \overline{\widehat{\phi}}(2^{-j}\xi) e\big(-2^{-j}k\xi\big)\, d\xi = 2^{-j/2} \int_{-1/2}^{1/2} \widehat{f}(2^j\xi)\overline{\widehat{\phi}}(\xi) e(-k\xi)\, d\xi. \tag{3.65}$$

This means that it is the k-th Fourier coefficient of the periodic extension of the function $2^{-j/2}\widehat{f}(2^j\xi)\overline{\widehat{\phi}}(\xi)$. Hence, Parseval identity again but this time for Fourier series, proves

$$\|f - \pi_j f\|_2^2 = \|f\|^2 - 2^{-j} \int_{-1/2}^{1/2} |\widehat{f}(2^j\xi)|^2 |\widehat{\phi}(\xi)|^2\, d\xi = \|f\|^2 - \int_{-M}^{M} |\widehat{f}(\xi)|^2 |\widehat{\phi}(2^{-j}\xi)|^2\, d\xi. \tag{3.66}$$

By the continuity, $\widehat{\phi}(2^{-j}\xi) \to \widehat{\phi}(0)$, therefore $\|f - \pi_j f\|_2^2 \to \|f\|_2^2 - |\widehat{\phi}(0)|^2 \|f\|_2^2$ and (3.63) follows.

Finally, note that $\overline{\bigcup V_j} = L^2(\mathbb{R})$ implies $\|f - \pi_j f\|_2 \to 0$ then necessarily $|\widehat{\phi}(0)| = 1$. □

Let us see how the application of Theorem 3.1.9 gives rise to specific examples of wavelets that one can compute numerically with high precision. We illustrate it with the *Meyer wavelet* and the *Franklin wavelet*. In both cases one looks for a trick to get automatically the needed scaling relation between $\phi(x)$ and $\phi(2x-k)$.

We can force ϕ to satisfy (3.29) adjusting $\widehat{\phi}$ as the square root of a real function. The idea of doing this fulfilling at the same time the scaling relation leads to *Meyer wavelets*. Following [Pin02, §6.4.4] we take any nonincreasing function $f:[0,1]\longrightarrow[0,1]$ satisfying

$$f(x)+f(1-x)=1 \qquad \text{and} \qquad f(x)=1 \quad \text{for } x\in[0,1/3]. \tag{3.67}$$

Then one defines

$$\phi(x)=\int_{-\infty}^{\infty}F(\xi)e(x\xi)\,d\xi \qquad \text{with} \quad F(\xi)=\begin{cases}\sqrt{f(|\xi|)} & \text{for } |\xi|\le 1,\\ 0 & \text{for } |\xi|\ge 1.\end{cases} \tag{3.68}$$

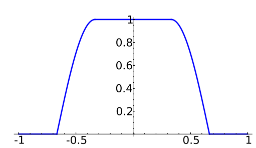
It is clear that F is as smooth as f that is our choice. Then $\phi=\widehat{F}$ is as smooth and as quickly decaying as we wish. Note that (3.29) reduces to

$$\sum_{k\in\mathbb{Z}}|F(\xi+k)|^2=f(|\xi|)+f(|1+\xi|)=f(-\xi)+f(1+\xi)=1 \qquad \text{for } -1\le\xi\le 0 \tag{3.69}$$

and by the periodicity it holds for every $\xi\in\mathbb{R}$. It is not difficult to check that under (3.67) we have $\widehat{\phi}(2\xi)=m_0(\xi)\widehat{\phi}(\xi)$ where m_0 is the periodic extension of $F(2\xi)$ for $|\xi|\le 1/2$. The key point is that $m_0(\xi)$ and $\widehat{\phi}(2\xi)$ both vanish for $\xi\in[1/3,2/3]$ and they coincide for $\xi\in[0,1/3]$ (see the details in [Pin02, §6.4.4]). Taking inverse Fourier transforms one concludes that we have a scaling relation for ϕ as required in Theorem 3.1.9.

For $f=\chi^*_{[-1/2,1/2]}$ restricted to $[0,1]$ one gets $\phi(x)=\operatorname{sinc} x$ and from it the Shannon wavelet (up to a translation), as we have seen. In this sense, the Shannon wavelet is the simplest of the Meyer wavelets. A nontrivial continuous choice of f due to Meyer is

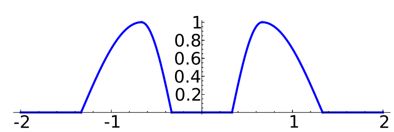
$$f(x)=\begin{cases}1 & \text{if } x\in[0,1/3],\\ \cos^2\big((3x-1)\pi/2\big) & \text{if } x\in[1/3,2/3],\\ 0 & \text{if } x\in[2/3,1].\end{cases} \qquad F(x)= \tag{3.70}$$



Using (3.68) and (3.38) with $\nu=1$ one obtains (see the calculations in [Bré02, D4·1])

(3.71)

$$\widehat{\psi}(\xi)=\begin{cases}-e^{i\pi\xi}\cos\big(\frac{3\pi}{2}|\xi|\big) & \text{if } \frac{1}{3}\le|\xi|\le\frac{2}{3},\\ e^{i\pi\xi}\sin\big(\frac{3\pi}{4}|\xi|\big) & \text{if } \frac{2}{3}\le|\xi|\le\frac{4}{3},\\ 0 & \text{otherwise.}\end{cases} \qquad e^{-i\pi\xi}\widehat{\psi}=$$

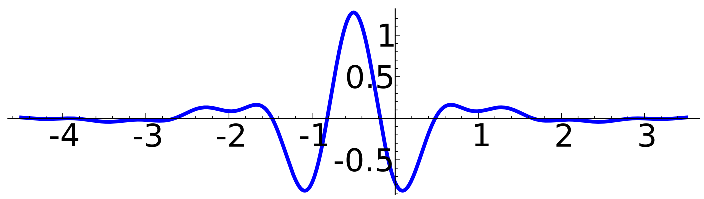


Inverting this Fourier transform is a tedious trivial computation giving

(3.72)

$$\psi(x)=\frac{1}{3\pi}F\Big(\frac{2x+1}{3}\Big) \qquad \text{with} \qquad F(x)=\frac{4\cos(\pi x)}{1-4x^2}+\frac{8\cos(4\pi x)}{1-16x^2}+\frac{24x\sin(2\pi x)}{(1-4x^2)(1-16x^2)}.$$

This is sometimes called the *Meyer wavelet*. The aspect of this fully explicit wavelet $\psi \in C^\infty \cap L^1$ is:



Another way of using Theorem 3.1.9 is looking for a function f with an easy relation between $f(x)$ and $f(2x-k)$, as in (3.39) to get the Haar wavelet, and adjust in some way (3.29). Let us choose the *tent function* $f(x) = \max(1-|x|, 0)$. It satisfies

$$f(x) = \frac{1}{2}f(2x+1) + f(2x) + \frac{1}{2}f(2x-1). \tag{3.73}$$

The set $\{f(\cdot - k)\}_{k\in\mathbb{Z}}$ generates the piecewise linear L^2 functions that equal their linear interpolation at the integers. This set is almost orthogonal because only contiguous elements have nonzero scalar product. It does not satisfy (3.29) but the Riesz system condition (3.31). In fact using Poisson summation formula it can be proved

$$\sum_{k\in\mathbb{Z}} |\widehat{f}(\xi+k)|^2 = \frac{1}{3} + \frac{2}{3}\cos^2(\pi\xi). \tag{3.74}$$

Now the trick is to force (3.29) dividing the Fourier transform by the square root of this quantity say $r(\xi)$. In this way our candidate for scaling function is $\widehat{\phi}(\xi) = \widehat{f}(\xi)/r(\xi)$. In general a relation like (3.73) is equivalent to $\widehat{f}(2\xi) = p(\xi)\widehat{f}(\xi)$ for $p \in L^2(\mathbb{T})$. Then $\widehat{\phi}(2\xi) = q(\xi)\widehat{\phi}(\xi)$ with $q(\xi) = r(\xi)p(\xi)/r(2\xi) \in L^2(\mathbb{T})$ or equivalently $\phi(x)$ can be expanded in terms of $\phi(2x-k)$ and Theorem 3.1.9 proves that we have a MRA (see also [Bré02, D4·2]). In our case, we have

$$\widehat{\phi}(\xi) = \frac{\sqrt{3}\widehat{f}(\xi)}{\sqrt{1+2\cos^2(\pi\xi)}} = \frac{\sqrt{3}\sin^2(\pi\xi)}{\pi^2\xi^2\sqrt{1+2\cos^2(\pi\xi)}}. \tag{3.75}$$

Note that Fourier inversion in the first equality allows to write $\phi(x) = \sum \lambda_k f(x-k)$ and this implies that ϕ is linear on each interval $[k, k+1]$, hence any $g \in V_1$ can be got from linear interpolation of the values $g(k/2)$. It particular it applies to the wavelet associated to ϕ.

Once we have the scaling function, (3.35) gives

$$m_0(\xi) = \sqrt{\frac{1+2\cos^2(\pi\xi)}{1+2\cos^2(2\pi\xi)}}\,\cos^2(\pi\xi). \tag{3.76}$$

Now, substituting in (3.1.4) with $\nu = 1$ and replacing ξ by $\xi/2$, we get

$$\widehat{\psi}(\xi) = e(\xi/2)h(\xi) \qquad \text{with} \quad h(\xi) = \frac{4\sqrt{3}\sin^4(\pi\xi/2)}{\pi^2\xi^2\sqrt{1+2\cos^2(\pi\xi)}}\sqrt{\frac{1+2\sin^2(\pi\xi/2)}{1+2\cos^2(\pi\xi/2)}}. \tag{3.77}$$

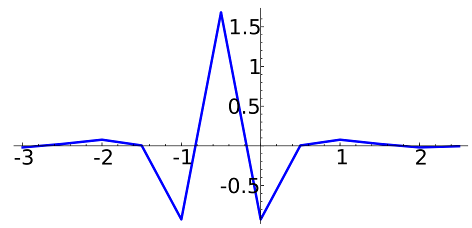
Finally, ψ can be obtained with linear interpolation of the values

$$\psi(k/2) = \int_{-\infty}^{\infty} h(\xi) e((k+1)\xi/2)\, d\xi = \int_{-\infty}^{\infty} h(\xi) \cos\big(\pi(k+1)\xi\big)\, d\xi. \tag{3.78}$$

Numerical calculations give
(3.79)

$$\begin{cases} \psi(-3) = \psi(2) &= -0.020 \\ \psi(-5/2) = \psi(3/2) &= 0.024 \\ \psi(-2) = \psi(1) &= 0.077 \\ \psi(-3/2) = \psi(1/2) &= 0.005 \\ \psi(-1) = \psi(0) &= -0.927 \\ \psi(-1/2) &= 1.682 \end{cases}$$

$\psi =$



The resulting wavelet is called the *Franklin wavelet* because in [Fra28], published in 1928, the ortogonalization of translations of the tent function was employed to get a continuous orthonormal system.

One could also apply the previous method to a piecewise quadratic function or to a cubic B-spline or to higher degree. This gives wavelets with higher regularity (they are called in general *spline wavelets*).

We finish this section mentioning the allegedly best method to construct wavelet. We do not deepen into it because it will reappear in some way in the discrete setting. The idea is that if we normalize $\phi(0) = 1$, which is harmless after (3.60), the iteration of (3.35) allows to define ϕ out of m_0 with

$$\widehat{\phi}(\xi) = \prod_{j=1}^{\infty} m_0(2^{-j}\xi). \tag{3.80}$$

Even if there are not convergence problems in principle it is not clear whether (3.29) holds. If we come back to the proof of Lemma 3.1.8 we see that it is equivalent to impose (3.50). Once this is assured the key point is the convergence. It is clear that we need $m_0(0) = 1$. In [Pin02, §6.5.1] it is required a non vanishing condition to avoid the product to diverge to zero and some strong continuity around zero. Namely, if we invent any $m_0 \in L^2(\mathbb{T})$ with $m_0(0) = 1$ satisfying (3.50) and the technical conditions

$$|m_0(\xi)| \ge c > 0 \quad \text{for } |\xi| \le \frac{1}{4} \qquad \text{and} \qquad |m_0(\xi) - 1| \log^2 \frac{1}{|\xi|} < c' \quad \text{for } |\xi| \le \frac{1}{2} \tag{3.81}$$

then (3.80) makes sense and gives the scaling function of a MRA and we can produce a wavelet with Theorem 3.1.4 or its corollary. Daubechies exploited the case in which $m_0(\xi)$ is a trigonometric polynomial. In this situation, the inverse Fourier transform of $m_0(2^{-j}\xi)$ are, having in mind (1.37), sums of Dirac deltas concentrated in smaller zones when j grows. Inverting (3.80) we get a convolution of them leading to a compactly supported scaling function ϕ and Corollary 3.1.5 proves that the corresponding wavelets has also compact support. The regularity of them is related to the degree. In this way one obtains localized wavelets with arbitrary regularity. The only drawback is that they do not admit explicit formulas. As we will see this is less important in practice when working in the discrete setting.

Suggested Readings. In the references already mentioned [Bré02], [Pin02] and [HW96], one can find the theory of the construction of wavelets. The case of compactly supported wavelets is discussed with detail in the two latter. In [Kam07, Ch.10] it is achieved a good balance between the development of the theory and the practice as a motivation (and the same applies to the whole book). One can find in the literature *biorthogonal wavelets*, a more general definition of wavelet dropping the condition of giving an orthonormal basis allowing more freedom for constructing them. They are discussed in [Mal09, §7.4].