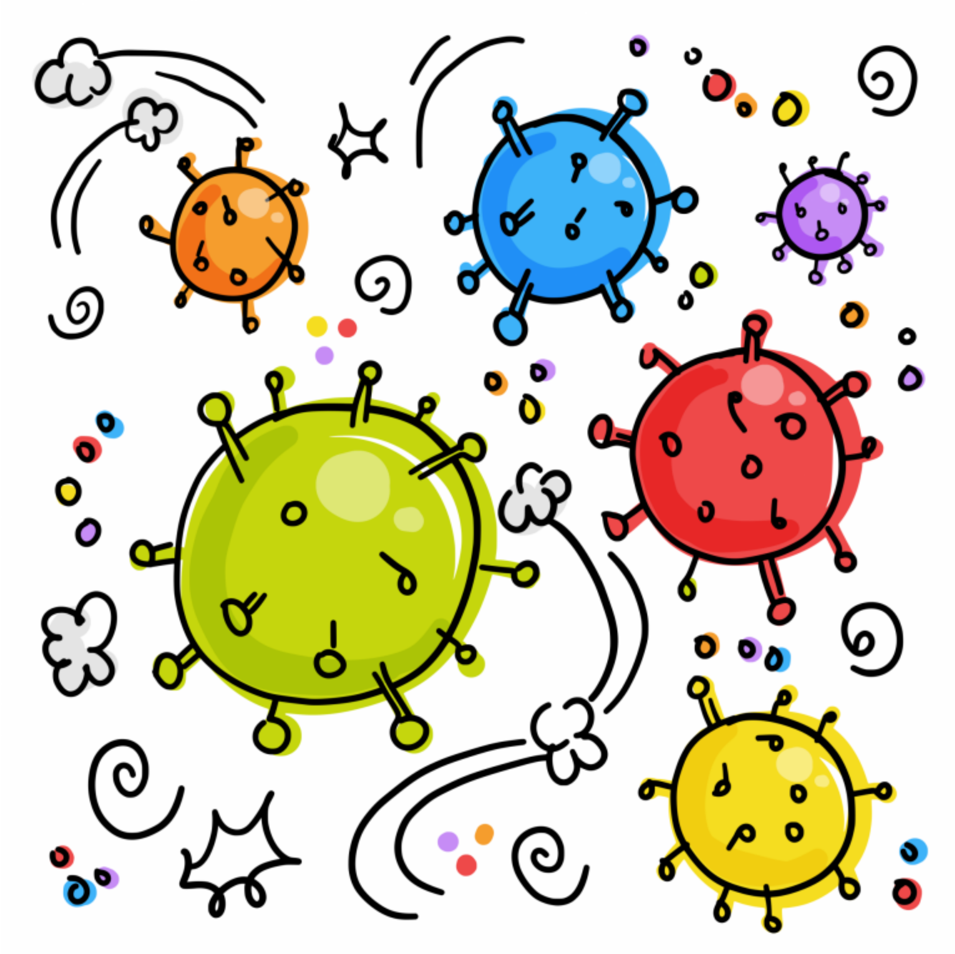
TMA4170: Corona Week 5

Wavelets

Miłosz Krupski



Based on
– A. Bogges, F. J. Narcowich *A First Course in Wavelets with Fourier Analysis*, Prentice-Hall 2001,
– M. A. Pinsky, *Introduction to Fourier Analysis and Wavelets*, AMS, 2009,
– S. Radhakrishnan (editor and author) *Wavelet Theory and Its Applications*, IntechOpen, 2018,
– A. Pietsch, *History of Banach spaces and linear operators.* Birkhäuser 2007.

1. Preliminaries

As we saw in several examples, Fourier analysis can be used to analyse verious measurements progressing in time – which works well if we deal with a situation where the observed data is periodic, with a period fixed. This is, however, an idealized situation:

- A seismic tremor (natural of artificial) consists of "fast" and "slow" waves, each providing valuable information, both in terms of their strenth and time of occurence, but they require separate "resolutions" for analysis.
- An ECG gives a periodic (rhytmic) reading in a healthy person, but once there is an *arrythmia*, it becomes even more important to obtain useful information from the cardiogram.
- A symphony consists of a large number of individual notes, each of which can be considered as a wave (periodic oscillation of air), but the result as a whole is seldom *repetitive.*

The wavelet analysis provides tools for addressing those issues. First, let us look at an example (still idealized) Here, we simply have frag-

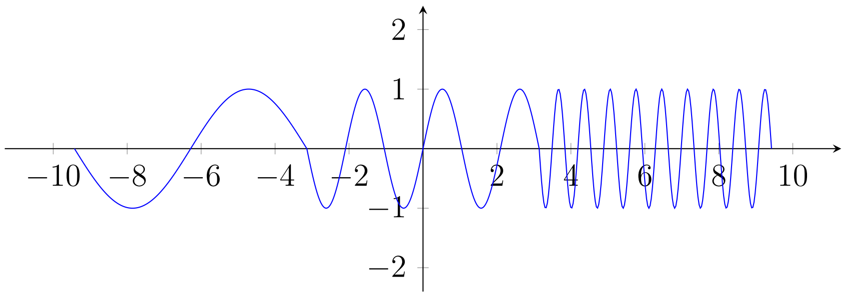


FIGURE 1. "Locally periodic" signal.

ments of $\sin x$, $\sin 3x$ and $\sin 9x$, glued together on consecutive intervals (and we can continue this over the entire line in some fashion). We can thus easily divide the line into separate intervals and perform the Fourier transform to discover the individual frequencies on each such interval (by "periodicizing" the function and using the framework of the **Fourier series**).

We can only do this, however, because we *know* how to divide the line (we can easily see where each "simple" signal begins and ends). If we didn't know this, we could imagine a moving "window" of a fixed length.

Mathematically, this procedure is the following:

$$g_t(\xi) = \int_{\mathbb{R}} f(x) \mathbb{1}_{[t,t+w]}(x) e^{ix\xi}\, dx,$$

where t is allowed to change and w is the fixed size of the window, while observing the change of values of g with respect to ξ allows us to discover when we have a "correct" reading of the frequency (thanks to the orthogonality relations). In our example, such a window still allows us to measure the frequencies when contained in each "base" interval, but it produces a garbage outcome when passing over a transition point – as a result, we want the window to be small, so that it does not happen too often.

On the other hand, when the window is too small, the frequency reading becomes unreliable – a small fragment of the sine function does not look like the sine function when extended in a periodic way to the whole line. In each window we need to observe several full oscillations, so that when the end values "do not match", it doesn't affect the reading too much.

We thus have to deal with two opposing forces: a larger window provides a more reliable measurement of the frequency spectrum and a smaller window allows us to better localize those measurements in time.

Instead of relying on the Fourier transform[1], there is a better approach.

2. Haar wavelets

2.1. Scaling function. Consider the function

$$\phi(x) = \mathbb{1}_{[0,1]}(x).$$

DEFINITION 2.1. For $n \in \mathbb{N}$, the space of step functions at level n, denoted by V_n, is defined to be the space of finite (real/complex-valued) linear combinations of functions in the set

$$\{\phi(2^n x - k)\}, \quad k \in \mathbb{Z}.$$

In other words, V_n is the space of piecewise constant (simple) functions of finite support whose discontinuities are contained in the set

$$\{\ldots - \tfrac{2}{2^n}, -\tfrac{1}{2^n}, 0, \tfrac{1}{2^n}, \tfrac{2}{2^n}, \tfrac{3}{2^n} \ldots\}.$$

We call those points the **dyadic** points.

A function in V_0 is a piecewise constant function with discontinuities contained in the set of integers. A function in V_n is also contained in V_{n+1}, hence we have

$$V_0 \subset V_1 \subset V_2 \subset \ldots V_{n-1} \subset V_n \subset V_{n+1} \ldots.$$

PROPOSITION 2.2.

- *$f(x) \in V_0$ if and only if $f(2^n x) \in V_n$.*
- *$f(x) \in V_n$ if and only if $f(2^{-n} x) \in V_0$.*

[1]The Fourier transform is still one of the most important tools for developing this theory and proving the theorems.

Notice that the functions $\phi(x-k)$ each have unit L^2-norm, namely

$$\|\phi(x-k)\|_2^2 = \int_{\mathbb{R}} \phi(x-k)^2\,dx = \int_k^{k+1} 1\,dx = 1.$$

Moroever,

$$\langle \phi(x-j), \phi(x-k)\rangle_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \phi(x-j)\phi(x-k)\,dx = \delta_{jk},$$

where $\delta_{jk} = 1$ if $j = k$ and $\delta_{jk} = 0$ if $j \neq k$.

Thus $\{\phi(x-k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis for V_0. In the same way, the set $\{2^n\phi(2^n x - k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis of V_n.

2.2. Haar wavelets. The spaces V_n and their orthonormal bases are only half of what we need. In order to "isolate" the behaviour of a function "on level n", we have to describe spaces $V_{n+1}\backslash V_n$, or better $V_{n+1} \cap (V_n)^{\perp}$.

The idea is to decompose V_{n+1} as an orthogonal sum of V_n and its complement. Let's start with $n = 0$. Since V_0 is generated by ϕ and its translates, it is reasonable to expect that the orthogonal complement of V_0 is generated by the translates of some function ψ. In such situation,

(1) $\psi \in V_1$ and so $\psi(x) = \sum_k a_k\phi(2x-k)$ for some choice of (finitely many) $a_k \in \mathbb{R}$;
(2) ψ is orthogonal to V_0, i.e. $\int_{\mathbb{R}} \psi(x)\phi(x-k)\,dx$ for all $k \in \mathbb{Z}$.

The simplest ψ satisfying both of these requirements is the function

$$\psi(x) = \phi(2x) - \phi(2x-1) = \mathbb{1}_{[0,1/2]}(x) - \mathbb{1}_{[1/2,1]}(x).$$

It is easy to notice that $\psi \in V_1$ and

$$\begin{aligned}\int_{\mathbb{R}} &\big(\mathbb{1}_{[0,1/2]}(x) - \mathbb{1}_{[1/2,1]}(x)\big)\mathbb{1}_{[k,k+1]}(x)\,dx \\ &= \lambda\big([k,k+1]\cap[0,1/2]\big) - \lambda\big([k,k+1]\cap[1/2,1]\big) = 0,\end{aligned}$$

since it can only be $0-0$ or $\frac{1}{2}-\frac{1}{2}$. Thus ψ is orthogonal to V_0.

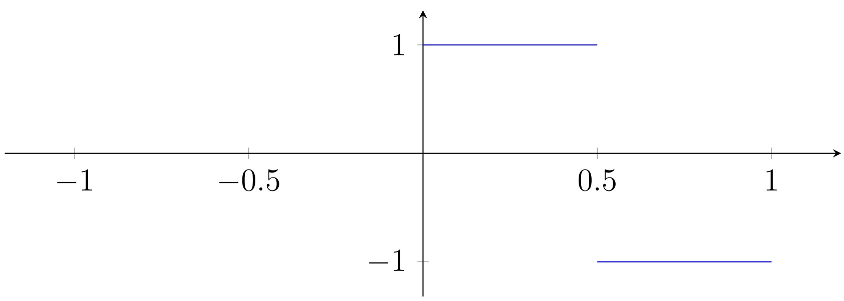


FIGURE 2. The Haar wavelet

DEFINITION 2.3. The function $\psi(x) = \phi(2x+1) - \phi(2x)$ is called the **Haar wavelet** (see Figure 2).

In the same way we may show that $\psi(x-k) \in V_1$ is orthogonal to V_0 for every $k \in \mathbb{Z}$. In fact, it is not difficult to notice that the space $W_0 = \left\{ \sum_k a_k \psi(x-k) \right\}$, spanned by finite linear combinations of $\psi(x-k)$ is the orthogonal complement of V_0 in V_1. Thus

$$V_1 = V_0 \oplus W_0,$$

and by the same scaling property as before,

$$V_n = V_{n-1} \oplus W_{n-1},$$

where $W_n = \left\{ \sum_k a_k 2^n \psi(2^n x - k) \right\}$ with $\{2^n \psi(2^n x - k)\}_{k \in \mathbb{Z}}$ being the orthonormal basis.

By induction we obtain

$$V_n = V_0 \oplus W_0 \oplus W_1 \oplus \ldots \oplus W_{n-1},$$

or in other words, each function $f \in V_n$ may be *uniquely* written as

$$f = f_0 + w_0 + w_1 + \cdots + w_{n-1},$$

where $w_j \in W_j$ and $f_0 \in V_0$.

As n increases to infinity, we have the following result.

THEOREM 2.4. *The space $L^2(\mathbb{R})$ can be decomposed as an infinite orthogonal direct sum*

$$L^2(\mathbb{R}) = V_0 \oplus W_0 \oplus W_1 \oplus \ldots = V_0 + \mathrm{cl}\left(\bigoplus_{n=0}^{\infty} W_n \right)$$

In particular, each $f \in L^2(\mathbb{R})$ can be written uniquely as

$$f = f_0 + \sum_{n=0}^{\infty} w_n,$$

where $f_0 \in V_0$ and $w_n \in W_n$.

This theorem may be proved directly, or you may use the Stone-Weierstrass theorem – but since the Haar functions are not continuous, and hence are not a sub-algebra of $C([a,b])$, you have to first "isolate" the dyadic points with neighbourhoods of small measure and then combine two (or three – the last one for the "tails" at infinities) different types of estimates. It's an optional exercise.

We can use this representation to "filter" the necessary information out of the given function by considering only the relevant "band" of spaces W_n. If, for example, we expect the signal to have a regular shape with short "spikes" representing noise (see Figure 3), we can cut the tail of the series, which only accounts for the spikes of limited width, but leave the rest of the function, which can be well approximated by wider blocks relatively unchanged.

Let us also notice that while we use V_0 as a "starting" space with "resolution" 1, and then descend to finer and finer divisions with spaces

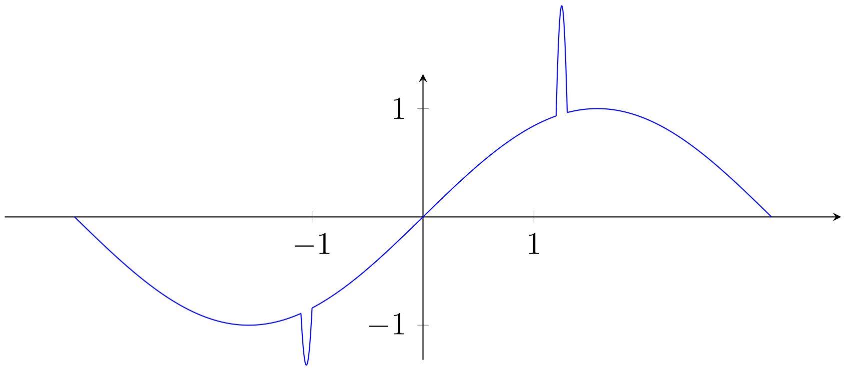


FIGURE 3. The signal and the noise.

V_n and W_n, in principle we may also go in the opposite direction and for $n \in \mathbb{N}$ define spaces V_{-n} and W_{-n} of step functions of increasing base size 2^n. Then

$$L^2(\mathbb{R}) = \mathrm{cl}\left(\bigcup_{n=-\infty}^{\infty} V_n\right) = \mathrm{cl}\left(\bigoplus_{n=-\infty}^{\infty} W_n\right).$$

3. Abstract wavelets

Haar wavelets are the simplest example, but we can abstract the basic ideas we developed and then study different sets of wavelets, which may be relevant in their specific applications.

The main problem with the Haar wavelets is the fact that they are not continuous (which is mostly fine for digital processing and they play a huge role in e.g. compression algorithms).

On the other extreme, we could think of having smooth wavelets. It turns out, however, that in such case they cannot be compactly supported (as the Haar wavelets are). Hence the moving "window" is never bounded, but it acts more as a lense which can magnify certain areas. The best we can do is to have smooth wavelets in the Schwartz class.

But perhaps the most interesting case is the intermediate one, when – according to our needs – we may trade-off the size of the window for the regularity of the function (the more derivatives the wavelet has, the wider its support has to be).

We will discuss several useful examples at the end of this section, but now let us abstract the general framework.

DEFINITION 3.1. An orthogonal **multiresolution analysis**. is a family of closed subspaces V_n of $L^2(\mathbb{R})$ indexed by $\mathbb{Z}$ that has the following properties:

(1) $\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots$;

(2) $\bigcap_{n\in\mathbb{Z}} V_n = \{0\}$ and $\mathrm{cl}\left(\bigcup_{n\in\mathbb{Z}} V_n\right) = L^2(\mathbb{R})$;

(3) $f(x) \in V_n$ if and only if $f(2x) \in V_{n+1}$ for $n \in \mathbb{Z}$.
(4) $f(x) \in V_0$ if and only if $f(x-k) \in V_0$ for $k \in \mathbb{Z}$.
(5) There exists a scaling function ϕ whose translates $\phi(x-k)$ with $k \in \mathbb{Z}$ form an orthonormal basis of V_0.

Of course, all these properties are satisfied for the sequence of spaces V_n we discussed in the Haar case. The other important ingredient are the wavelets themselves, which define (or are defined by) spaces W_n.

DEFINITION 3.2. A function ψ on $\mathbb{R}$ is called a **wavelet** if $\psi \in L^2(\mathbb{R})$ and the family of functions

$$\psi_{n,k}(x) = \sqrt{2^n}\psi(2^n x - k),$$

where $n, k \in \mathbb{Z}$, is an orthonormal basis of $L^2(\mathbb{R})$.

The wavelet function ψ may be obtained from the scaling function ϕ. Let $\phi \in V_0 \subset V_1$ be the scaling function of the multiresolution analysis $\{V_n\}_{n\in\mathbb{Z}}$. Since $\{\phi(2x-k)\}_{k\in\mathbb{Z}}$ is the orthonormal basis of V_1, there exists a sequence $\{c_k\} \in \ell^2(\mathbb{Z})$ such that

$$\phi(x) = \sum_{k\in\mathbb{Z}} c_k\, \phi(2x-k).$$

Then the required wavelet is obtained by

$$\psi(x) = \sum_{k\in\mathbb{Z}} (-1)^{k-1}\overline{c_k}\, \phi(2x+k-1).$$

Of course, the main problem is finding the correct coefficients c_k.

In the literature the (scaled) functions ϕ are sometimes called *fathers* and the functions ψ – *mothers*, but from the biological perspective it makes no sense. It's rather a process of mitosis.

Different examples of wavelets include (**here without** the correct normalizing constants that make them have a unit L^2-norm!)

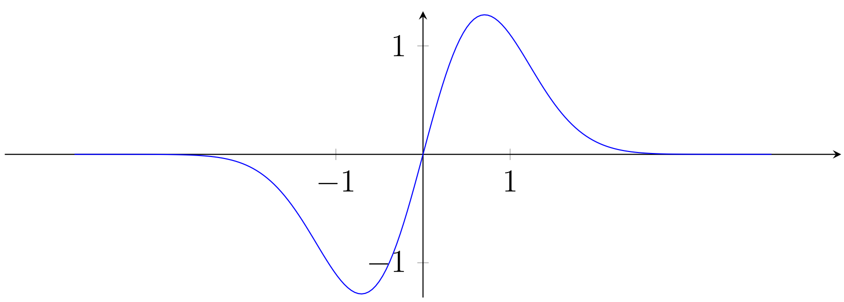


FIGURE 4. The Gaussian wavelet $\psi(x) = xe^{-x^2}$.

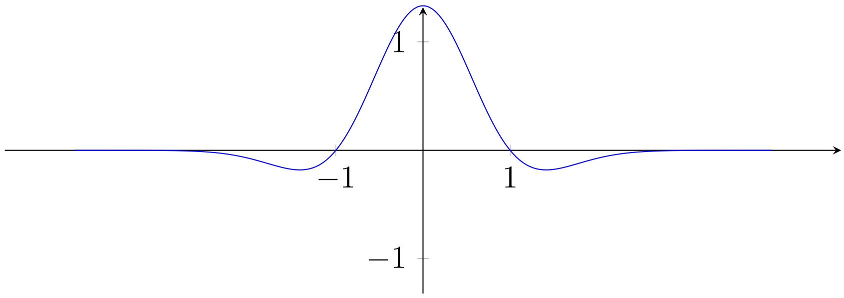


FIGURE 5. The Ricker wavelet (also known as the Mexican hat wavelet) $\psi(x) = (1 - x^2)e^{-x^2}$.

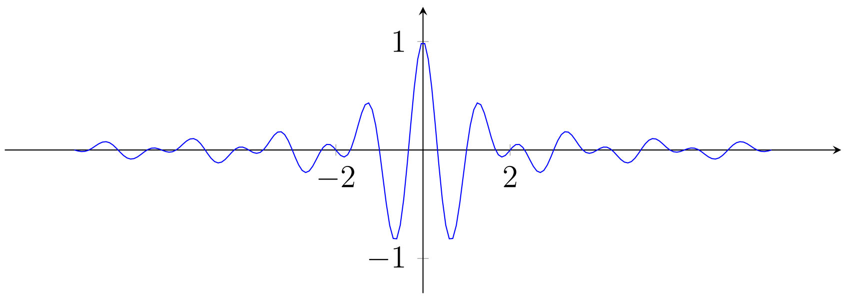


FIGURE 6. The Shannon wavelet (Fourier dual of the Haar wavelet).

4. Non-orthogonal multiresolution analysis

This is an optional section. The assumption of orthogonality of the basis defined by the scaling function ϕ is sometimes to restrictive, and an unnecessary one. To resolve this issue, we define a Riesz system.

DEFINITION 4.1. In a Hilbert space $\mathcal{H}$ a set of vectors $\{x_n\}$ is **a Riesz system**, if there exist constants $0 < A \leqslant B < \infty$ such that for any finite set of complex numbers $\{a_n\}$ we have

$$A \sum_n |a_n|^2 \leqslant \left\| \sum_n a_n x_n \right\|_{\mathcal{H}}^2 \leqslant B \sum_n |a_n|^2.$$

Clearly, an orthonormal system is a Riesz system if and only if $A = B = 1$ (for the optimal choice of A and B). A Riesz system is a Riesz basis if in addition it is a basis, i.e. the relevant linear combinations are dense.

Then we can define a non-orthogonal multiresolution analysis by replacing condition (5) with

(5') There exists a scaling function ϕ whose translates $\phi(x - k)$ with $k \in \mathbb{Z}$ form a Riesz basis of V_0.

As an example we may consider as the scaling function ϕ a tent function, which is compactly supported, like the Haar wavelet, but also continuous.

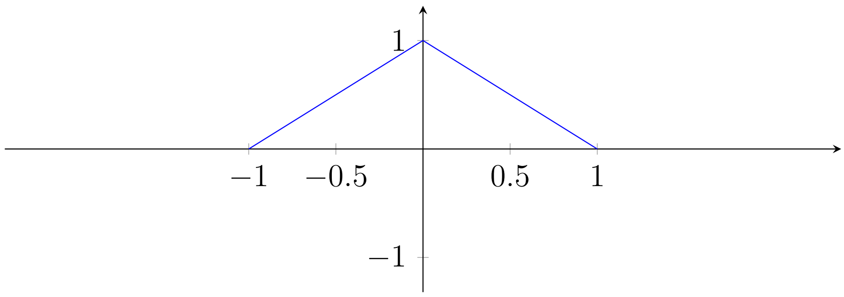


FIGURE 7. The tent function.

In this case we have $A = 1/3$ and $B = 1$, so the resulting basis is not orthonormal.

The tent function is also the simplest example of a spline, which may be specifically designed for the given purpose. A spline is a continuous function with compact support, which is piecewise linear, or is built with other polynomials of a given degree.

5. Wavelet transform

To conclude, we define an analogue of the Fourer transform– the wavelet trasform. It is a sort of a generalization of MRA.

DEFINITION 5.1. Let ψ be continuous wavelet function. We define the **wavelet transform**

$$W_f(a,b) = \frac{1}{\sqrt{|a|}} \int_{\mathbb{R}} f(x) \overline{\psi\left(\frac{x-b}{a}\right)}\, dx.$$

As in the case of the Fourier transform, a basic question concerns the inverse transform. It is given by the formula

$$f(x) = \frac{1}{C_\psi} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{a^2\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right) W_f(a,b)\, db\, da,$$

where

$$C_\psi = 2\pi \int_{\mathbb{R}} \frac{|\widehat{\psi}(\xi)|^2}{|\xi|}\, d\xi.$$

Questions:

- Do you understand the problem of calculating the Fourier transform of a progressing signal, which is not periodic, but may be "locally periodic"?
- Do you understand the idea behind the multiresolution analysis and the difference between the sequences of spaces $\{V_n\}$ and $\{W_n\}$?

Problems:

PROBLEM 1. Prove that $L^2(\mathbb{R}) = V_0 + \mathrm{cl}\left(\bigoplus_{n\in\mathbb{N}} W_n\right)$, where V_0 and W_n are defined for the Haar wavelets.

PROBLEM 2. Let $\phi(x) = (1-|x|)^+$ be the tent function. Show that $\{\phi(x-k)\}_{k\in\mathbb{Z}}$ is not an orthonormal basis, but it is a Radon basis.