EECS 123 Digital Signal Processing University of California, Berkeley: Fall 2007
Gastpar November 14, 2007

Transforms II - Wavelets

Preliminary version — please report errors, typos, and suggestions for improvements

We follow an approach similar to [1, 2]. The emphasis of this document is on a conceptual understanding. Not all statements are mathematically entirely correct, but they definitely hold for sufficiently "nice" functions, such as the ones that are typically used in time-frequency and wavelet analysis.

1 The General Transform

1.1 Definition

A useful general way of thinking of transforms is in the shape of inner products with a set of "basis" functions:

$$T_x(\gamma) = \langle x(t), \phi_\gamma(t) \rangle \quad (1)$$

$$= \int_{-\infty}^{\infty} x(t)\phi_\gamma^*(t)dt, \quad (2)$$

where * denotes the complex conjugate.

The idea here is that 'T' denotes what kind of "basis" functions are being used and γ is the index of a basis function. The basis functions are $\phi_\gamma(t)$ for all values of γ.

A good way of thinking about this is that for a fixed γ, the transform coefficient $T_x(\gamma)$ is the result of projecting the original signal $x(t)$ onto the "basis" element $\phi_\gamma(t)$.

An example is the Fourier transform, where instead of the letter γ, we more often use the letter Ω, and where $\phi_\Omega(t) = e^{j\Omega t}$. Hence, in line with the above general notation, we could write

$$FT_x(\Omega) = \langle x(t), \phi_\Omega(t) \rangle \quad (3)$$

$$= \int_{-\infty}^{\infty} x(t)e^{-j\Omega t}dt, \quad (4)$$

Of course, we more often simply write $X(\Omega)$ (or $X(j\Omega)$) in place of $FT_x(\Omega)$.

1.2 Alternative Formulation

For our next step, we need the (general) Parseval/Plancherel formula, which asserts that

$$\int_{-\infty}^{\infty} f(t)g^*(t)dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(j\Omega)G^*(j\Omega)d\Omega. \quad (5)$$

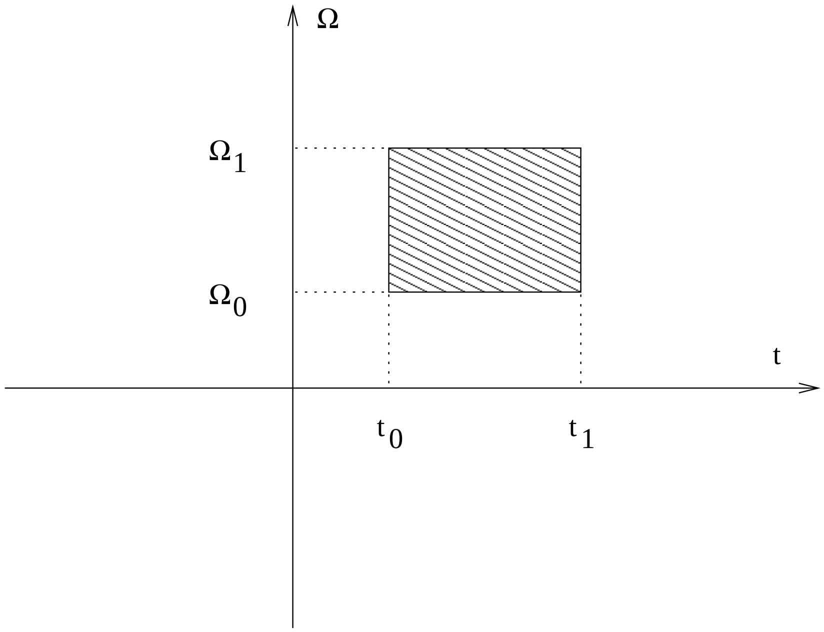

Figure 1: A conceptual picture: We imagine that the basis element $\phi_\gamma(t)$ only lives in the shaded box, i.e., that the signal is very small outside the interval $t_0 \leq t \leq t_1$, and that its spectrum $\Phi_\gamma(j\Omega)$ is very small outside of the interval $\Omega_0 \leq \Omega \leq \Omega_1$.

Using this, we can rewrite the general transform as

$$\begin{aligned} T_x(\gamma) &= \langle x(t), \phi_\gamma(t) \rangle && (6) \\ &= \int_{-\infty}^{\infty} x(t)\phi_\gamma^*(t)dt && (7) \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} X(j\Omega)\Phi_\gamma^*(j\Omega)dt && (8) \\ &= \langle X(j\Omega), \frac{1}{2\pi}\Phi_\gamma(j\Omega) \rangle && (9) \end{aligned}$$

Hence, we now have *two* good ways of thinking about transforms: For a fixed γ, the transform coefficient $T_x(\gamma)$ is the result of projecting the original signal $x(t)$ onto the "basis" element $\phi_\gamma(t)$, and *equivalently*, of projecting the original spectrum $X(j\Omega)$ onto the spectrum of the "basis" element $\phi_\gamma(t)$, which is $\frac{1}{2\pi}\Phi_\gamma(j\Omega)$.

Consider Figure 1: Merely as a though experiment, let us think of a "basis" element $\phi_\gamma(t)$ that lives[1] only inside the box illustrated in Figure 1. Then, a great way of thinking about the transform coefficient $T_x(\gamma)$ is that it tells us "how much" of the original signal $x(t)$ sits inside that box.

In lines with this intuition, for the Fourier transform, the transform coefficient $T_x(\Omega)$ tells us "how much" of the original signal $x(t)$ sits at frequency Ω, and the "box" shown in Figure 1 is infinitesimally thin in frequency and infinitely long in time.

[1] In the next section, we will make precise what "lives" means.

1.3 The Heisenberg Box Of A Signal

Reconsider the conceptual picture given in Figure 1. Now, we want to make this precise. In order to do so, consider any signal $\phi(t)$. For simplicity (and without loss of generality), we assume that the signal is "normalized" such that

$$\int_{-\infty}^{\infty} |\phi(t)|^2 dt \quad = \quad 1. \tag{10}$$

Note that by Parseval, this also means that $\frac{1}{2\pi}\int_{-\infty}^{\infty} |\Phi(j\Omega)|^2 d\Omega = 1$.

We define the following quantities. The "middle" of the signal $\phi(t)$ is given by

$$m_t \quad = \quad \int_{-\infty}^{\infty} t|\phi(t)|^2 dt. \tag{11}$$

If you have taken a class in probability, you will recognize this to be the mean value of the distribution $|\phi(t)|^2$.

Similarly, we define the "middle" of the spectrum $\Phi(j\Omega)$ to be

$$m_\Omega \quad = \quad \int_{-\infty}^{\infty} \Omega \frac{1}{2\pi} |\Phi(j\Omega)|^2 d\Omega, \tag{12}$$

with a similar probability interpretation.

Moreover, we define:

$$\sigma_t^2 \quad = \quad \int_{-\infty}^{\infty} (t-m_t)^2 |\phi(t)|^2 dt, \tag{13}$$

$$\sigma_\Omega^2 \quad = \quad \int_{-\infty}^{\infty} (\Omega - m_\Omega)^2 \frac{1}{2\pi} |\Phi(j\Omega)|^2 d\Omega. \tag{14}$$

Again, these can be understood as the respective *variances* of the two "probability distributions."

With these definitions, we can now draw a more precise picture of the time-frequency box of the signal $\phi(t)$, as given in Figure 2.

We should also point out that for the Fourier transform, the basis functions are of the form $\phi(t) = e^{j\Omega_0 t}$, and for those, the above integrals do not all converge, so special care is required mathematically. However, the right intuition is to say that the Heisenberg box (the term appears in [2], and perhaps earlier) of the function $\phi(t) = e^{j\Omega_0 t}$ is a horizontal line at frequency Ω_0.

1.4 The Uncertainty Relation

So, what are the possible Heisenberg boxes?

Theorem 1 (uncertainty relation)**.** *For any function* $\phi(t)$*, the Heisenberg box must satisfy*

$$\sigma_t \sigma_\Omega \quad \geq \quad \frac{1}{2}. \tag{15}$$

That is, Heisenberg boxes cannot be too small. Or: transforms cannot have a very high time resolution *and* a very high frequency resolution at the same time. (Proof: see class.)

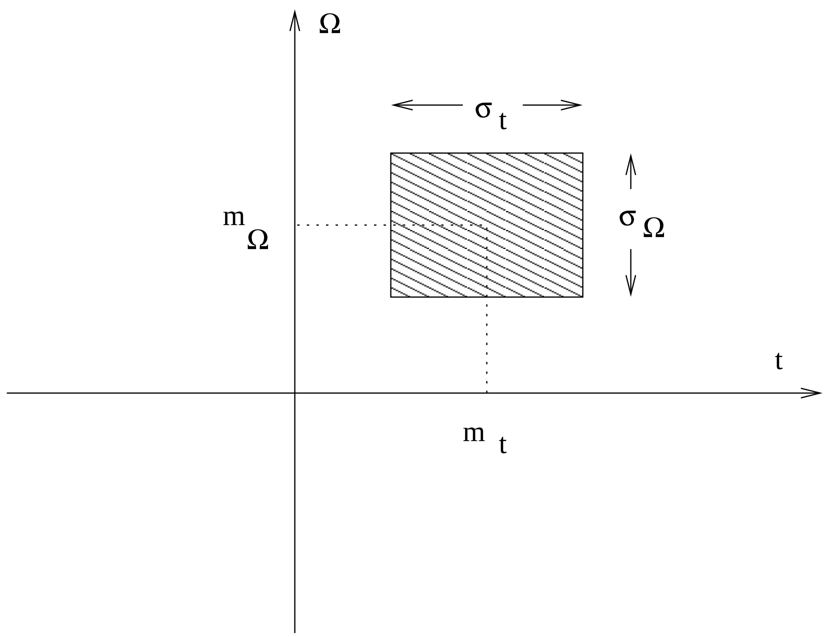

Figure 2: The Heisenberg box of the function $\phi(t)$ (i.e., the place in time and frequency where the function $\phi(t)$ is really alive).

2 The Short-time Fourier Transform

It has long been recognized that one of the most significant drawbacks of the Fourier transform is its lack of *time localization:* An event that is localized in time (such as a signal discontinuity) affects all of the frequencies (remember the Gibbs phenomenon). This feature is clearly undesirable for many engineering tasks, including compression and classification.

To regain some of the time localization, one could do a "short-time" Fourier transform, essentially chopping up the signal into "short" pieces and taking Fourier transforms separately for each piece. Kind of trivially, this gives back some time localization.

More generally, the following form can be given:

$$STFT_x(\tau, \Omega) = \int_{-\infty}^{\infty} x(t) g^*(t-\tau) e^{-j\Omega t} dt, \tag{16}$$

where the function $g(t)$ is an appropriate "window" function that cuts out a piece of the signal $x(t)$. With the parameter τ, we can place the window wherever we want.

With regard to the general transform, here, instead of the letter γ, we use the pair (τ, Ω), and

$$\phi_{\tau,\Omega}(t) = g(t-\tau) e^{j\Omega t}. \tag{17}$$

Many different window functions $g(t)$ are being used, but one of the easiest to understand is the Gaussian window:

$$g(t) = \frac{1}{\sqrt[4]{\pi\sigma^2}} e^{-\frac{t^2}{2\sigma^2}}. \tag{18}$$

Note that strictly speaking, this window is never zero, so it does not really "cut" the signal. However, if $|t|$ is large, $g(t)$ is tiny, so this is "almost the same as zero," but much easier to analyze. With this window, we find the "basis" elements to be

$$\phi_{\tau_0,\Omega_0}(t) \quad = \quad \frac{1}{\sqrt[4]{\pi\sigma^2}} e^{-\frac{(t-\tau_0)^2}{2\sigma^2}} e^{j\Omega_0 t}. \tag{19}$$

Now, we want to find explicitly the Heisenberg box of this "basis" function. To this end, we need the Fourier transform of the Gaussian window, which is known to be

$$G(j\Omega) \quad = \quad \sqrt[4]{4\pi\sigma^2} e^{-\frac{\Omega^2\sigma^2}{2}}, \tag{20}$$

and thus, using the standard time- and frequency-shift properties of the Fourier transform,

$$\Phi_{\tau_0,\Omega_0}(j\Omega) \quad = \quad \sqrt[4]{4\pi\sigma^2} e^{-\frac{(\Omega-\Omega_0)^2\sigma^2}{2}} e^{-j\Omega\tau_0}. \tag{21}$$

Now, we can find the corresponding parameters of the Heisenberg box as:

$$m_t \quad = \quad \tau_0, \tag{22}$$

$$m_\Omega \quad = \quad \Omega_0 \tag{23}$$

$$\sigma_t^2 \quad = \quad \frac{\sigma^2}{2} \tag{24}$$

$$\sigma_\Omega^2 \quad = \quad \frac{1}{2\sigma^2}, \tag{25}$$

and so, we can draw the corresponding Figure 2. It is also interesting to note that for the Gaussian window, the Heisenberg uncertainty relation (Theorem 1) is satisfied with equality. It can be shown that the Gaussian window is (essentially) the only function that satisfies the uncertainty relation with equality, see e.g. [2, p.31].

In class, we also saw the resulting time-frequency plots for speech signals (see the separate figures that were handed out in class).

3 Wavelet Transforms

This is an interesting case of the general transform where we start form a *single* function

$$\psi(t), \tag{26}$$

sometimes called the *mother wavelet.*

Then, we build up our "dictionary" by shifting and scaling the mother wavelet, specifically,

$$\psi_{m,n}(t) \quad = \quad 2^{-m/2}\psi(2^{-m}t - n), \tag{27}$$

where n and m are arbitrary *integers* (positive, negative, or zero). That is, in place of the parameter γ, we will use the pair of integers (m, n), where m is the scale (the bigger the coarser) and n is the shift. We will often denote the transform coefficient as

$$a_{m,n} \quad = \quad WT_x(m,n) = \langle x(t), \psi_{m,n}(t)\rangle. \tag{28}$$

The following key questions are of obvious interest:

- What are the conditions such that we can recover the original signal $x(t)$ from the wavelet coefficients $a_{m,n}$?
- How do we design good mother wavelets $\psi(t)$?
- How do we efficiently compute the wavelet coefficients $a_{m,n}$ for a given signal $x(t)$?
- and of course many more...

3.1 The Haar Wavelet

We will start with the Haar wavelet:

$$\psi(t) = \begin{cases} 1, & \text{for } 0 \le t < \frac{1}{2}, \\ -1, & \text{for } \frac{1}{2} \le t < 1, \\ 0, & \text{otherwise.} \end{cases} \tag{29}$$

Exercise: Sketch $\psi_{0,0}(t), \psi_{1,0}(t), \psi_{-1,0}(t)$, and $\psi_{-1,3}(t)$.

Facts:

1. The functions $\psi_{m,n}(t)$, taken over all integers m and n, are an orthonormal set. (Easy to verify.)
2. The functions $\psi_{m,n}(t)$, taken over all integers m and n, are in fact an orthonormal basis for $L^2(\mathbb{R})$, the space of all functions $x(t)$ for which $\int_{-\infty}^{\infty} |x(t)|^2 dt$ is finite. (This is more difficult to prove, see e.g. [2, ch.7].)

Due to these facts, we can express any square-integrable function $x(t)$ in the following form:

$$x(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} a_{m,n} \psi_{m,n}(t), \tag{30}$$

which will be called the *Haar expansion* (of more generally, the wavelet expansion) of the signal $x(t)$. Moreover, due to the orthogonality, we also know that

$$a_{m,n} = \langle x(t), \psi_{m,n}(t) \rangle = \int_{-\infty}^{\infty} x(t) \psi_{m,n}(t) dt. \tag{31}$$

To understand how this wavelet works, it is instructive to consider a piecewise constant function, as in Figure 3. We here follow the development in [1, p.211]. Specifically, consider the function $f^{(0)}(t)$ which is piecewise constant over intervals of length $2^0 = 1$, and assumes the values $\ldots, b_{-1}, b_0, b_1, b_2, \ldots$. As shown in Figure 3, we can write $f^{(0)}(t)$ as the sum of two components: A sequence of shifted versions of the Haar wavelet at scale $m = 1$ (i.e., of the functions $\psi_{1,n}(t)$) and a “residual” function $f^{(1)}(t)$, which is piecewise constant over intervals of length $2^1 = 2$.

Specifically, we can write

$$d^{(1)}(t) = \sum_{n=-\infty}^{\infty} \underbrace{\frac{b_{2n} - b_{2n+1}}{\sqrt{2}}}_{a_{1,n}} \psi_{1,n}(t), \tag{32}$$

and

$$f^{(1)}(t) = \sum_{n=-\infty}^{\infty} \underbrace{\frac{b_{2n} + b_{2n+1}}{\sqrt{2}}}_{c_n} \varphi_{1,n}(t), \tag{33}$$

where we use

$$\varphi_{1,n}(t) = \begin{cases} 1, & \text{for } n \le t < n+1, \\ 0, & \text{otherwise.} \end{cases} \tag{34}$$

The real boost now comes from observing that we can just continue along the same lines and decompose $f^{(1)}(t)$ into two parts.

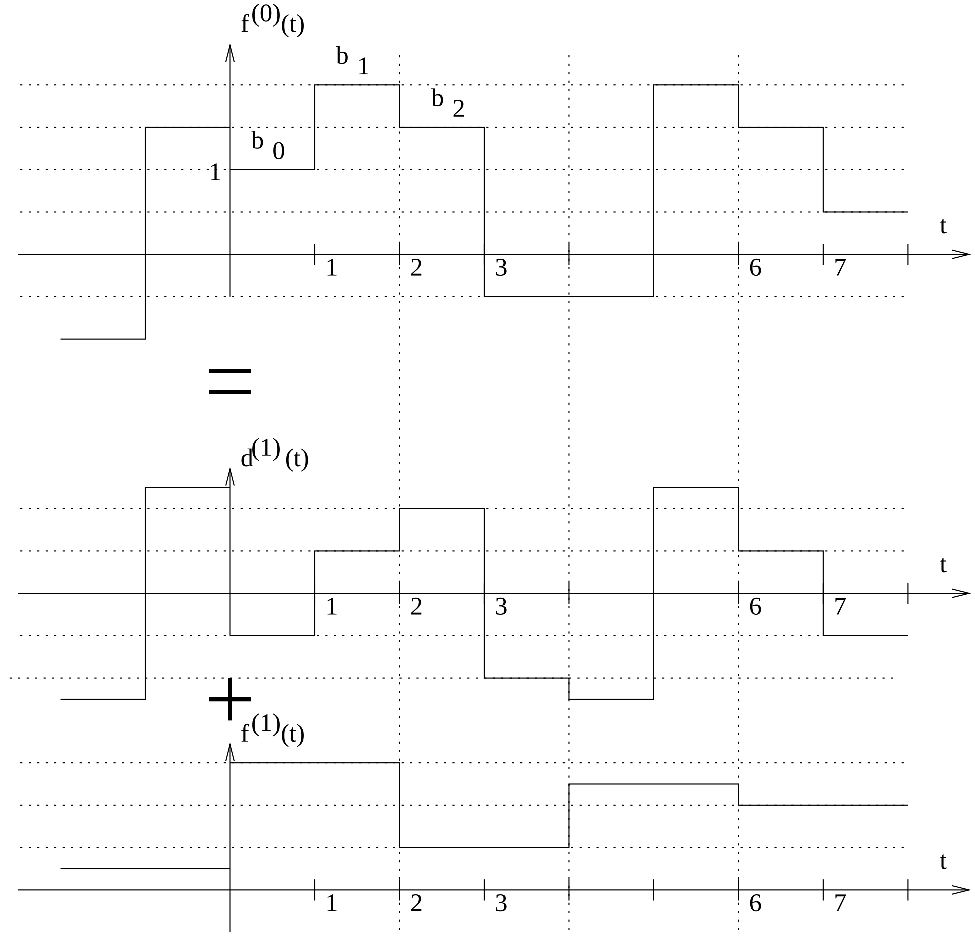

Figure 3: The Haar wavelet at work.

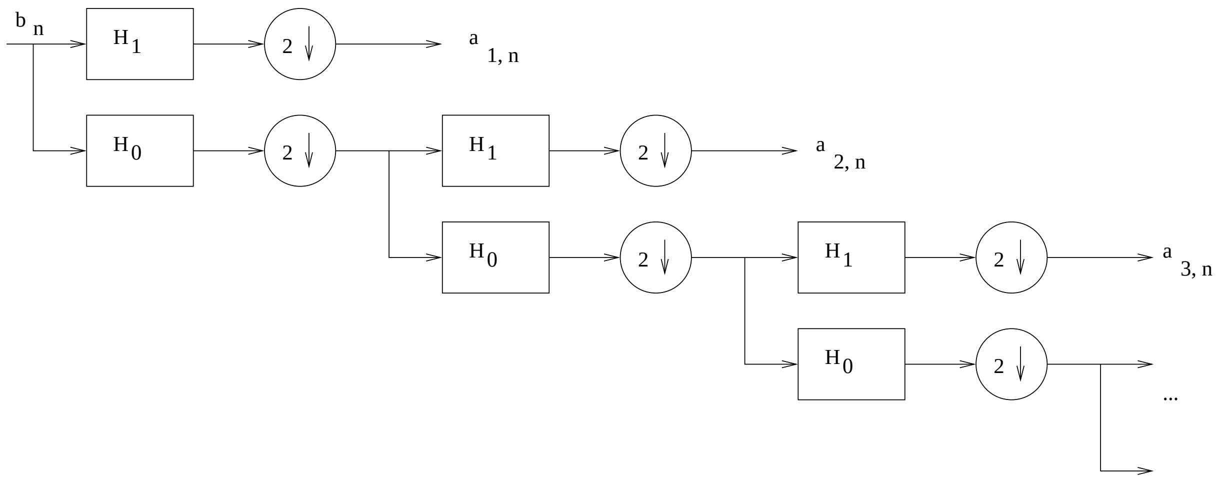

Figure 4: The Haar filter bank.

3.2 A Filter Bank Companion To The Haar Wavelet

Let us now reconsider Equations (32) and (33). We can rewrite them purely in terms of the coefficients as

$$a_{1,n} = \frac{b_{2n} - b_{2n+1}}{\sqrt{2}} \quad (35)$$

$$c_n = \frac{b_{2n} + b_{2n+1}}{\sqrt{2}}, \quad (36)$$

which we easily recognize as filtering the sequence b_n with two different filters and then downsampling by a factor of two! In other words, starting from the coefficients b_n, there is a simple filter bank structure that computes all the wavelet coefficients, illustrated in Figure 4. For the Haar example, the filters are

$$H_1(z) = \frac{1}{\sqrt{2}}(1 - z) \quad (37)$$

$$H_0(z) = \frac{1}{\sqrt{2}}(1 + z). \quad (38)$$

Note that these filters are not quite causal, but since they are FIR, this is not a problem at all.

3.3 Tilings Of The Time-frequency Plane

An interesting observation follows by observing that the Haar filter $H_1(z)$ is (a crude version of) a highpass filter, and $H_0(z)$ a (no less crude version of a) lowpass filter. Nevertheless, if we merely go with this highpass/lowpass picture, and merge it with the downsampling in time, we can draw a figure like the one given in Figure 5: consider for example the wavelet coefficient $a_{1,0}$. It results from highpass-filtering, so it pertains to the upper half of the frequencies. Moreover, it pertains only to the time interval from 0 to 2. This is how we found the rectangle labelled $a_{1,0}$ in Figure 5, and this is how you can find all the other rectangles in the figure.

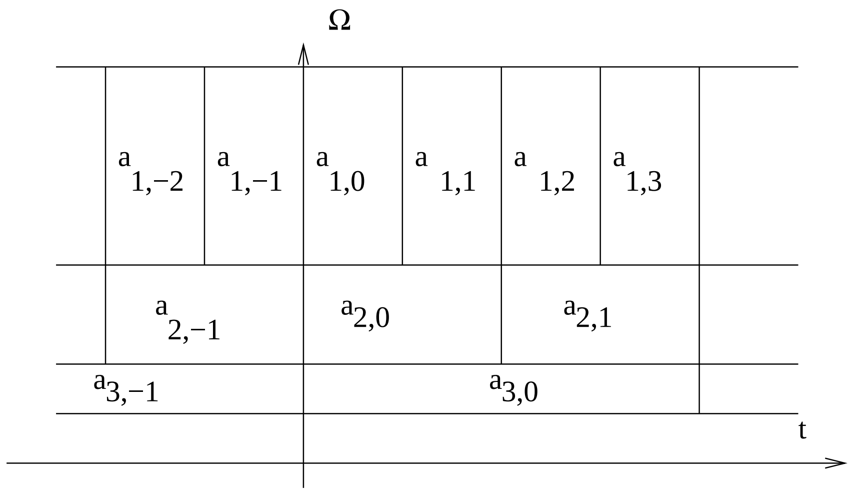

Figure 5: A tiling of the time frequency plane that somewhat mirrors what the Haar wavelet is doing.

References

[1] M. Vetterli and J. Kovacevic, *Wavelets and subband coding.* Upper Saddle River, NJ: Prentice Hall, 1995.

[2] S. Mallat, *A wavelet tour of signal processing.* San Diego, CA: Academic Press, 2nd ed., 1999.