Transform Methods & Signal Processing

lecture 09

Matthew Roughan
<matthew.roughan@adelaide.edu.au>

Discipline of Applied Mathematics
School of Mathematical Sciences
University of Adelaide

September 8, 2010

Wavelets

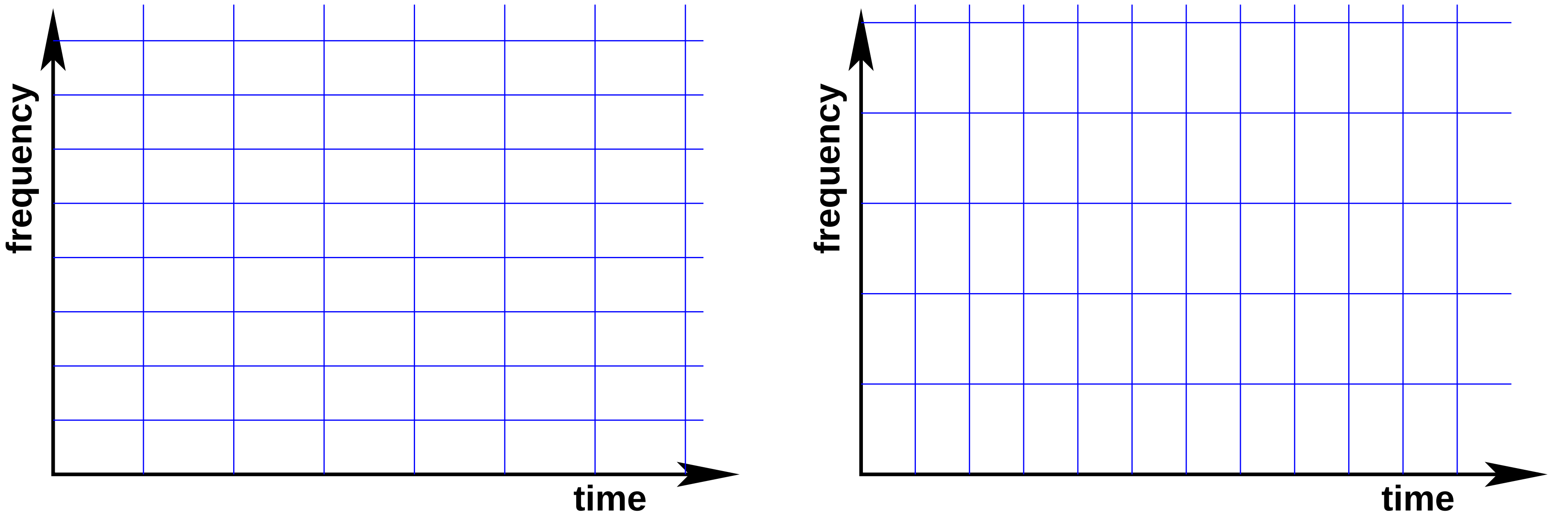
In previous lectures we saw that the STFT had problems. The Wavelet transform is the way to overcome these problems. One of the nicest aspects of wavelets is that they are so natural: they have been invented several times, each time from a different viewpoint, so we will consider several approaches that naturally result in a Wavelet transform, starting by extending our understanding of the uncertainty principle and Windowed Fourier Transforms.

Limitations of the STFT

- computational cost $O(nm \log m)$
- time/frequency resolution tradeoff
  - small m better time, worse frequency resolution
- time/frequency resolution tradeoff is fixed
  - higher freq. can change faster than low freq.
  - appropriate resolution for each frequency?
- how can we do better?
  - some improvement might be gained through using better window functions (I have just used rectangular windows above)
  - lets try to get a more theoretical understanding of windows, and uncertainty bounds

Cutting up the time-frequency space

STFT partition of time-frequency



Areas of boxes don't get smaller!

Scaling property of FT

If we scale a function in time, then

$$\mathcal{F}\{f(at)\} = \frac{1}{a} F\left(\frac{s}{a}\right)$$

- Reciprocal scaling in each domain
- Tighter in Time, makes it looser in Fourier domain
- This contributes to uncertainty!!!!
  - in the STFT we use a window function to restrict the support of basis functions
  - tighter support on window function (less uncertainty in the time domain) results in a wider function in the frequency domain (and so more uncertainty there).

Heisenberg's Uncertainty Principle

Heisenberg's inequality is

$$\Delta x \Delta p \geq \frac{h}{2\pi}$$

where Δx and Δp are the unknown errors in position and momentum, respectively. It arises because, when one measures, say the location of a particle, one must bounce a photon on the particle. The impact of the photon changes the momentum of the particle by an unknown amount. One can reduce the energy of the photon to reduce the range of uncertainty in this change in momentum, but only by reducing the photon's frequency, thereby reducing the accuracy of the localization gained through the measurement.

Uncertainty Principle

Given a transient signal $f(t)$, we want to localize this signal in time and frequency. We measure mean location of transient time and frequency by

$$u = \frac{1}{\|f\|^2}\int_{-\infty}^{\infty} t|f(t)|^2\,dt$$

$$\xi = \frac{1}{\|f\|^2}\int_{-\infty}^{\infty} s|F(s)|^2\,ds$$

Measure uncertainties in time and frequency by variance about the mean, e.g.

$$\sigma_t^2 = \frac{1}{\|f\|^2}\int_{-\infty}^{\infty} (t-u)^2|f(t)|^2\,dt$$

$$\sigma_s^2 = \frac{1}{\|f\|^2}\int_{-\infty}^{\infty} (s-\xi)^2|F(s)|^2\,ds$$

Uncertainty Principle

Theorem: For a function $f \in L^2$, the temporal and frequency variance satisfy

$$\sigma_t \sigma_s \geq \frac{1}{4\pi}$$

And this is an equality only if there exist $(u, \xi, a, b) \in \mathbb{R}^2 \times \mathbb{C}^2$ such that

$$f(t) = a e^{-b(t-u)^2} e^{i2\pi\xi t}$$

for which

$$\begin{aligned} \sigma_t^2 &= \frac{1}{4\pi b^2} \\ \sigma_s^2 &= \frac{b^2}{4\pi} \end{aligned}$$

Uncertainty Principle

Proof: It is sufficient to prove the theorem for f such that $u = \xi = 0$ as we can always perform shifts in time and frequency, e.g. by taking $\exp(i2\pi\xi t) f(t-u)$, to get the general case. In the case $u = \xi = 0$ we get

$$\sigma_t^2 \sigma_s^2 = \frac{1}{\|f\|^4} \int_{-\infty}^{\infty} t^2 |f(t)|^2 \, dt \int_{-\infty}^{\infty} s^2 |F(s)|^2 \, ds$$

Remember $\mathcal{F}\left\{\frac{df}{dt}\right\} = (i2\pi s)F(s)$, so Rayleigh's theorem implies

$$\int_{-\infty}^{\infty} |i2\pi s F(s)|^2 \, ds = 4\pi^2 \int_{-\infty}^{\infty} |f'(t)|^2 \, dt$$

Uncertainty Principle

Proof: Hence we can write

$$\sigma_t^2 \sigma_s^2 = \frac{1}{4\pi^2 \| f \|^4} \int_{-\infty}^{\infty} t^2 |f(t)|^2 \, dt \int_{-\infty}^{\infty} |f'(t)|^2 \, dt$$

Schwarz's inequality (for real functions)

$$\int_a^b \psi_1(x)^2 dx \int_a^b \psi_2(x)^2 dx \geq \left[\int_a^b \psi_1(x) \psi_2(x) \, dx \right]^2$$

with equality only if $\psi_2(x) = \alpha \psi_1(x)$ for some constant α.

$$\sigma_t^2 \sigma_s^2 \geq \frac{1}{4\pi^2 \| f \|^4} \left[\int_{-\infty}^{\infty} t \, f'(t) \, f(t) \, dt \right]^2$$

Uncertainty Principle

Proof: When $\psi_1(x)$ and $\psi_1(x)$ are complex, a more appropriate form of Schwarz's inequality is (from Bracewell, p.176) gives

$$4\int_a^b |\psi_1(x)|^2 dx \int_a^b |\psi_2(x)|^2 dx \geq \left[\int_a^b \left(\psi_1^*(x)\psi_2(x) + \psi_1(x)\psi_2^*(x)\right) dx\right]^2$$

So

$$\begin{aligned}
\sigma_t^2 \sigma_s^2 &\geq \frac{1}{16\pi^2 \|f\|^4} \left[\int_{-\infty}^{\infty} t \left(f'(t) f^*(t) + f^{*\prime}(t) f(t)\right) dt\right]^2 \\
&\geq \frac{1}{16\pi^2 \|f\|^4} \left[\int_{-\infty}^{\infty} t \frac{d}{dt} \left(f(t) f^*(t)\right) dt\right]^2 \\
&\geq \frac{1}{16\pi^2 \|f\|^4} \left[\left[t\,|f(t)|^2\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} |f(t)|^2 dt\right]^2
\end{aligned}$$

Uncertainty Principle

Proof: The Theorem holds for all $f \in L^2(\mathbb{R})$, but we are mainly interested in **transient** signals

- **transient** signals go to zero at some point
- lets have a fairly weak definition $\lim_{|t|\to\infty} \sqrt{t}\, f(t) = 0$
- in this case, the first term in the integration by parts is zero, so

$$
\begin{aligned}
\sigma_t^2 \sigma_s^2 &\geq \frac{1}{16\pi^2 \|f\|^4} \left[\int_{-\infty}^{\infty} |f(t)|^2 \, dt \right]^2 \\
&\geq \frac{1}{16\pi^2 \|f\|^4} \left[\|f\|^2 \right]^2 \\
&\geq \frac{1}{16\pi^2}
\end{aligned}
$$

Uncertainty Principle

Proof: To obtain an equality, note that Schwarz's inequality requires $\psi_2(x) = \alpha\psi_1(x)$ for some constant α, which in this case implies that

$$f'(t) = -2bt\,f(t)$$

which is true only for

$$f(t) = ae^{-bt^2}$$

This is the result for $(u,\xi) = (0,0)$. We perform a frequency and time translation to freq. ξ and time u to get

$$f(t) = ae^{-b(t-u)^2}e^{i2\pi\xi t}$$

$\square$

Gabor function

Definition: A Gabor function

$$f_{a,b,u,\xi}(t) = ae^{-b\pi(t-u)^2}e^{i2\pi\xi t}$$

It has FT

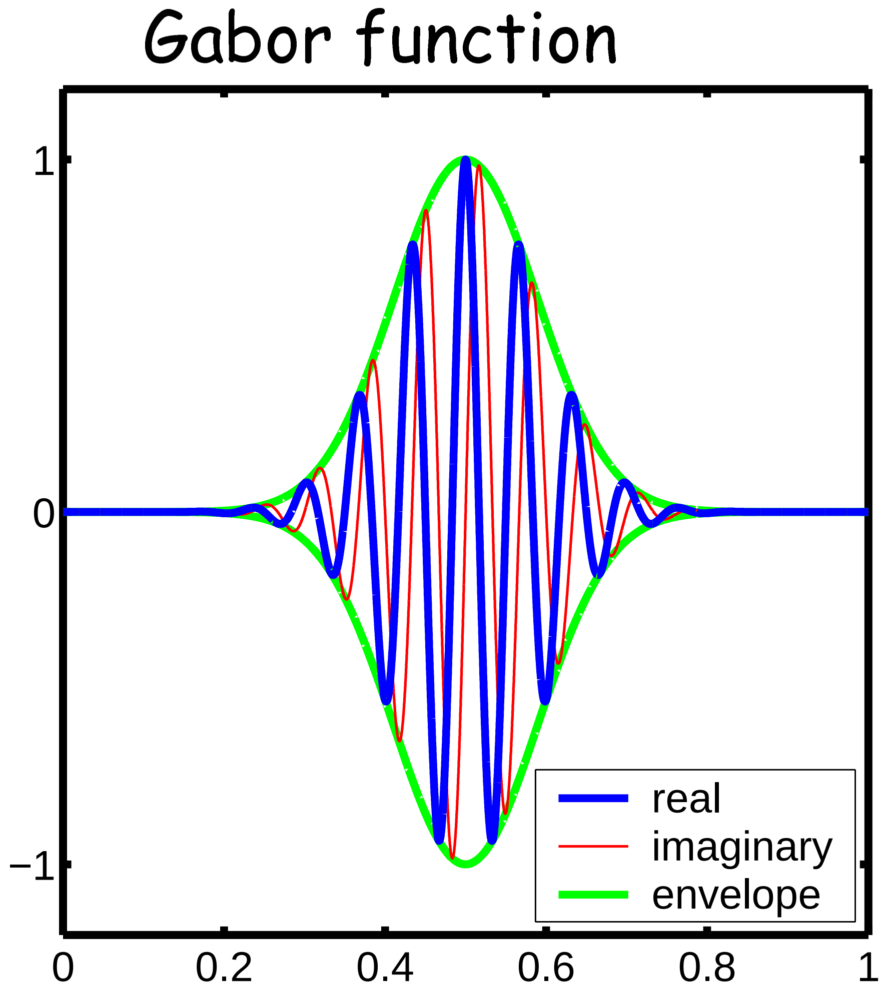
$$F_{a,b,u,\xi}(s) = \frac{a}{\sqrt{b}}e^{-\pi(s-\xi)^2/b}e^{-i2\pi su}$$

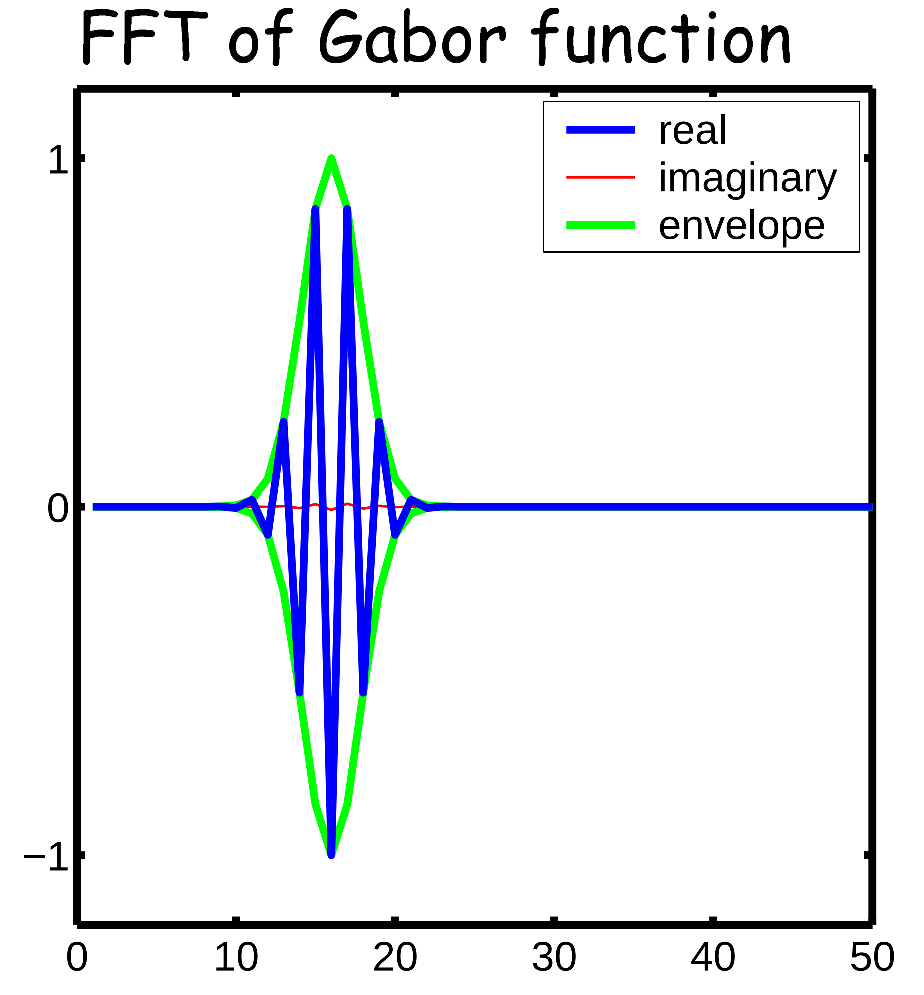
Mean position and frequency are u and ξ, and the uncertainty in location is

$$\sigma_t^2 = \frac{1}{\|f\|^2}\int_{-\infty}^{\infty}(t-u)^2|f(t)|^2\,dt = \frac{1}{b}\int_{-\infty}^{\infty}t^2e^{-2b\pi t^2}\,dt = \frac{1}{4\pi b^2}$$

$$\sigma_s^2 = \frac{1}{\|f\|^2}\int_{-\infty}^{\infty}(s-\xi)^2|F(s)|^2\,ds = \frac{1}{b}\int_{-\infty}^{\infty}t^2e^{-2b\pi t^2}\,dt = \frac{b^2}{4\pi}$$

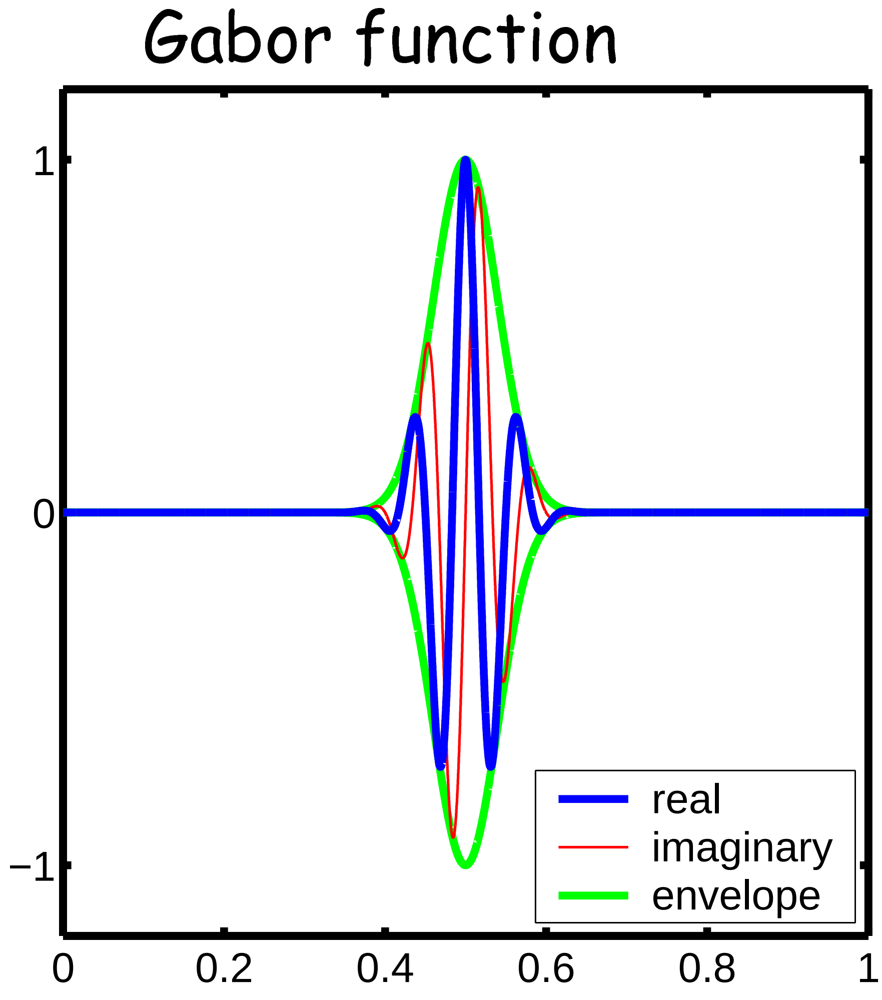
Gabor function

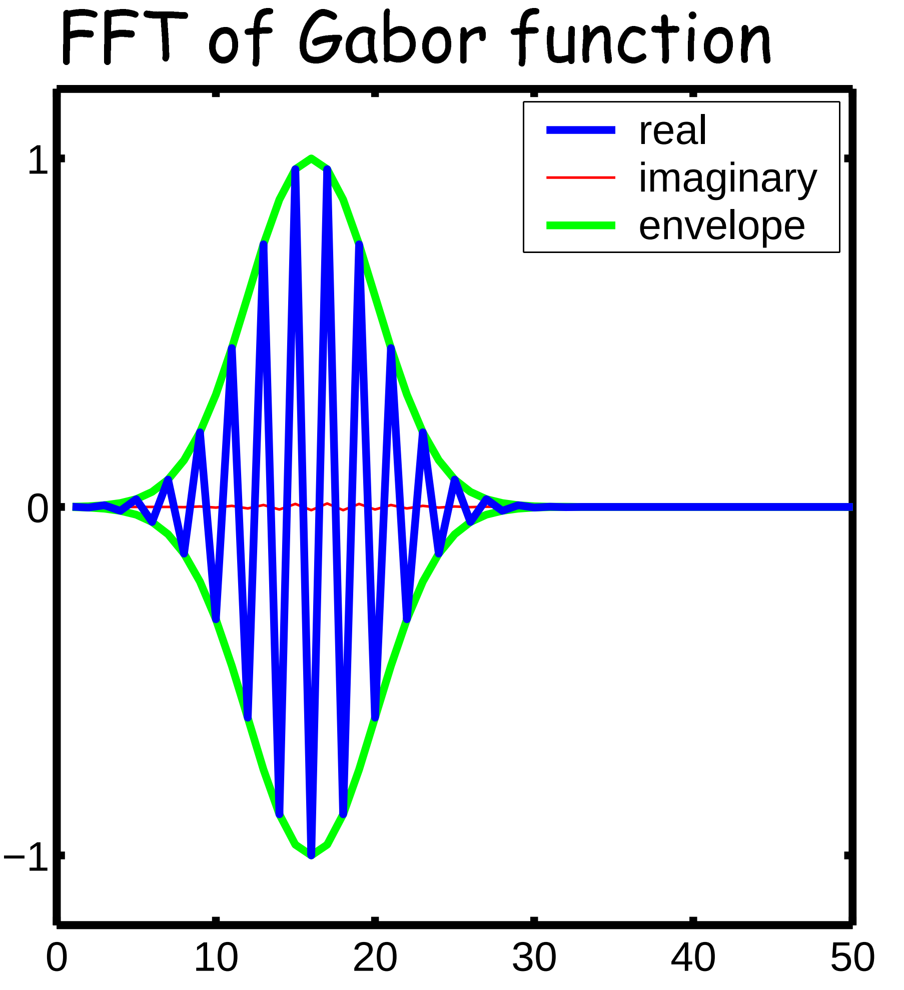




Gabor function
= Gaussian window applied to a complex sinusoid

Gabor function

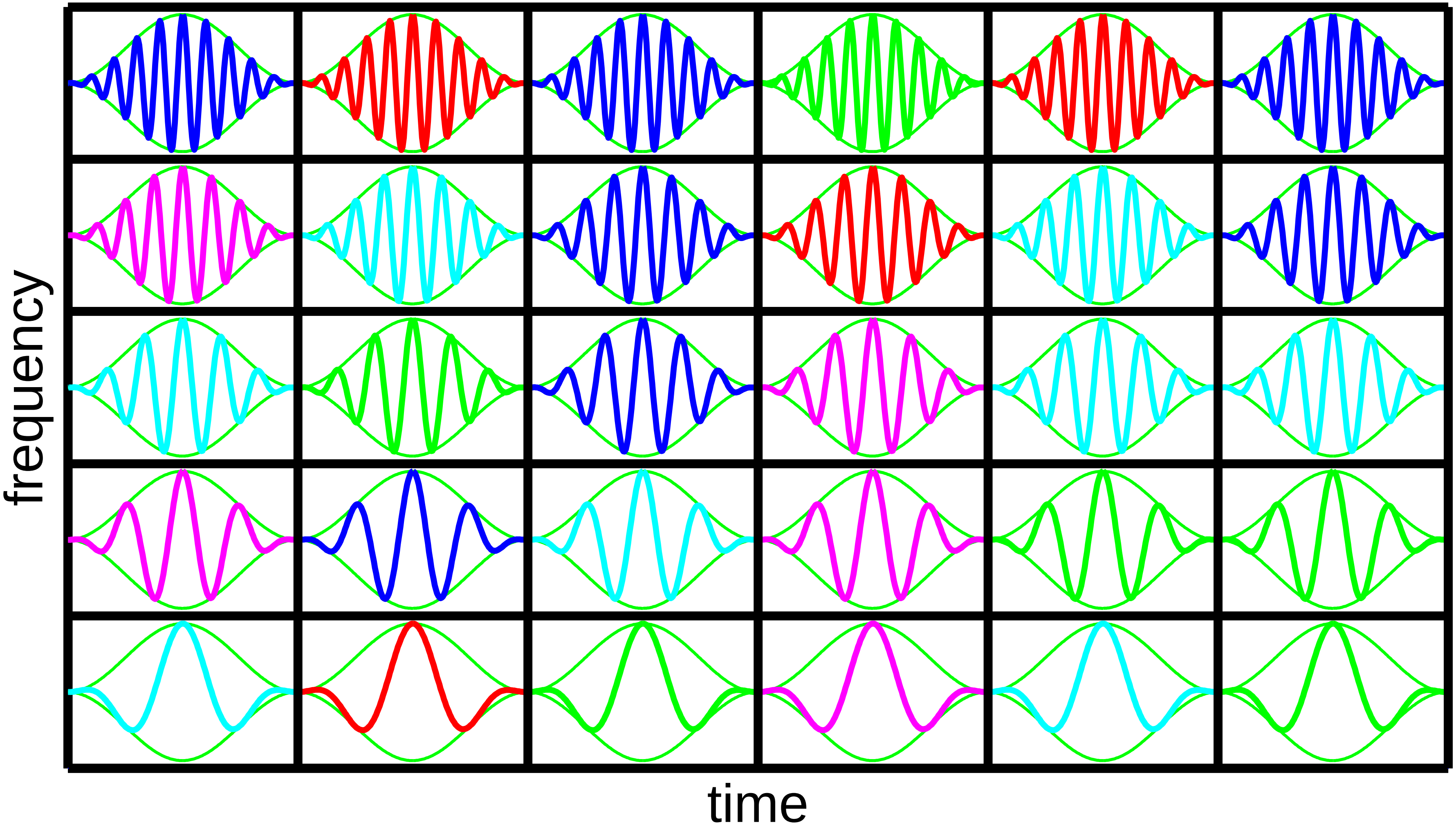




If we make the time-domain function narrower
the frequency domain function gets wider

Cutting up the time-frequency space

Basis-like functions for a STFT with a window function



An alternative

Remember that we can scale window functions to change the resolution in time and frequency.

- higher frequencies can change more quickly
- why not change frequency resolution to match the frequency?
- just have to make the window width a function of frequency
- e.g. for the Gabor functions $f(t) = ae^{-b\pi(t-u)^2} e^{i2\pi\xi t}$ make the window frequency dependent by making b a function of ξ
  - higher frequencies make the window narrower
  - so for larger ξ we want smaller b.

Wavelets

Wavelet are the natural result of this idea.

- start with a function we call the **Mother Wavelet**
  - e.g. a rectangular pulse, or a Gabor function
  - denote by $\psi(t)$
  - require $\psi \in L^2$, $\|\psi\| = 1$ and $\int_{-\infty}^{\infty} \psi(t)\, dt = 0$
- construct a set of atomic functions $\psi_{u,s}$ (atoms) from this function by
  - dilation (stretching and shrinking by s)
  - translation (shifting in time by u)

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right)$$

- e.g. could generate any Gabor function this way

Definition: Atoms

Time-frequency **atoms** $\{\phi_\gamma\}$, underly many transforms

- $\phi_\gamma \in L^2$
- $\|\phi_\gamma\| = 1$
- Transform $F(\gamma) = \langle f(t), \phi_\gamma(t) \rangle$

For example the STFT

$$\phi_\gamma(t) = g_{\xi,u}(t) = e^{-i2\pi\xi t} g(t-u)$$

where $g(t)$ is the (suitably normalized) window function.

Continuous wavelet transform

Wavelet Transform (analysis)

$$\mathcal{W}\{f(t)\} = W_f(u,s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}} \psi^* \left(\frac{t-u}{s} \right) dt$$

Wavelet Reconstruction (synthesis), choose a complete, orthogonal set of wavelets $\{\psi_{j,n}\}$, then

$$f = \sum_j \sum_n \langle f, \psi_{n,j} \rangle \psi_{n,j}$$

Similar to the generalized Fourier transform.

Wavelets

There are many possible Mother Wavelets

- Haar
- Daubechies
- Mexican hat
- Gabor
- ...

Each has slightly different properties – much the same as when we considered window functions.

Wavelet Example

Mexican Hat wavelets are given by the second derivative of a Gaussian function, e.g.

$$\psi(t) = \frac{2}{\pi^{1/4}\sqrt{3\sigma}}\left(\frac{t^2}{\sigma^2} - 1\right)\exp\left(\frac{-t^2}{2\sigma^2}\right)$$

Its FT is

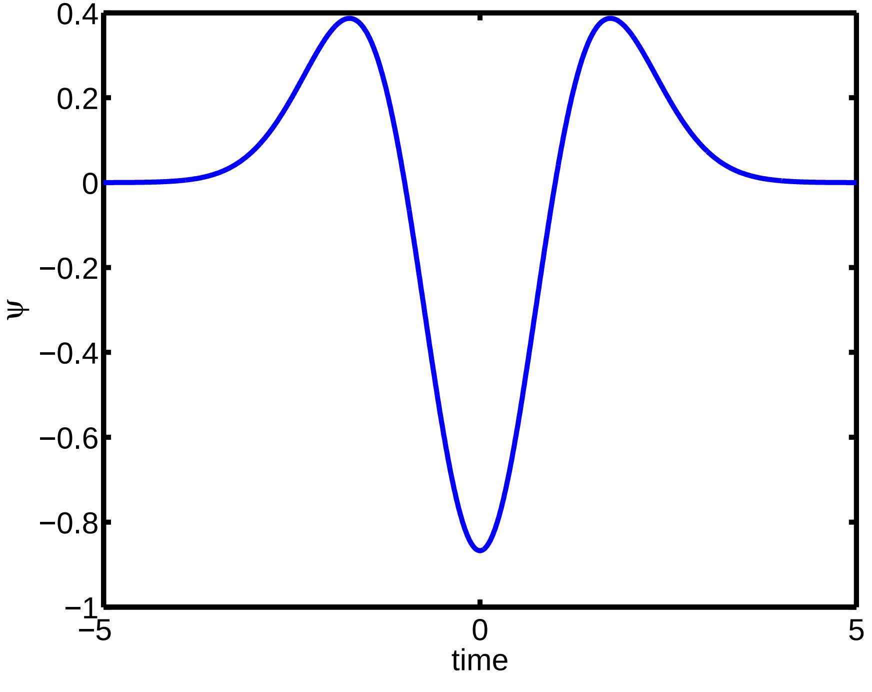
$$\Psi(\omega) = \frac{-\sqrt{8}\sigma^{5/2}\pi^{1/4}}{\sqrt{3}}\omega^2 \exp\left(\frac{-\sigma^2\omega^2}{2}\right)$$

where $\omega = 2\pi s$ is frequency in radians per time unit.

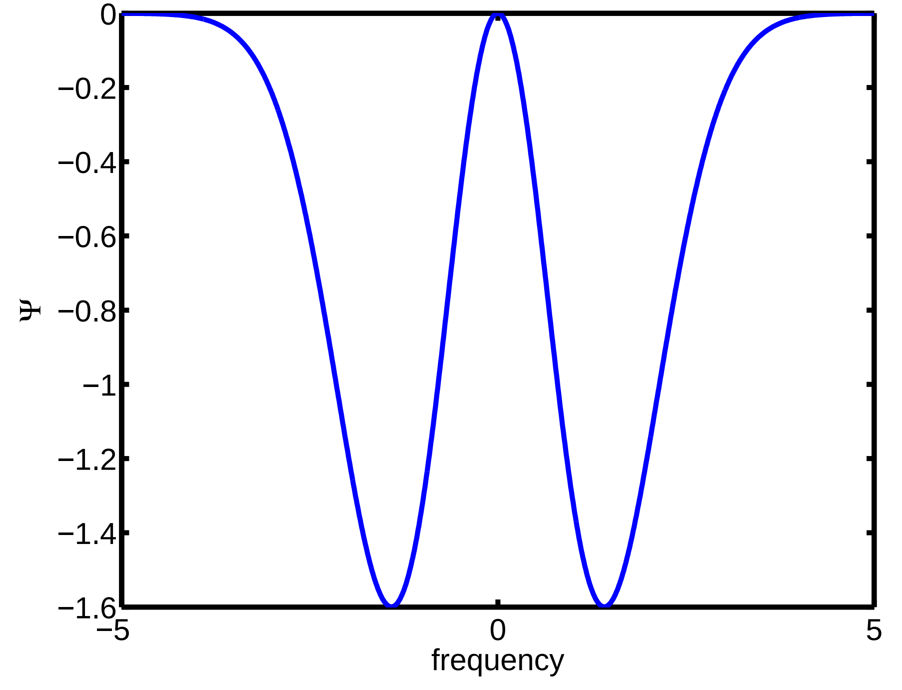
Wavelet Example

Mexican Hat wavelets

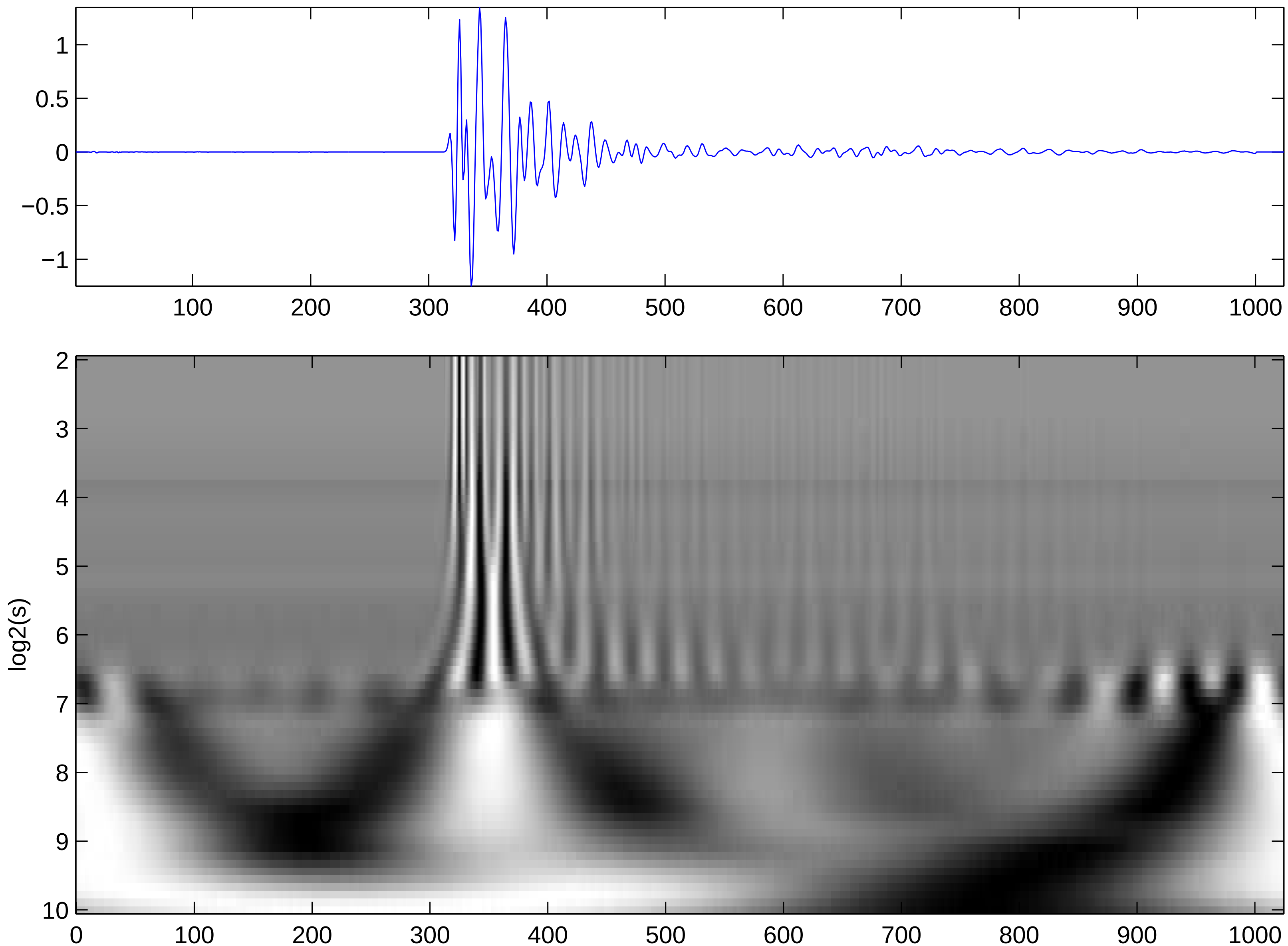
wavelet



FT of the wavelet



Example wavelet transform



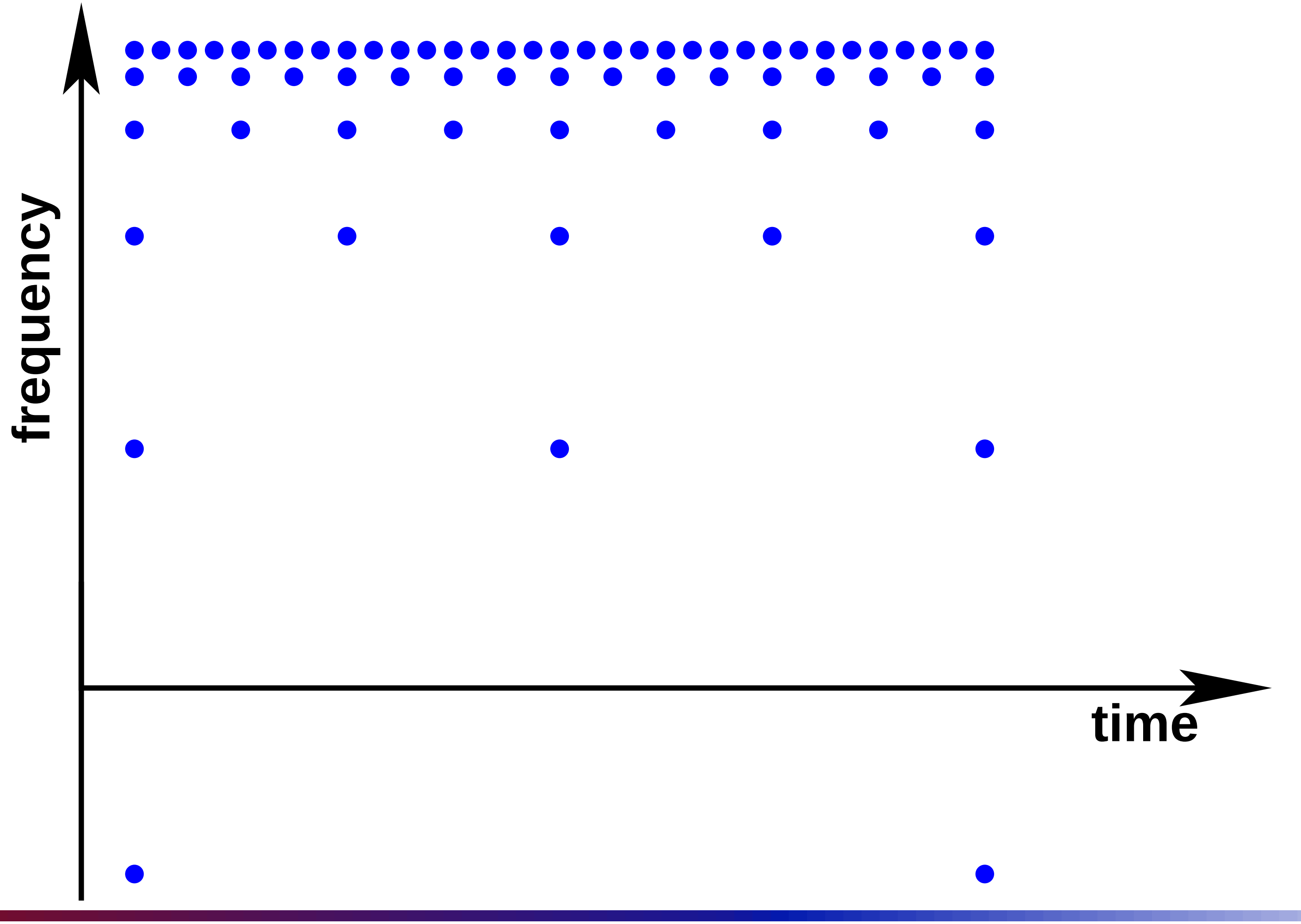
Wavelet Basis

We don't need to consider all possible wavelet translations and dilations:

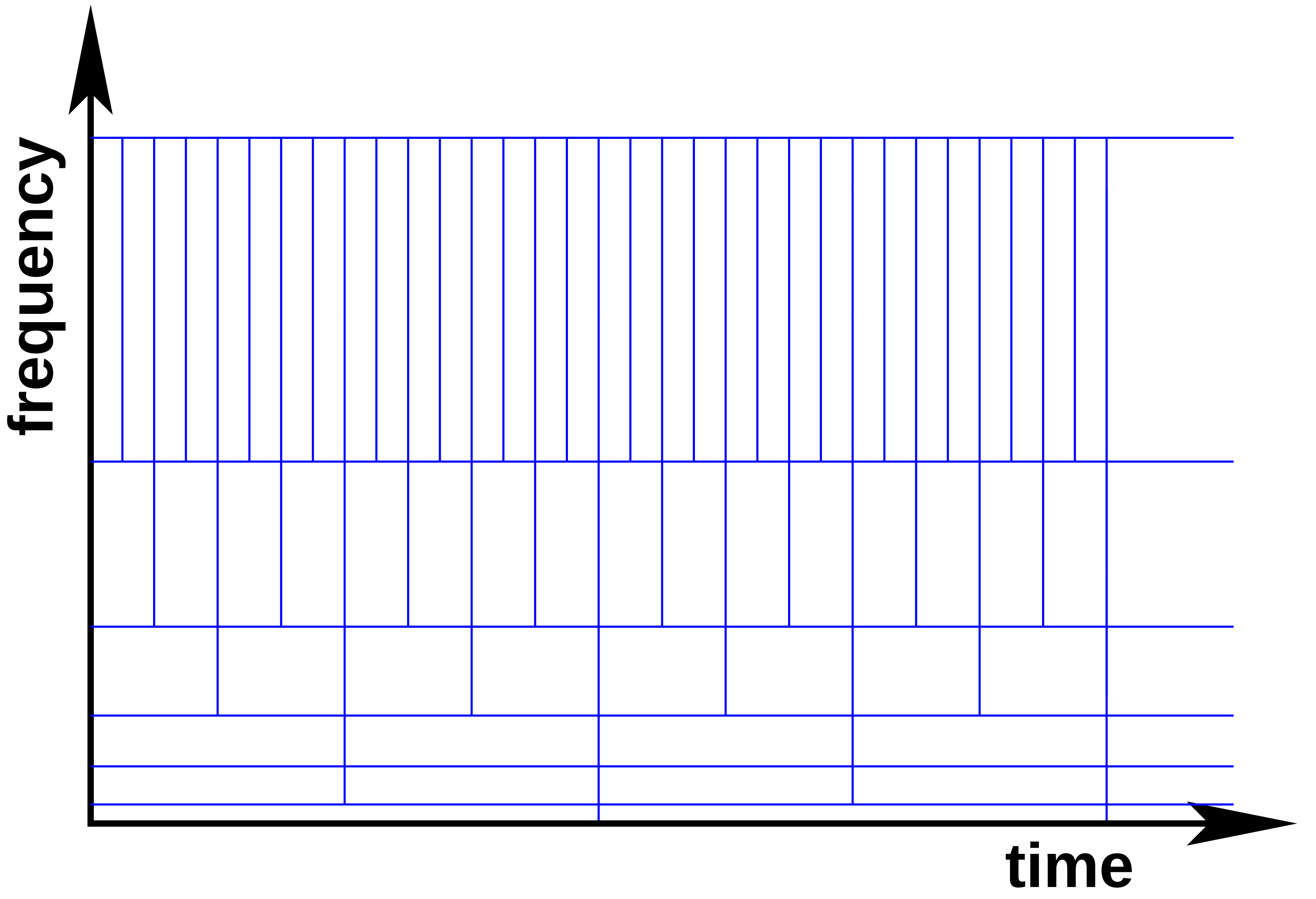
- We can think of the wavelet transform as a generalized FT
- So we want to find an orthogonal basis
- Also want time resolution tuned to frequency
- Choose a set of wavelets such that we get this
- Choose points on the **dyadic grid**

Dyadic grid

Higher frequencies change more rapidly than low frequencies and so need to be sampled at a higher rate.

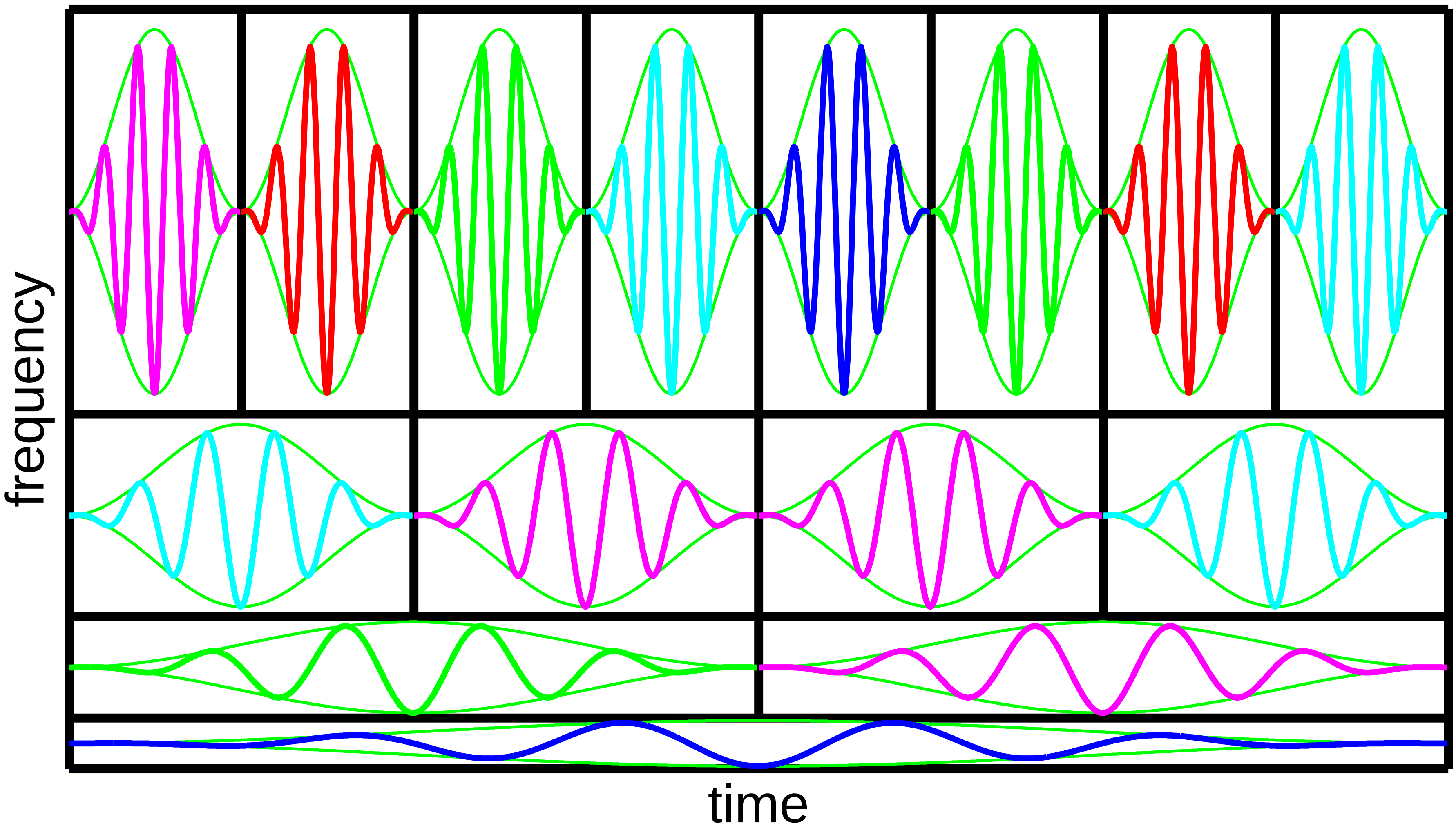


Wavelet Partition



Cutting up the time-frequency space

Basis functions for a wavelet(-like) transform



Wavelet transforms

- Continuous Wavelet Transform (CWT) is the transform onto the whole space (u, s).
- Discrete Wavelet Transform (DWT) is the continuous transform, onto the discrete space given by the dyadic grid.
  - wavelet basis on dyadic grid defined by

$$
\begin{aligned}
s &= 2^j \\
u &= 2^j n
\end{aligned}
$$

    where n and j are integers. So we get the basis

$$
\psi_{n,j}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t}{2^j} - n \right)
$$

Scalogram

Take the power in each wavelet coefficient, e.g.

$$|W_f(u,s)|^2$$

and call this the **scalogram**

- analogous to periodogram (power of Fourier transform)
- analogous to spectrogram (power of STFT)

Time-Frequency Measurement

We can perform transform in either time or frequency domain

$$\mathcal{W}\{f\} = W_f(u,s) = \int_{-\infty}^{\infty} f(t)\psi_{u,s}^*(t)\,dt = \int_{-\infty}^{\infty} F(r)\Psi_{u,s}^*(r)\,dr$$

where $\Psi_{u,s}^*(r) = \mathcal{F}\left\{\psi_{u,s}^*(t)\right\}$

Note that

$$\Psi_{u,s}(r) = e^{-i2\pi ur}\sqrt{s}\Psi(sr)$$

using the scaling and time-translation properties.

Time-Frequency resolution

Time-frequency resolution of a wavelet

$$\mathcal{W}\{f(t)\} = W_f(u,s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}} \psi^* \left(\frac{t-u}{s} \right) dt$$

Suppose WLOG that ψ is centered at 0, which implies $\psi_{u,s}$ is centered at u, then

$$\int_{-\infty}^{\infty} (t-u)^2 |\psi_{u,s}|^2 \, dt = \int_{-\infty}^{\infty} t^2 |\psi_{0,s}|^2 \, dt = s^2 \int_{-\infty}^{\infty} t^2 |\psi(t)|^2 \, dt = s^2 \sigma_t^2$$

So the energy spread of a wavelet atom $\psi_{u,s}$ is a "box" $s\sigma_t$ wide in time.

- σ_t depends on the particular mother wavelet

Time-Frequency resolution

The FT of a wavelet is

$$\Psi_{u,s}(r) = e^{-i2\pi ur}\sqrt{s}\Psi(sr)$$

The center frequency is therefore η_ψ/s, where η_ψ is the center frequency of the mother wavelet.

- hence we call s the **scale**, and note that is it proportional to one over the frequency.
- the center frequency of the mother wavelet is given by

$$\eta_\psi = \int_{-\infty}^{\infty} \omega |\Psi(\omega)|^2 \, d\omega$$

Time-Frequency resolution

The energy spread of the wavelet about the central frequency η_ψ/s is

$$\frac{1}{2\pi}\int_0^\infty \left(\omega - \frac{\eta}{s}\right)^2 |\Psi_{u,s}(\omega)|\, d\omega = \frac{\sigma_\omega^2}{s^2}$$

where

$$\sigma_\omega^2 = \frac{1}{2\pi}\int_0^\infty (\omega-\eta)^2 |\Psi(\omega)|\, d\omega$$

So the energy spread of a wavelet atom $\psi_{u,s}$ is a "box"

- $s\sigma_t$ wide in time (wider for lower frequencies)
- σ_ω/s in frequency (finer for lower freq.)

MultiResolution Approximation and Wavelets

Wavelets were independently invented from several different viewpoints. In this section we start by considering how we can approximate functions at different levels of detail, and by doing so come up again with the notion of wavelets.

MultiResolution Analysis

- a noted, we call s scale
- time-resolution at a particular scale s is fixed
- at different scales, the time resolution is proportional to the scale
- like observing the data at multiple scales
- hence the name **multiresolution analysis**
  - we can take this concept further by considering multiresolution approximation

Approximation

Definition: An approximation of a function $f \in L^2$ in subspace $\mathbf{V}$ is defined as the orthogonal projection of f onto $\mathbf{V}$ (e.g. the projection $\hat{f} \in \mathbf{V}$ that minimizes $\|f - \hat{f}\|$).

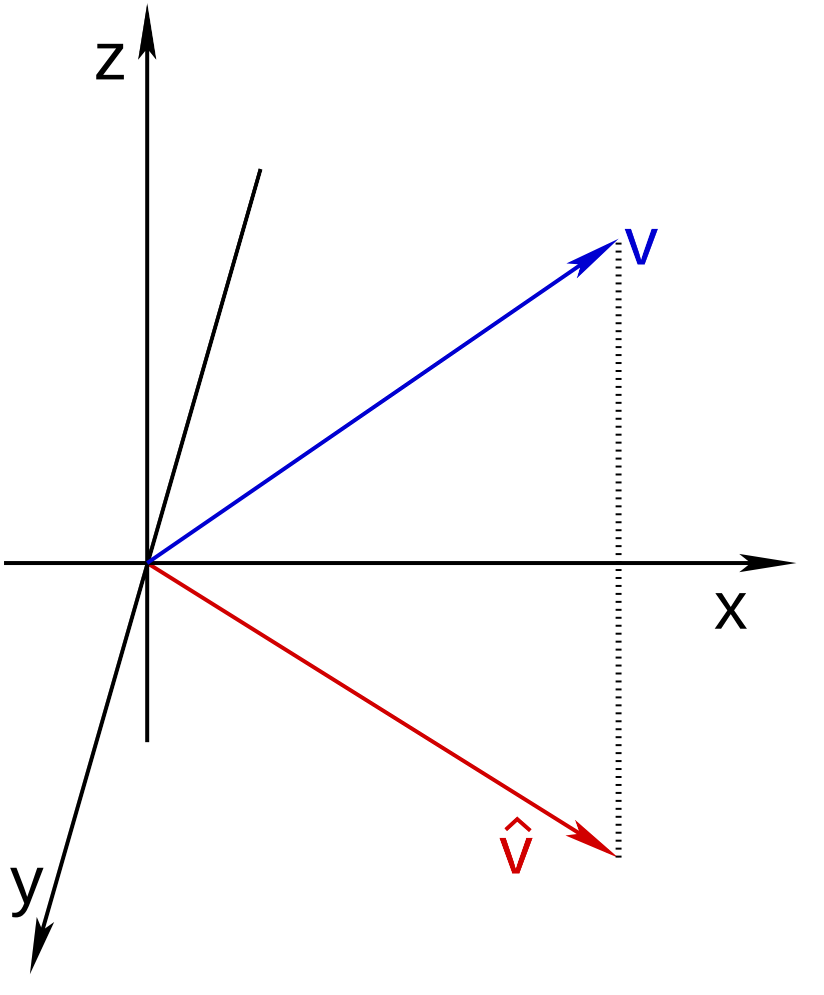
If an orthonormal basis $\{\phi_\gamma\}$ for $\mathbf{V}$ exists, then the projection into the space is given by

$$\hat{f} = \sum_\gamma \langle f, \phi_\gamma \rangle \phi_\gamma$$

Projection

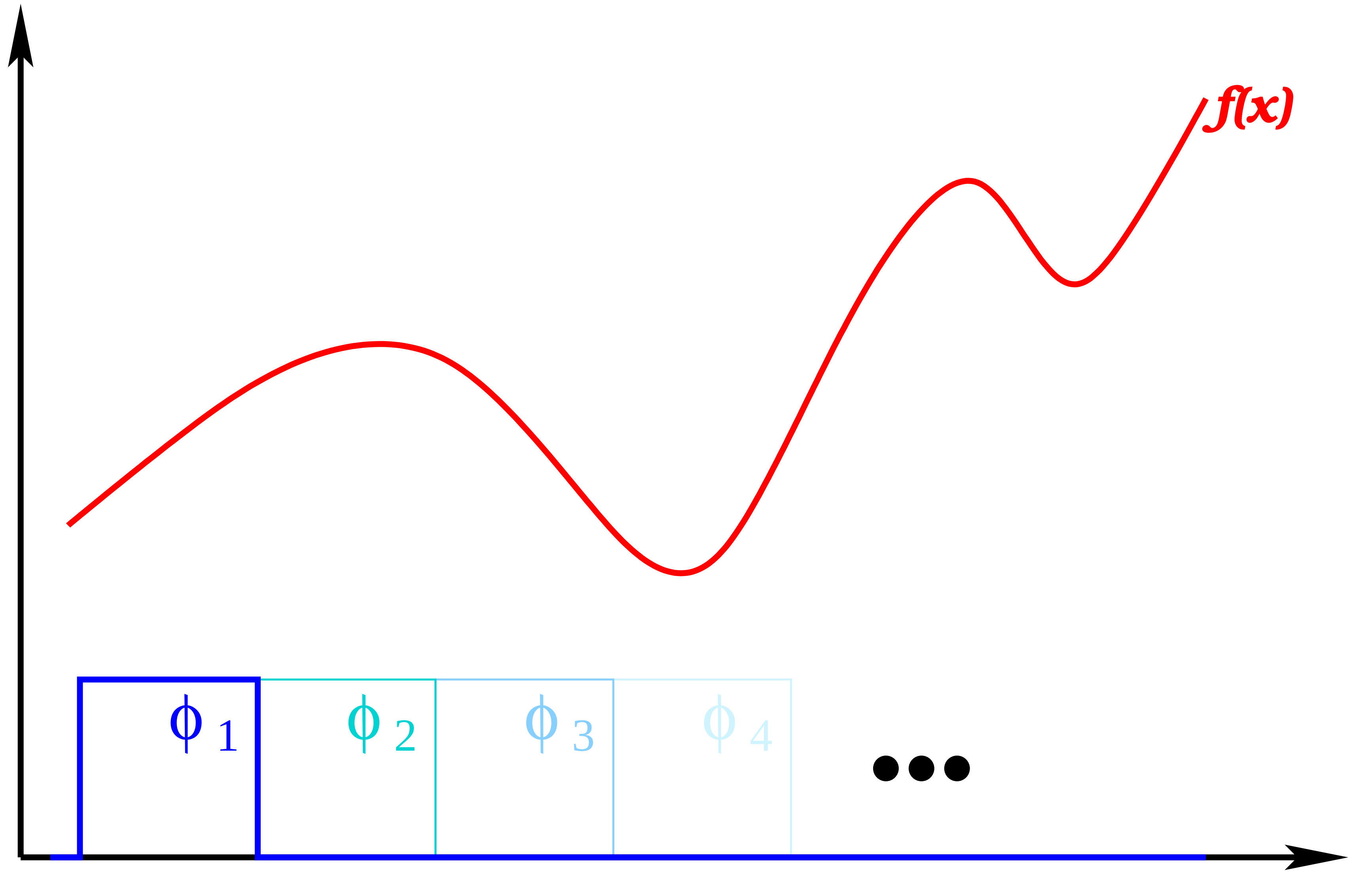
Simple example of projection:

- projecting an (x, y, z) vector into the $x - y$ plane.
- vector $v \in \mathbb{R}^3$ is projected to $\hat{v} \in \mathbb{R}^2$
- take $(1,0,0)$ and $(0,1,0)$ as the basis vectors of the $x - y$ plane.
- inner product is just vector dot product

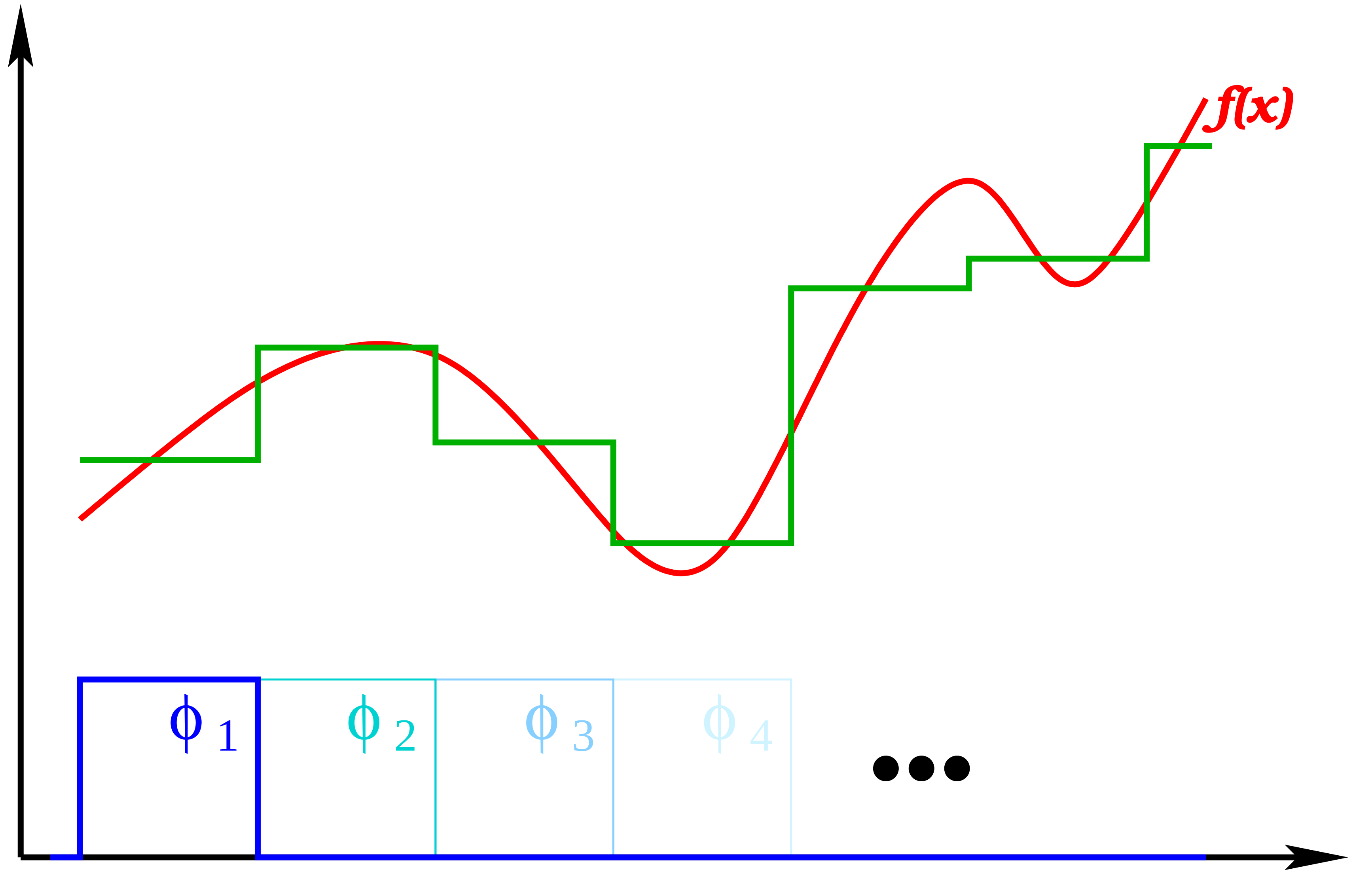


$$
\begin{aligned}
\hat{v} &= [v.(1,0,0)](1,0,0) + [v.(0,1,0)](0,1,0) \\
&= (v_1, v_2, 0)
\end{aligned}
$$

Approximation



Approximation



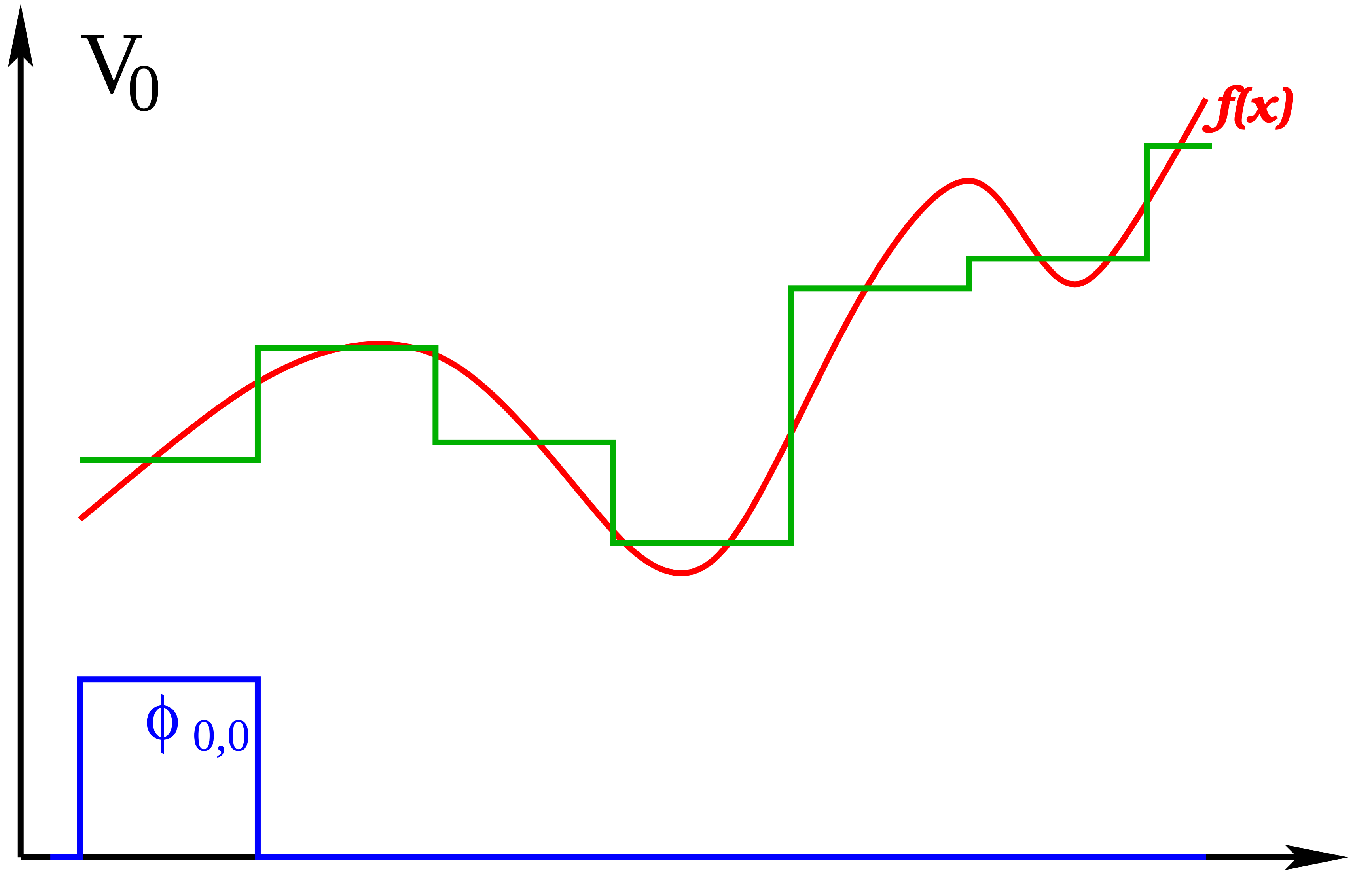
MultiResolution Approximation (MRA)

A sequence $\{\mathbf{V}_j\}_{j\in\mathbb{Z}}$ of closed subspaces of $L^2(\mathbb{R})$ is called a MultiResolution Approximation (MRA) if

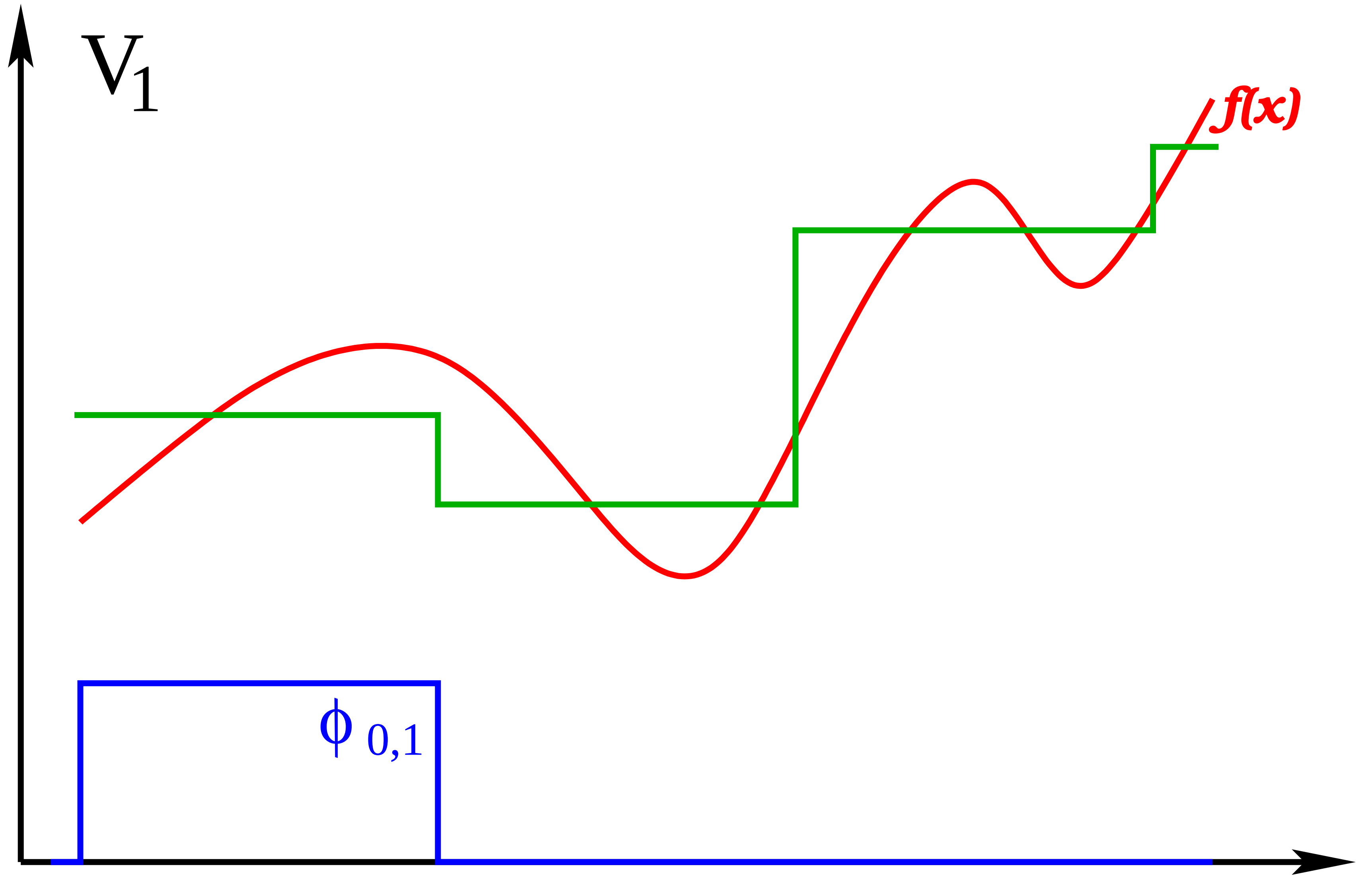
1. $\mathbf{V}_{j+1} \subset \mathbf{V}_j$ for all $j \in \mathbb{Z}$
2. $f(t) \in \mathbf{V}_j \Leftrightarrow f(t - 2^j k) \in \mathbf{V}_j$ for all $j, k \in \mathbb{Z}$
3. $f(t) \in \mathbf{V}_j \Leftrightarrow f(t/2) \in \mathbf{V}_{j+1}$ for all $j, k \in \mathbb{Z}$
4. $\lim_{j\to\infty} \mathbf{V}_j = \{0\}$
5. $\lim_{j\to-\infty} \mathbf{V}_j = L^2(\mathbb{R})$
6. $\exists\theta$ such that $\{\theta(t-n)\}_{n\in\mathbb{Z}}$ is a Riesz basis of $\mathbf{V}_0$.

We can think of $\mathbf{V}_j$ grouping together the approximations at **scale** 2^j. Sometimes call j the **octave** (through analogy to music).

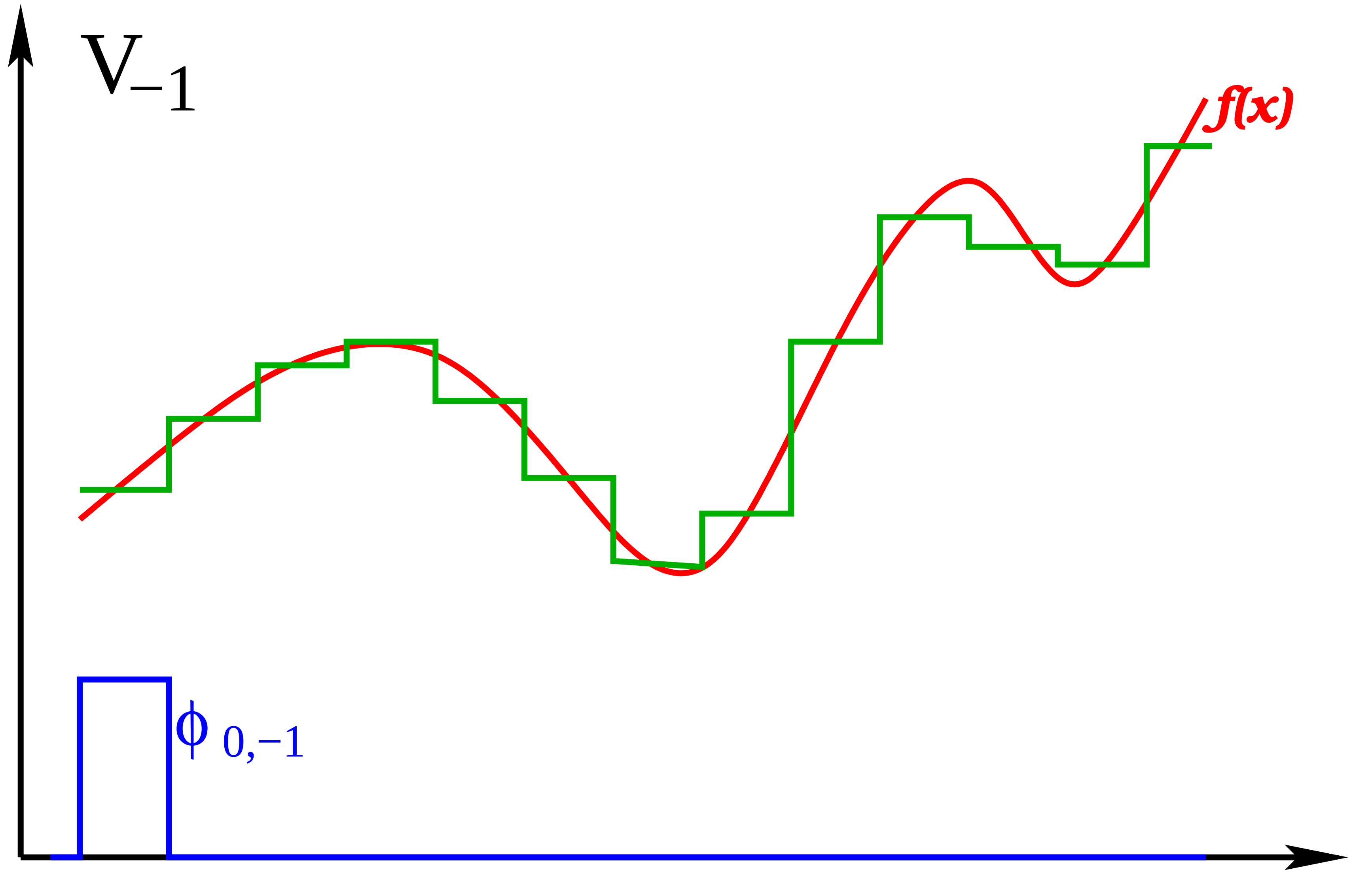
MRA example



MRA example



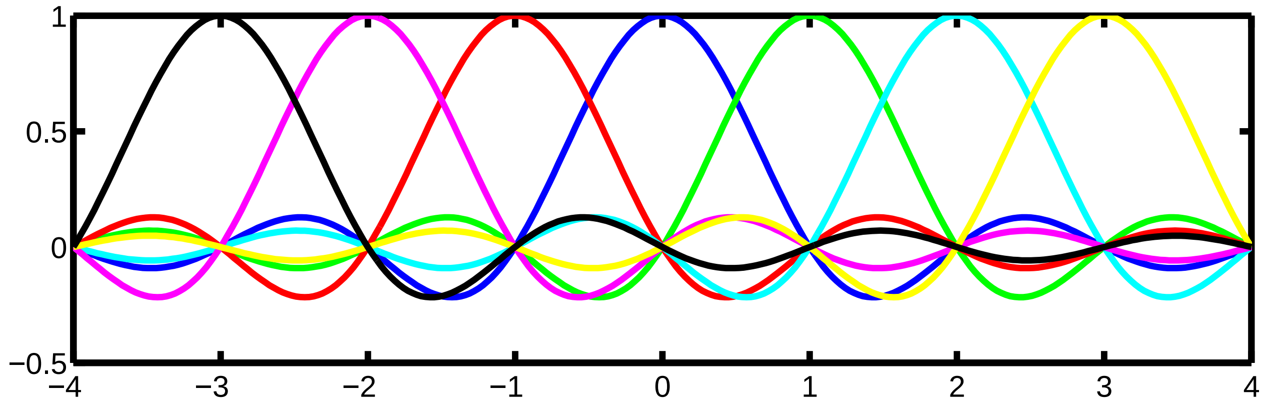
MRA example



MRA examples

Examples:

- **piecewise constant:** see above.
- **Shannon approximation:** using frequency band-limited functions (which hence must have infinite support in the time domain). Orthonormal basis $\mathrm{sinc}(t-n)$.



- **Spline approximation:**

MRA and scaling functions

From the Riesz basis $\exists\theta$ for the MRA, we can derive an orthonormal basis $\{\phi_{n,j}(t)\}_{n\in\mathbb{Z}}$ for $\mathbf{V}_j$. The functions ϕ are called **scaling functions**, and can be derived from a mother scaling function as with wavelets, e.g.

$$\phi_{n,j}(t) = \frac{1}{\sqrt{2^j}}\phi\left(\frac{t}{2^j} - n\right)$$

The approximation of a function $f \in L^2(\mathbb{R})$ is given by

$$\hat{f}_j(t) = \sum_{n\in\mathbb{Z}} \langle f, \phi_{n,j}\rangle \phi_{n,j}(t)$$

where

$$\langle f, \phi_{n,j}\rangle = \int_{-\infty}^{\infty} f(t)\phi_{n,j}(t)\,dt = \int_{-\infty}^{\infty} f(t)\frac{1}{\sqrt{2^j}}\phi\left(\frac{t}{2^j} - n\right)dt = \left[f * \bar{\phi}_j\right](n)$$

The Approximation

The approximation of a function $f \in L^2(\mathbb{R})$ is given by

$$\hat{f}_j(t) = \sum_{n \in \mathbb{Z}} \langle f, \phi_{n,j} \rangle \phi_{n,j}(t) = \sum_{n \in \mathbb{Z}} a_j(n) \phi_{n,j}(t)$$

where $a_j(n) = \langle f, \phi_{n,j} \rangle = \left[f * \overline{\phi}_j \right](n)$

- frequency response of the approximation coefficients $a_j(n)$ depends on the frequency response of the scaling function
- scaling function typically a low-pass, so this becomes a low-frequency approximation.
- larger scale gives a coarse approx, so lower-freq.
- consistent with scaling law (as we dilate scaling function, the filter pass-band is reduced)

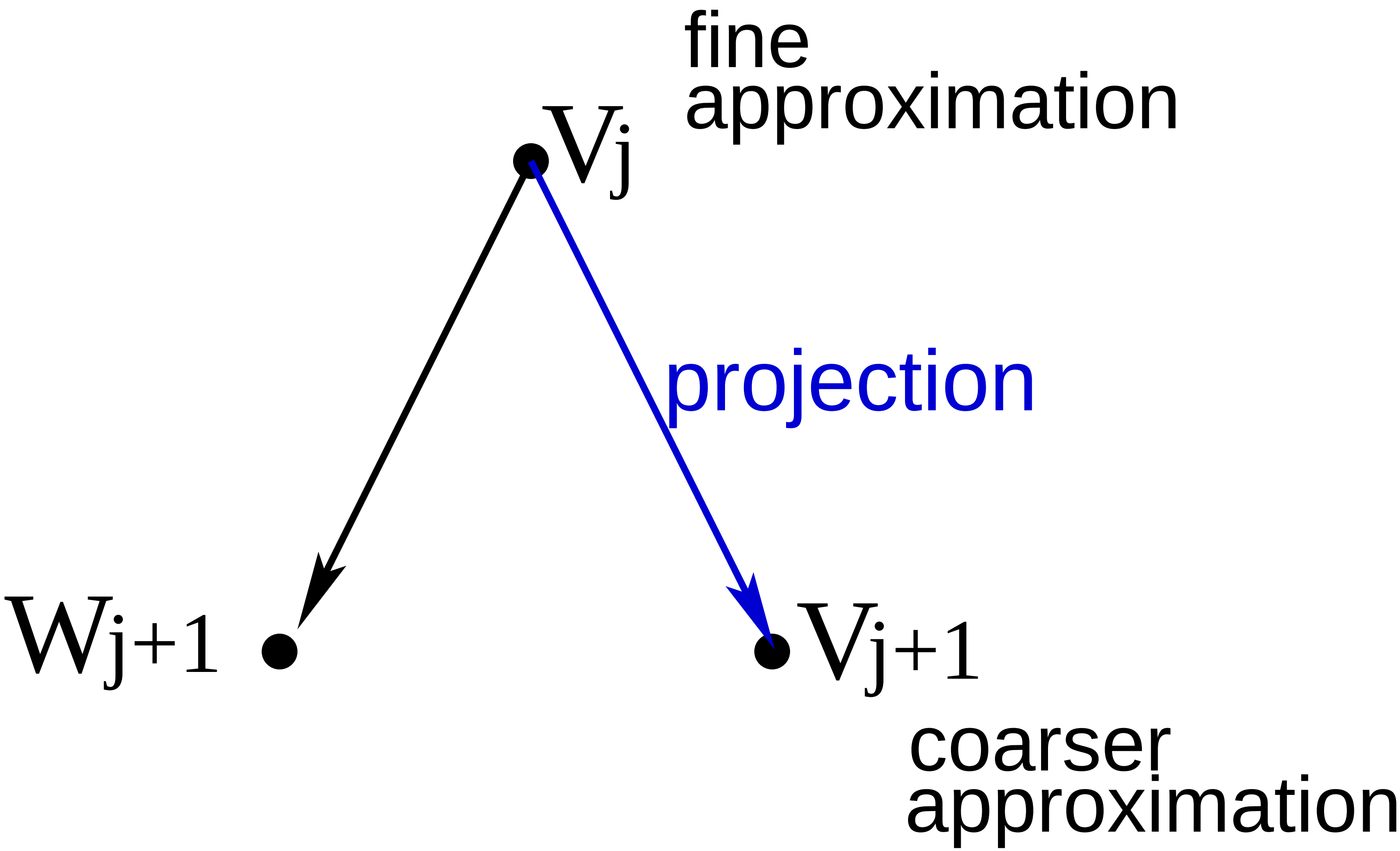
Relationship to wavelets

The approximation of a function $\hat{f}_j \in V_j$ into V_{j+1} is

$$\hat{f}_{j+1}(t) = \sum_{n \in \mathbb{Z}} \left\langle \hat{f}_j, \phi_{n,j} \right\rangle \phi_{n,j}(t)$$

- when we approximate a function $f \in V_j$ with a coarser approximation $f \in V_{j+1}$ we lose detail
- prefer a decomposition of V_j into an orthogonal sum of V_{j+1} and W_{j+1}
  - W_{j+1} are the bits we lost in the approximation
  - should be able to recombine V_{j+1} and W_{j+1} to get back to $f \in V_{j+1}$
- natural to associate W_{j+1} somehow with the wavelet

Relationship to wavelets



Relationship to wavelets

Properties imposed by the relationship

1. $W_{j+1} \subset V_j$, so the basis vectors of W_{j+1} must be $\in V_j$.
   - we want the basis of W_{j+1} to be wavelets, so

$$\psi_{j+1} \in W_{j+1} \subset V_j$$

   - hence we can represent ψ_{j+1} in terms of ψ_j, i.e.,

$$\psi_{0,j+1}(t) = \sum_n a_j(n)\phi_{n,j}(t)$$

2. V_j is an orthogonal sum of V_{j+1} and W_{j+1}, so

$$\langle \phi_{0,j+1}(t), \psi_{n,j+1}(t) \rangle = 0$$

Relationship to wavelets

Take the properties above (for $j = 0$), and work out relationships between mother wavelet, and mother scaling function. First take the property that

$$\psi_{0,j+1}(t) = \sum_n a_j(n)\phi_{n,j}(t)$$

for $j = 0$

$$\begin{aligned}
\psi_{0,1}(t) &= \sum_n a_1(n)\phi_{n,0}(t) && (1)\\
\psi(t/2)/\sqrt{2} &= \sum_n a_1(n)\phi(t-n) && (2)\\
\psi(t) &= \sum_n a(n)\phi(2t-n) && (3)
\end{aligned}$$

Relationship to wavelets

Combining the first and second properties (from p.51)

$$\psi(t) = \sum_n a(n)\phi(2t-n)$$

$$\langle \psi(t), \phi(t-n) \rangle = \int_{-\infty}^{\infty} \psi(t)\phi(t-n)\,dt = 0$$

we get

$$\int_{-\infty}^{\infty} \sum_k a(k)\phi(2t-k)\phi(t-n)\,dt = \sum_k a(k) \int_{-\infty}^{\infty} \phi(2t-k)\phi(t-n)\,dt = 0$$

which defines possible values for $a(k)$

Example: Haar wavelets

Piecewise constant approximation: so take

$$\phi(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Basis functions for approximations are rectangular pulses.

$$\begin{aligned} \sum_k a(k) \int_{-\infty}^{\infty} \phi(2t-k)\phi(t-n)\, dt &= 0 \\ \sum_k a(k) \int_{n}^{n+1} \phi(2t-k)\, dt &= 0 \end{aligned}$$

Example: Haar wavelets

Now, $\phi(2t-k)$ is only positive in the interval $[n, n+1]$ for $k = 2n$ or $2n+1$

$$
\begin{aligned}
\sum_k a(k) \int_n^{n+1} \phi(2t-k)\, dt &= 0 \\
a(2n) + a(2n+1) &= 0
\end{aligned}
$$

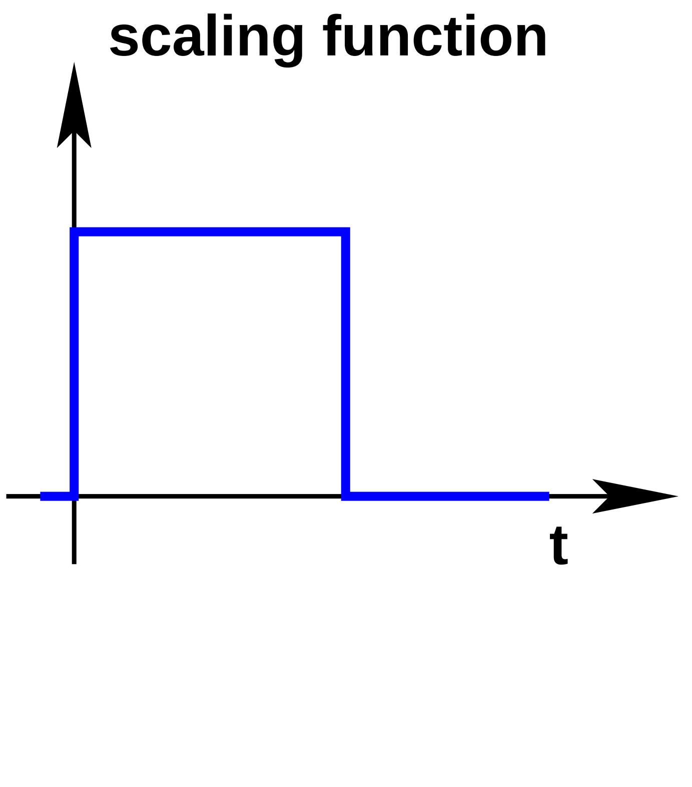
because in both cases the integral is 1.

The function with minimal support that satisfies this relationship has $a(0) = 1$ and $a(1) = -1$ and all other $a(k) = 0$, so

$$
\psi(t) = \phi(2t) - \phi(2t-1)
$$

Haar wavelets

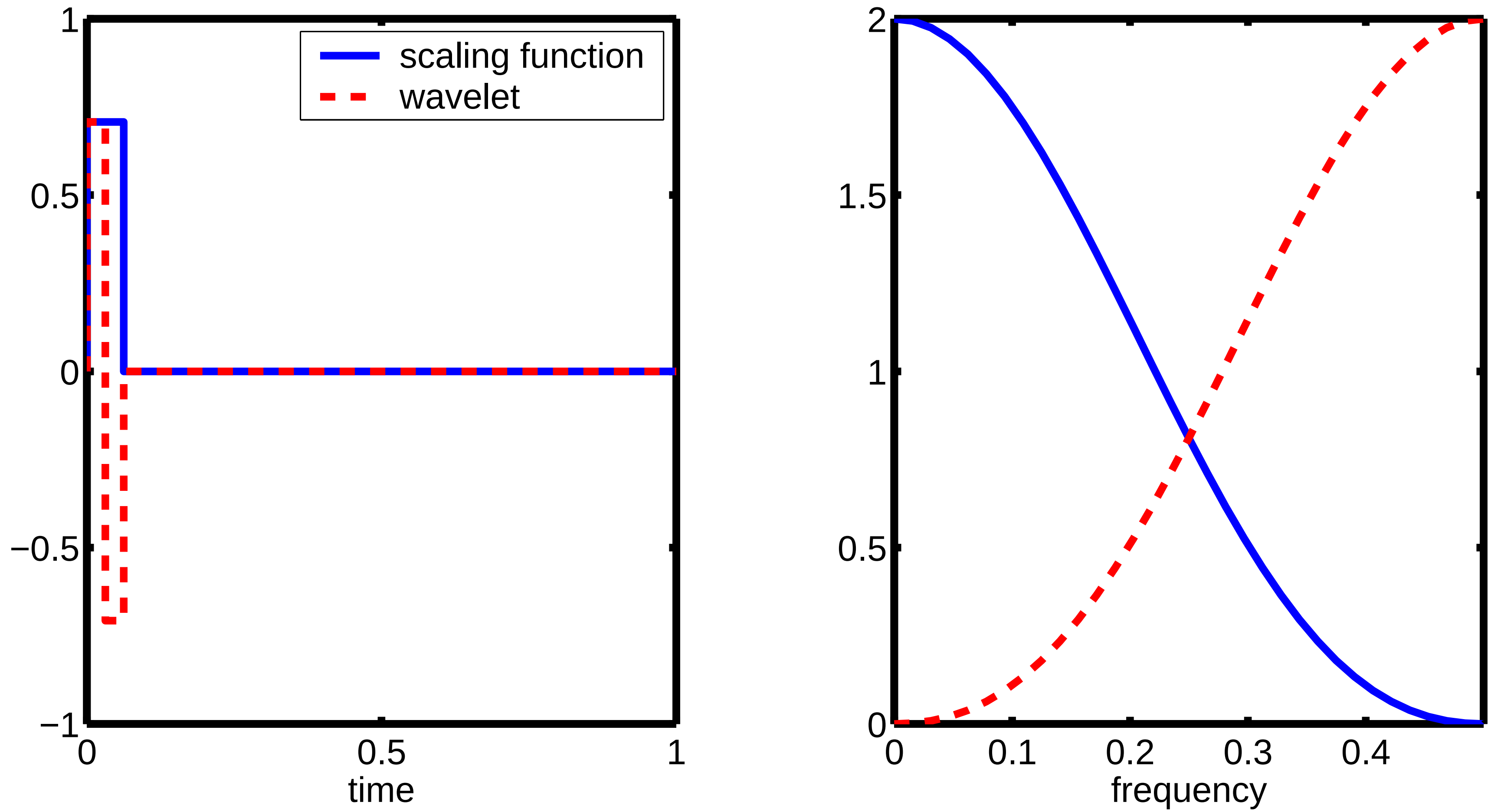
Scaling and wavelet functions for the Haar transform shown below



Approximations are piecewise constant curves.

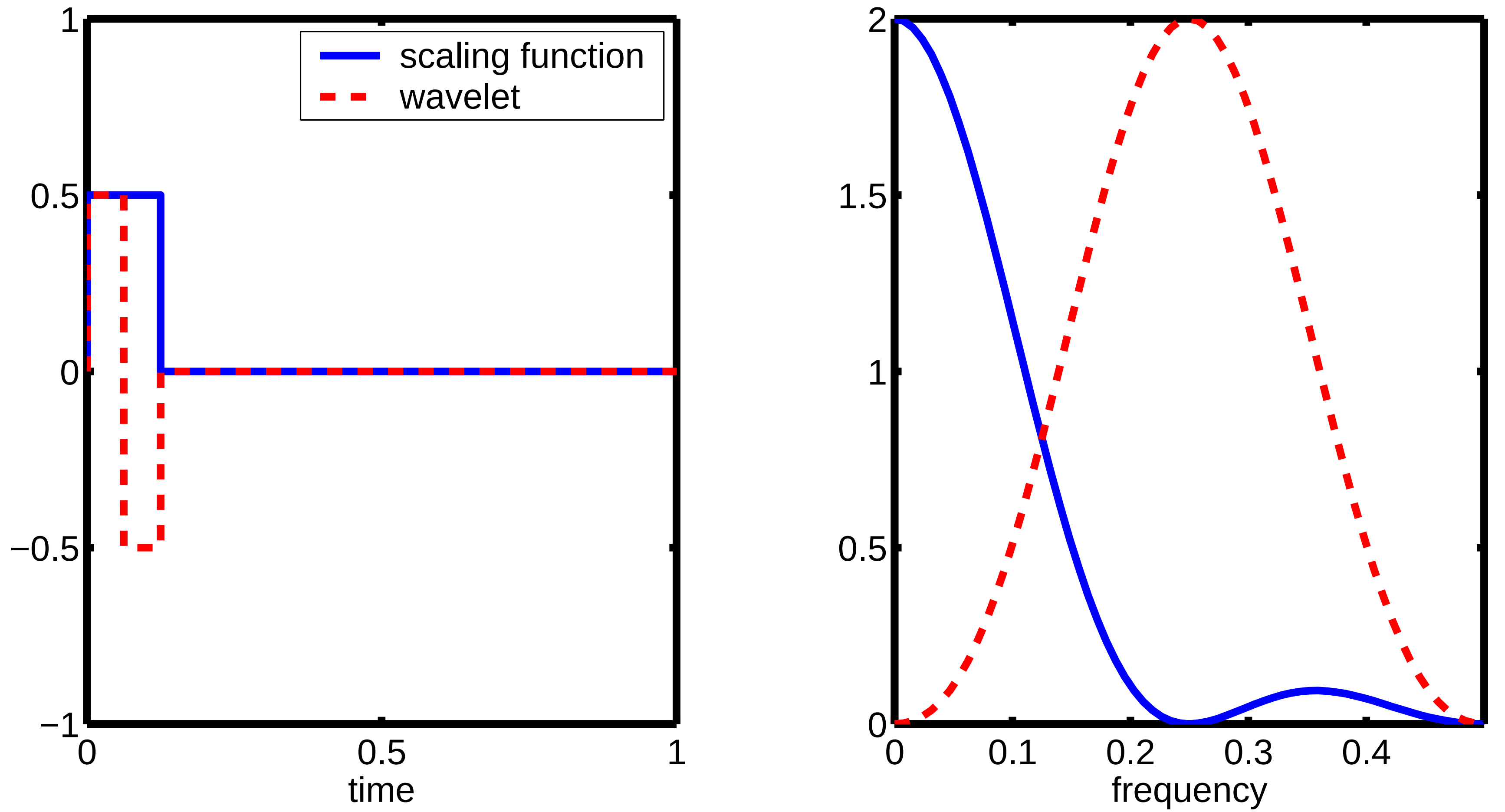
Haar wavelets: freq. representation

At scale $j = 0$, scale by 2^0 $\left(\psi_{0,j}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t}{2^j}\right)\right)$



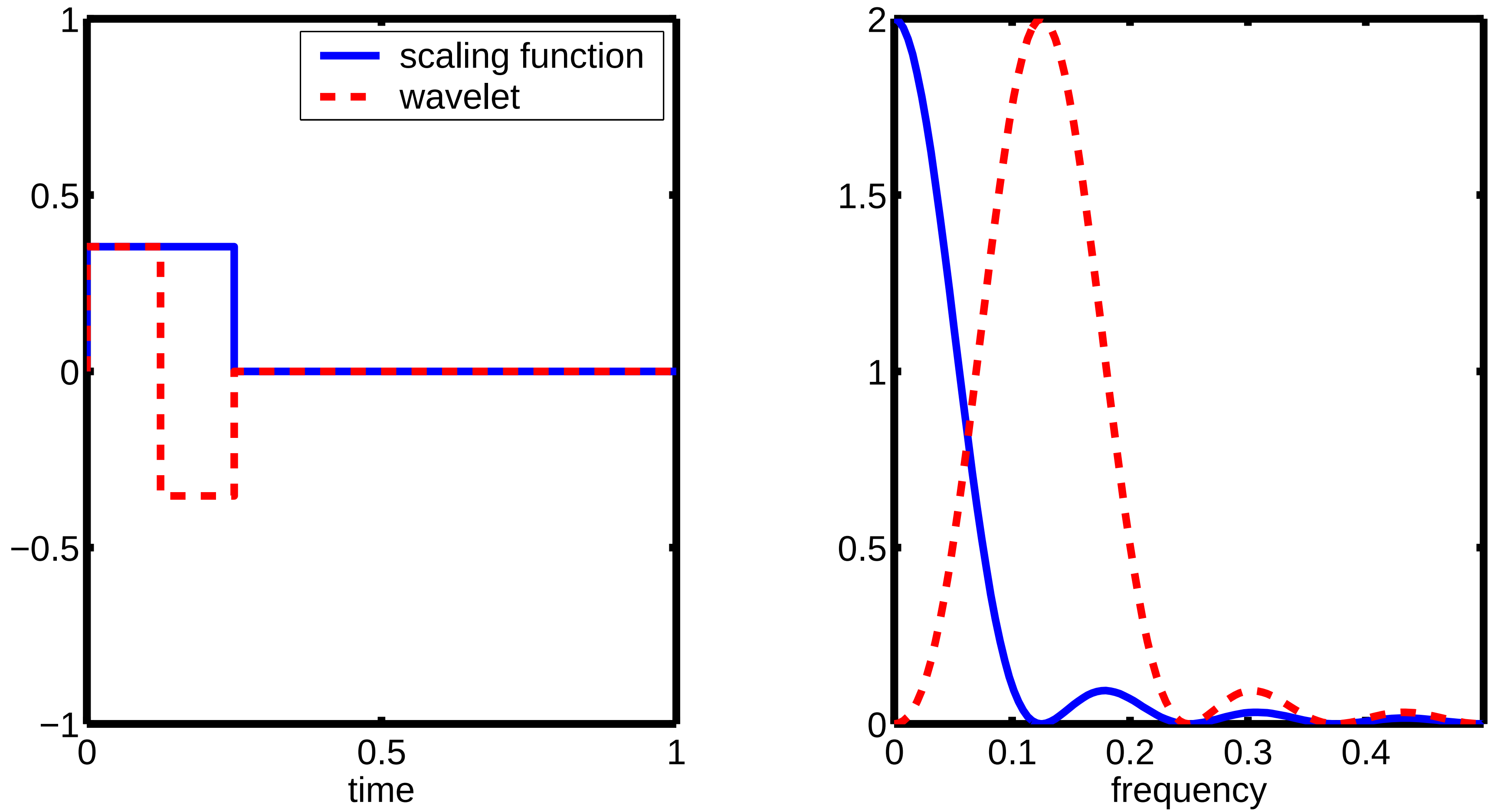
Haar wavelets: freq. representation

At scale $j = 1$, scale by 2^1 $\left(\psi_{0,j}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t}{2^j}\right)\right)$



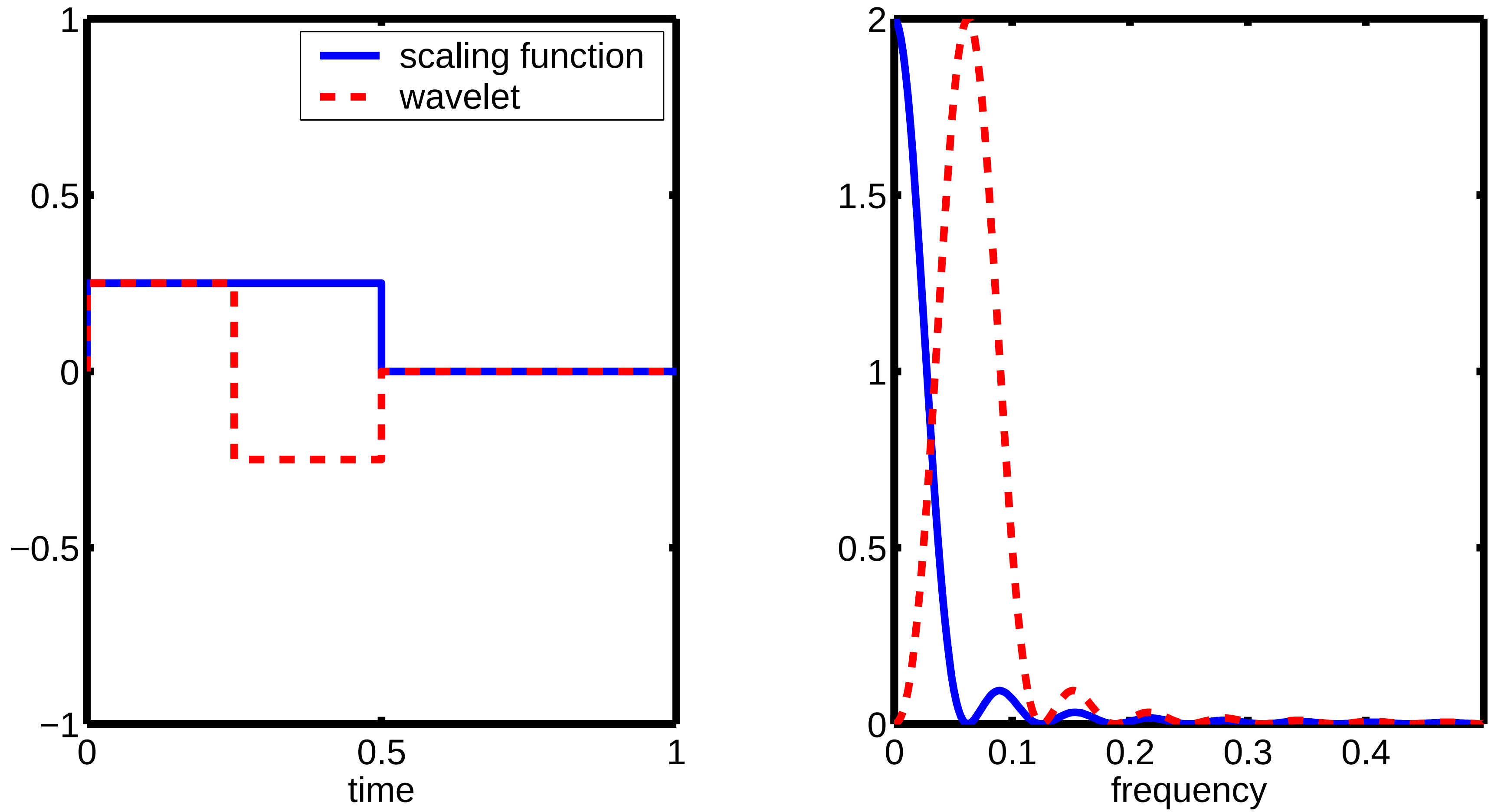
Haar wavelets: freq. representation

At scale $j = 2$, scale by 2^2 ($\psi_{0,j}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t}{2^j}\right)$)



Haar wavelets: freq. representation

At scale $j = 3$, scale by 2^3 $\left(\psi_{0,j}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t}{2^j}\right)\right)$

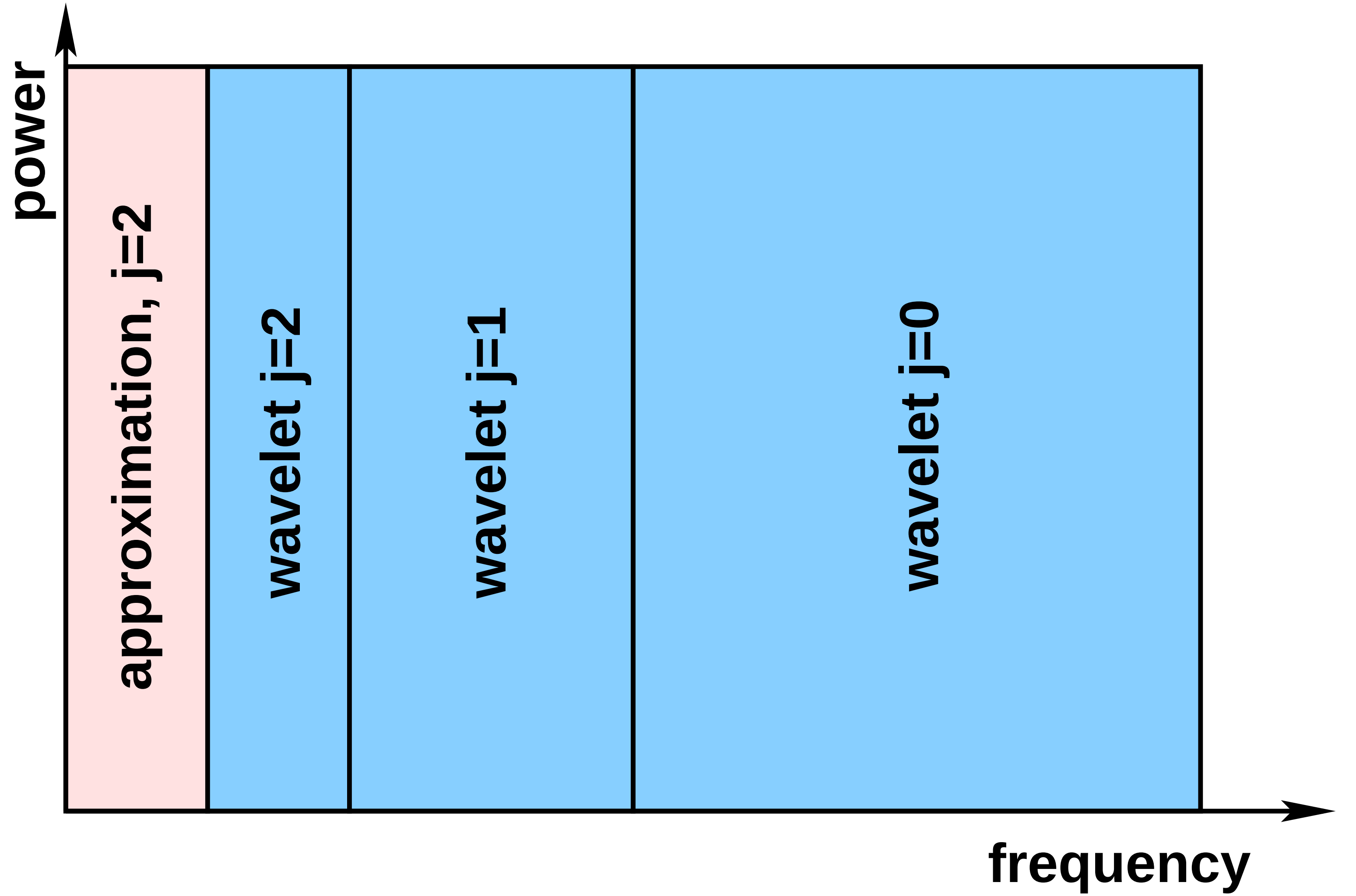


Haar wavelets: freq. representation

- scaling function is a low-pass
  - approximations are low-freq. approximations
  - larger scale, low-frequency stop-band
- wavelet function is a band-pass
  - together with scaling they break up a block of the frequency spectrum

Subband coding

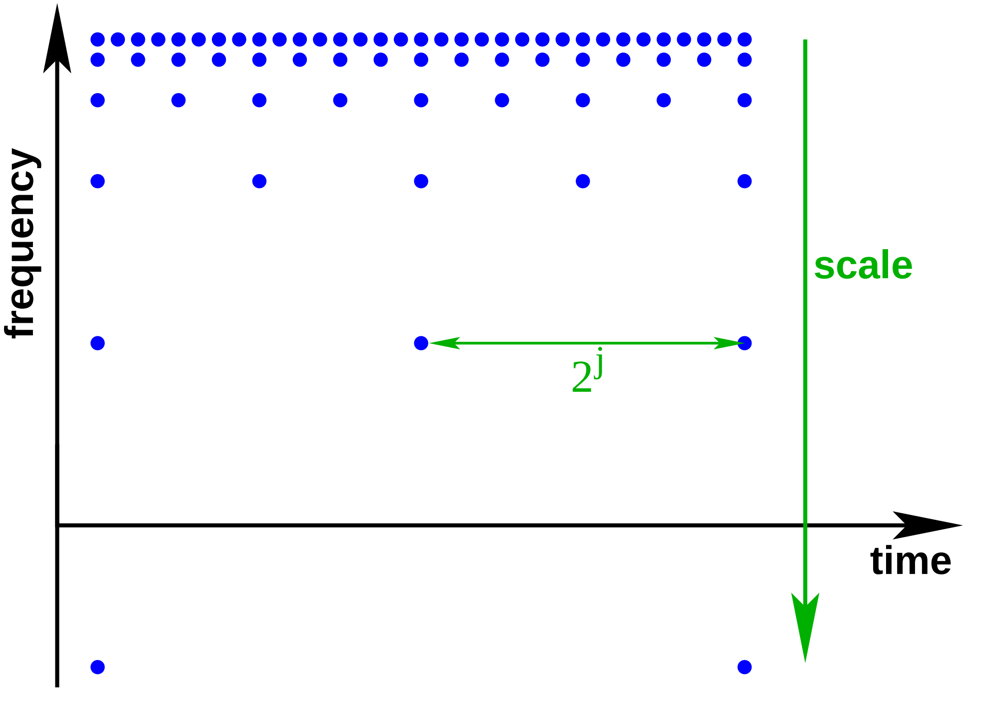
The idea (looking across frequencies or scales) is that the transform breaks frequency spectrum into bands.



MRA and wavelets

Take mother wavelet $\psi(t)$, with orthogonal discrete wavelet basis on the dyadic grid

$$\psi_{n,j}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t}{2^j} - n\right)$$



Form closed subspaces

$$W_j = \mathrm{Sp}\{\psi_{n,j} | n \in \mathbb{Z}\}$$

As noted earlier,

$$V_j = \oplus_{i=j}^{\infty} W_i$$

is a MRA and the scaling function ϕ was also given earlier, and $V_{j-1} = V_j \oplus W_j$ so an orthogonal projection into V_{j-1} can be decomposed into projections into V_j and W_j.

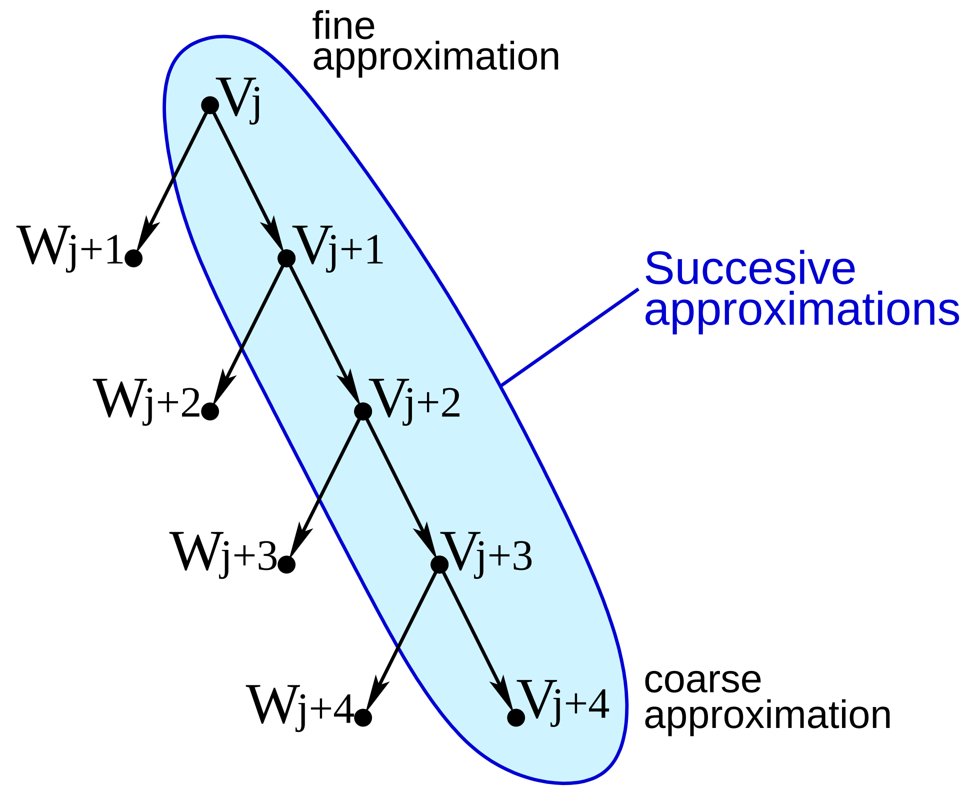
Successive decompositions

We can iteratively decompose approximation V_j into a wavelet part (the **details**) and a coarser scale **approximation** $V_{j-1} = V_j \oplus W_j$ using the projection operation

Form $f_{j-1} \in V_{j-1}$ by

$$\hat{f}_{j+1} = \sum_{n\in\mathbb{Z}} \langle \hat{f}_j, \phi_{n,j+1} \rangle \phi_{n,j+1}$$
$$= \sum_{n\in\mathbb{Z}} a_{n,j+1} \phi_{n,j+1}$$

$$\dot{f}_{j+1} = \sum_{n\in\mathbb{Z}} \langle \hat{f}_j, \psi_{n,j+1} \rangle \psi_{n,j+1}$$
$$= \sum_{n\in\mathbb{Z}} d_{n,j+1} \psi_{n,j+1}$$



MRA and wavelets

$$\begin{aligned}\hat{f}_j &= \hat{f}_{j+1} + \dot{f}_{j+1} \\ &= \sum_{n\in\mathbb{Z}} a_{n,j+1}\,\phi_{n,j+1} + \sum_{n\in\mathbb{Z}} d_{n,j+1}\,\psi_{n,j+1}\end{aligned}$$

- $\hat{f}_{j+1}$ is a coarser scale approximation of f
- it loses some "detail"
- **details** are captured in the wavelet component $\dot{f}_{j+1}$
- often call the coefficients
  - $a_{n,j}$ the approximation
  - $d_{n,j}$ the details
- As $j \to -\infty$ the approximation $\hat{f}_j \to f$

The Scaling Function

The above representation requires wavelet coefficients for $s = -\infty, \ldots, \infty$ and $u = -\infty, \ldots, \infty$. We can still manage if we have coefficients $\langle f, \psi_{u,s} \rangle$ for $s < s_0$, by using a **scaling function** $\phi(t)$.

- can be thought of as a low frequency (high scale) approximation of the signal
- form scaling functions $\phi_{u,s}(t)$ by the same dilations and translation used to form wavelets
- scaling function $\phi(t)$ brings in info from scales $s > 1$, so it is the aggregation of wavelets above this scale

$$|\Phi(\omega)|^2 = \int_1^\infty |\Psi(s\omega)|^2 \frac{1}{s}\, ds = \int_\omega^\infty |\Psi(\xi)|^2 \frac{1}{\xi}\, d\xi$$

The Scaling Function

- DWT representation

$$f = \sum_{j=j_0}^{\infty} \sum_{n=-\infty}^{\infty} \langle f, \psi_{n,j} \rangle \psi_{n,j} + \sum_{n=-\infty}^{\infty} \langle f, \phi_{n,j_0} \rangle \phi_{n,j_0}$$

Wavelet Properties

Potential wavelet properties

- finite support
- vanishing moments
- orthogonal/ bi-orthogonal
- complex(analytic) or real
- redundant (framelets)

Applications

- edge (and anomaly) detection
- motion detection
- denoising
- compression (JPEG 2000)

To do these, we will need to

- perform wavelet transforms on discrete data.
- make the algorithms efficient (as with FFT)

Appendices

Riesz basis

A family of elements $\{e_n\}_{n\in\mathbb{Z}}$ from a Hilbert space $\mathbf{H}$ is said to be a Riesz basis of $\mathbf{H}$ if it is linearly independent and there exists $A > 0$ and $B > 0$ such that for any $f \in \mathbf{H}$ one can find λ_n with

$$f(t) = \sum_{n=-\infty}^{\infty} \lambda_n e_n$$

which satisfies

$$\frac{1}{B}\|f\|^2 \leq \sum_{n=-\infty}^{\infty} |\lambda_n|^2 \leq \frac{1}{A}\|f\|^2$$

If $A = B$ the frame is said to be tight.