Document: Math 271 (Winter 2012)
Professor: Strohmer
Latest Update: October 12, 2012
Author: Jeff Irion
http://www.math.ucdavis.edu/~jlirion

Contents

1 1-12-12

1.1 Basic Fourier Analysis

Let $f \in L^1(\mathbb{R})$, i.e. $\|f\|_{L^1(\mathbb{R})} = \int_{\mathbb{R}} |f(x)|\, dx < \infty$. Its *Fourier transform* is $\mathcal{F} : \mathbb{R} \to \hat{\mathbb{R}} \approx \mathbb{R}$, given by

$$(\mathcal{F}f)(\omega) = \hat{f}(\omega) = \int_{\mathbb{R}} f(x) e^{-2\pi i x \omega}\, dx$$

where ω is the frequency (math & engineering) or momentum (physics).

Theorem 1.1. ***Plancherel***

If $f \in L^1 \cap L^2(\mathbb{R})$, then $\|f\|_{L^2(\mathbb{R})} = \|\hat{f}\|_{L^2(\mathbb{R})}$ and the Fourier transform extends to a unitary operator on $L^2(\mathbb{R})$.

$\Rightarrow \quad \mathcal{F}^* = \mathcal{F}^{-1}$

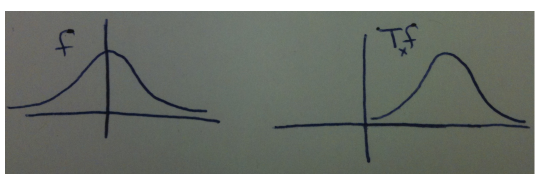
Theorem 1.2. ***Parseval***

$$\langle f, g \rangle = \left\langle \hat{f}, \hat{g} \right\rangle \qquad \text{for } f, g \in L^2(\mathbb{R})$$

1.2 Two Basic Operators

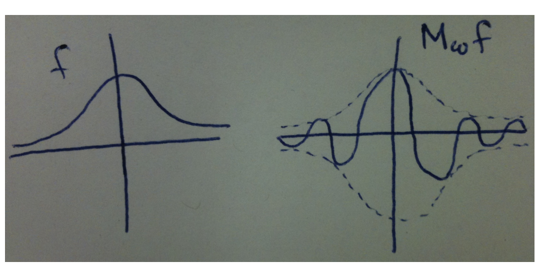
The *translation operator* (also called the *shift operator*):

$$T_x f(t) = f(t - x), \qquad x, t \in \mathbb{R}$$



The *modulation operator* (also called the *frequency shift operator*):

$$M_\omega f(t) = e^{2\pi i \omega t} f(t)$$



Fundamental Calculation:

$$\begin{aligned} T_x M_\omega f(t) &= M_\omega f(t-x) \\ &= e^{2\pi i\omega(t-x)} f(t-x) \\ &= e^{-2\pi i\omega x} e^{2\pi i\omega t} f(t-x) \\ &= e^{-2\pi i\omega x} M_w T_x f(t) \\ \Rightarrow \qquad T_x M_w &= e^{-2\pi i\omega x} M_w T_x \end{aligned}$$

Thus, T_x and M_ω commute iff $x \cdot \omega \in \mathbb{Z}$.

T_x and M_ω are unitary:

$$\widehat{(T_x f)} = M_{-x}\hat{f} \quad \Leftrightarrow \quad \mathcal{F}(T_x f) = M_{-x}\hat{f}$$
$$\widehat{(M_\omega f)} = T_\omega \hat{f} \quad \Leftrightarrow \quad \mathcal{F}(M_\omega f) = T_\omega \hat{f}$$

Lemma 1.3.

1. T_x and M_ω are unitary operators.
2. $$\widehat{(T_x f)} = M_{-x}\hat{f}$$
3. $$\widehat{(M_x f)} = T_x \hat{f}$$

Definition 1.4. ***Convolution***

Let $f, g \in L^1(\mathbb{R})$. Then the *convolution* $f * g$ is

$$(f * g)(x) = \int_{\mathbb{R}} f(y) g(x-y)\, dy$$

The following hold:

$$\|f * g\|_1 \le \|f\|_1 \|g\|_1$$
$$\widehat{(f * g)} = \hat{f} \cdot \hat{g}$$

Theorem 1.5. ***Inversion of Fourier Transform***

If $f \in L^1(\mathbb{R})$ and $\hat{f} \in L^1(\mathbb{R})$, then

$$f(x) = \int_{\mathbb{R}} \hat{f}(\omega) e^{2\pi i x\omega}\, d\omega$$

for all $x, \omega \in \mathbb{R}$.

Fourier Transform and Derivatives:
Let $\mathcal{D}^\alpha$ be the derivative operator of order α. Then

$$\widehat{(\mathcal{D}^\alpha f)}(w) = (2\pi i\omega)^\alpha \hat{f}(\omega)$$

and

$$(\widehat{(-2\pi i\omega)^\alpha} f)(w) = \mathcal{D}^\alpha \hat{f}(\omega).$$

Some functions and their Fourier Transforms:

Example 1.6. ***Indicator Function Fourier Transform***

$$f(x) = \mathbf{1}_{[-T/2,T/2]}$$
$$\hat{f}(\omega) = \frac{\sin \pi\omega t}{\pi\omega} = \text{ sinc function (sinus cardinalus = cardinal sine)}$$

Theorem 1.7. ***Gaussian Fourier Transform***

$$\varphi_a(x) = e^{-\pi x^2/a}$$

(If $a = 1$, we write $\varphi(x)$.)

For all $a > 0$:

$$\begin{aligned}\hat{\varphi}_a(\omega) &= \sqrt{a}\varphi\left(\frac{\omega}{a}\right)\\ &= \sqrt{a}\varphi_{1/a}(\omega)\end{aligned}$$

So $\hat{\varphi}(\omega) = \varphi(x)$ is a fixed point.

Proof.

$$\begin{aligned}\frac{d}{d\omega}\hat{\varphi}_a(\omega) &= (\widehat{-2\pi i x\varphi_a})(\omega)\\ \widehat{\left(ia\frac{d}{dx}\varphi_a\right)}(\omega) &= ia(2\pi i\omega)\hat{\varphi}_a(\omega)\\ \frac{d}{d\omega}\hat{\varphi}_a(\omega) &= -2\pi a\omega\hat{\varphi}_a(\omega)\\ \hat{\varphi}_a(\omega) &= Ce^{-\pi a\omega^2}\end{aligned}$$

Now we need to show that $C = \sqrt{a}$.

$$C = \hat{\varphi}_a(0) = \int_{\mathbb{R}} e^{-\pi x^2/a}\,dx$$

□

Definition 1.8. ***Dilation***

$$D_a f(x) = \frac{1}{\sqrt{a}} f\left(\frac{x}{a}\right)$$

Dilation is unitary.

It is easily verified that $\|f\|_2 = \|D_a f\|_2$.

Theorem 1.9.

$$\widehat{(D_a f)} = D_{1/a}\hat{f}$$

1.3 Smoothness and the Fourier Transform

Rule of thumb: smoothness of $f \Rightarrow$ decay of $\hat{f} \rightarrow$ Sobolev spaces.

Lemma 1.10.

$$\mathcal{D}^\alpha f \in L^2(\mathbb{R}) \; \forall \; \alpha \leq n \quad \Leftrightarrow \quad \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \underbrace{(1+|\omega|)^2}_{\text{weight function}} \; d\omega < \infty$$

This says that $|\hat{f}(\omega)|$ must decay faster than $|\omega|^n$ if f is n times differentiable.

Fourier series are useful for analyzing periodic functions. Assume f is 1-periodic:

$$f(x) = f(x+k), \qquad k \in \mathbb{Z}.$$

Identify f with an interval... the torus, $\mathbb{T}$.

Theorem 1.11. ***Plancherel (on*** $\mathbb{T}$***)***

Let $f \in L^2(\mathbb{T})$ and let $\hat{f}(n) = \int_{\mathbb{T}} f(x)e^{-2\pi ixn}\, dx$ be the nth Fourier coefficient. Then

$$f(x) = \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{2\pi inx}$$

and

$$\|f\|_{L^2(\mathbb{T})} = \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2$$

Poisson's Summation Formula (PSF, equivalent to the Shannon sampling theorem) relates Fourier series to Fourier transforms.

Lemma 1.12. ***Periodization Trick***

If $f \in L^1(\mathbb{R})$ then for all $\alpha > 0$,

$$\int_{\mathbb{R}} f(x)\,dx = \int_0^{\alpha} \underbrace{\sum_{k\in\mathbb{Z}} f(x+\alpha k)}_{\alpha\text{-periodic function}}\,dx$$

Proof. The translated intervals are disjoint except for a set of measure 0, and their union is $\mathbb{R}$. Thus,

$$\begin{aligned}\int_{\mathbb{R}} f(x)\,dx &= \sum_{k\in\mathbb{Z}} \int_{\alpha k}^{\alpha(k+1)} f(x)\,dx \\ &= \int_0^{\alpha} \sum_{k\in\mathbb{Z}} f(x+\alpha k)\,dx.\end{aligned}$$

□

2 1-17-12

2.1 Poisson Summation Formula

Theorem 2.1. ***Poisson Summation Formula (PSF), Version 1***

Assume that for some $\epsilon > 0$ and $c > 0$, we have

$$|f(x)| \leq c(1+|x|)^{-1-\epsilon} \qquad \text{and} \qquad |\hat{f}(\omega)| \leq c(1+|\omega|)^{-1-\epsilon}$$

Then

$$\underbrace{\sum_{n\in\mathbb{Z}} f(x+n)}_{\text{periodization in time domain}} = \underbrace{\sum_{n\in\mathbb{Z}} \hat{f}(n)e^{2\pi inx}}_{\text{sampling in frequency domain}} . \qquad \text{(PSF)}$$

This identity holds pointwise for all $x \in \mathbb{R}$, and both sums converge absolutely for all $x \in \mathbb{R}$.

Proof. Set

$$g(x) := \sum_{n\in\mathbb{Z}} f(x+n).$$

Then g is a 1-periodic function. The decay assumption on f implies that $f \in L^1(\mathbb{R}) \Rightarrow g \in L^1(\mathbb{T})$.

$$\begin{aligned}
\hat{g}(n) &= \int_{[0,1]} g(x)e^{-2\pi ixn}\,dx \\
&= \int_{[0,1]} \sum_{k\in\mathbb{Z}} f(x+k)e^{-2\pi in(x+k)}\,dx \\
&= \int_{[0,1]} f(x)e^{-2\pi inx}\,dx \qquad \text{(Fubini) Periodization Lemma} \\
&= \hat{f}(n)
\end{aligned}$$

The decay assumption on $\hat{f}$ implies that $\sum_{n\in\mathbb{Z}} |\hat{f}(n)| < \infty \Rightarrow g$ has an absolutely convergent Fourier series:

$$\begin{aligned}
g(x) &= \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{2\pi inx} \\
\Rightarrow \qquad \sum_{n\in\mathbb{Z}} f(x+n) &= \sum_{n\in\mathbb{Z}} \hat{f}(n)e^{2\pi inx}
\end{aligned}$$

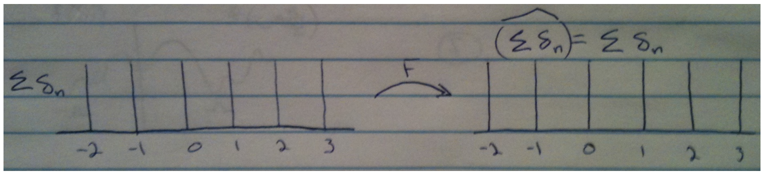
□

Remarks:

1. We can replace absolute convergence by convergence in L^2 and pointwise equality by equality a.e. If $\sum_n f(x+n) \in L^2(\mathbb{T})$ and $\sum_n |\hat{f}(n)|^2 < \infty$, then (PSF) holds a.e.

2. Denote $\delta_x : \langle \delta_x, f\rangle = \overline{f(x)}$. Then PSF says

$$\widehat{(\sum_{n\in\mathbb{Z}} \delta_n)} = \sum_{n\in\mathbb{Z}} \delta_n$$

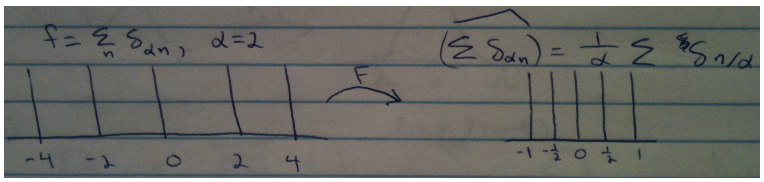


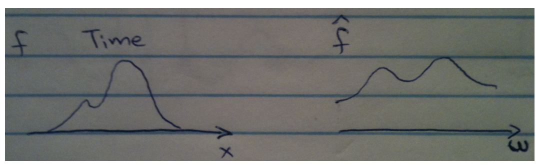
3. More general PSF:

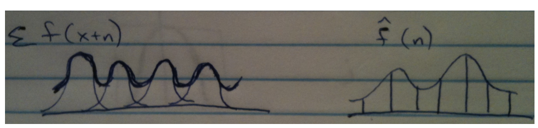
Theorem 2.2. ***Poisson Summation Formula (PSF), Version 2***

Let $\alpha > 0$. Then

$$\sum_{n\in\mathbb{Z}} f(x+n\alpha) = \frac{1}{\alpha}\sum_{n\in\mathbb{Z}} \hat{f}\left(\frac{n}{\alpha}\right) e^{2\pi i x n/\alpha}$$







2.2 Shannon's Sampling Theorem

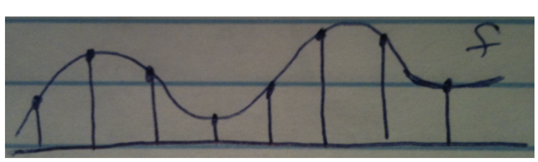
Given a function $f \in L^2(\mathbb{R})$, we want to store f in digital form. $\Rightarrow$ Need to represent f by a discrete set of data (as fast as possible). The Shannon Sampling Theorem is an intermediate step.

Naive approach: Sample f at equally spaced points, $n \cdot \tau,\ n \in \mathbb{Z}$, sampling interval $= \tau > 0$.

(Note: in everything we've done so far, we can replace $\mathbb{R}$ with $\mathcal{G}$, a locally compact group. Examples and their corresponding "Fourier spaces" include $\mathbb{R}^d \to \hat{\mathbb{R}}^d$, $\mathbb{Z}^d \to \mathbb{T}^d$, $\mathbb{T}^d \to \mathbb{Z}^d$, $\mathbb{C}^d \to \mathbb{C}^d$, $\mathbb{R}^d \times \mathbb{Z}^n \times \mathbb{C} \to \mathbb{R}^d \times \mathbb{T}^n \times \mathbb{C}$.)

We can associate to each sampling value $f(n\tau)$ a Dirac distribution $f(n\tau) \Leftrightarrow \delta(x - n\tau)$. By doing this, we get a discretized function ("signal") f_d:

$$f_d(x) = \sum_n f(n\tau)\delta(x - n\tau)$$



The Fourier Transform of $\delta(x - n\tau)$ is $e^{-2\pi i n x}$ $\Rightarrow$

$$\hat{f}_d(\omega) = \sum_n f(n\tau) e^{-2\pi i n \tau \omega}.$$

How is $\hat{f}_d$ related to $\hat{f}$?

PSF implies that

$$\hat{f}_d(\omega) = \frac{1}{\tau} \sum_{k \in \mathbb{Z}} \hat{f}\left(\omega - \frac{k}{\tau}\right) \tag{2.1}$$

Thus, discretizing a function in time corresponds to periodizing its Fourier transform.

Can we recover f from f_d?
In general, NO.
But sometimes, YES.

Definition 2.3. ***Bandlimited Function***

A function $f \in L^2(\mathbb{R})$ is called *bandlimited* if $\hat{f}(\omega) = 0$ for all $|\omega| > \Omega$.

$$B_\Omega = \left\{ f \in L^2(\mathbb{R}) \;\middle|\; \operatorname{supp} \hat{f} \subseteq [-\Omega, \Omega] \right\}$$

For example, in phone conversations, we assume that voice is limited to 8 kHz. Music typically goes up to 20 kHz (we can't hear higher frequencies).

Theorem 2.4. ***Shannon's Sampling Theorem***

(Shannon 1948, Whittaker 1915, Kotelnikov 1933, Roabe 1939, ...)

Let $f \in B_\Omega$. Then if we have discrete sampling, we can recover f from the sampling points $f(n\tau)$ by

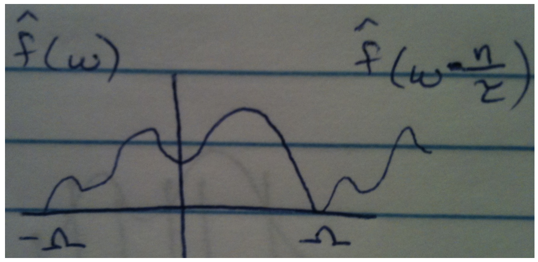
$$f(x) = \sum_{n=-\infty}^{\infty} f(n\tau) h_\tau(x - n\tau),$$

where

$$h_\tau(x) = \tau \cdot \frac{\sin \frac{\pi x}{\tau}}{\pi x}$$

and $\tau = \frac{1}{2\Omega}$. (Note: h_τ is more or less the sinc function.)

Proof. If $n \neq 0$, then $\operatorname{supp} \hat{f}\left(\omega - \frac{n}{\tau}\right)$ does <u>not</u> intersect with $\operatorname{supp} \hat{f}(\omega)$ because $\hat{f}(\omega) = 0$ for $|\omega| > \Omega = \frac{1}{2\tau}$.



Hence, (2.1) implies that

$$\hat{f}_d(\omega) = \frac{1}{\tau}\hat{f}(\omega) \qquad \text{if } |\omega| \leq \frac{1}{2\tau}$$

The Fourier Transform of h_τ is

$$\hat{h}_\tau = \tau \mathbf{1}_{[-\frac{1}{2\tau}, \frac{1}{2\tau}]}$$

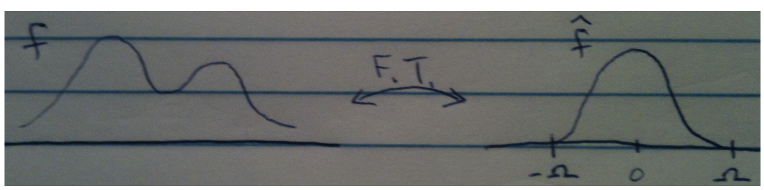
Since supp $\hat{f} \in \left[-\frac{1}{2\tau}, \frac{1}{2\tau}\right]$, we have that

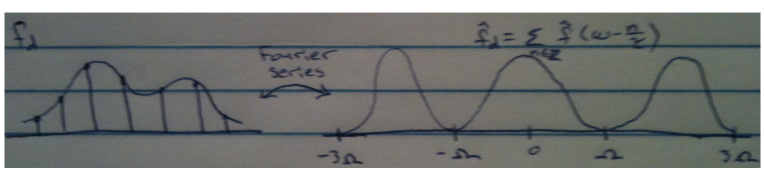
$$\hat{f}(\omega) = h_\tau(\omega) \cdot \hat{f}_d(\omega) \qquad \text{if } |\omega| \leq \frac{1}{2\tau}$$

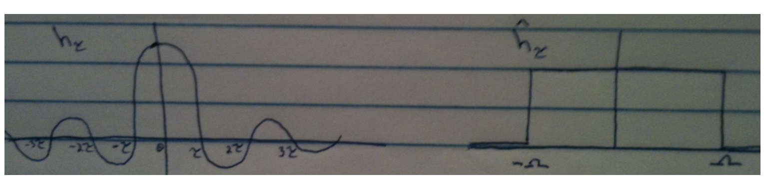
Now apply the Inverse Fourier Transform and get

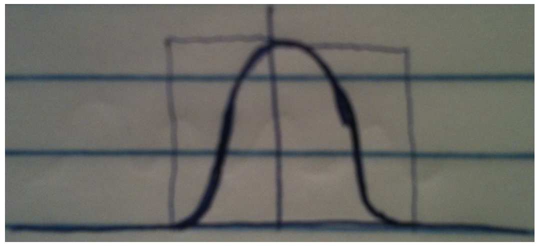
$$\begin{aligned} f(x) &= \mathcal{F}^{-1}(\hat{f})(x) = \mathcal{F}^{-1}(\hat{h}_\tau(\omega) \cdot \hat{f}_d(\omega)) \\ &= (h_\tau * f_d)(x) \\ &= h_\tau * \sum_{n=-\infty}^{\infty} f(n\tau)\delta(x - n\tau) \\ &= \sum_{n=-\infty}^{\infty} f(n\tau) h_\tau(x - n\tau) \end{aligned}$$

□

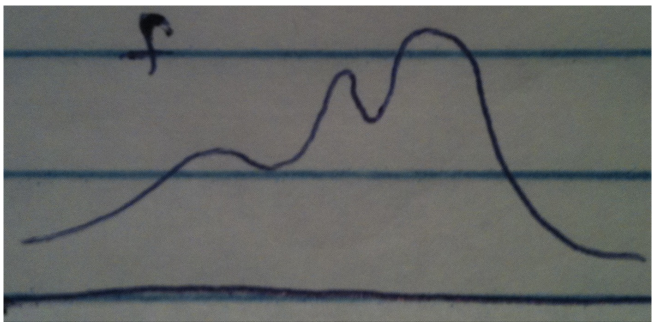


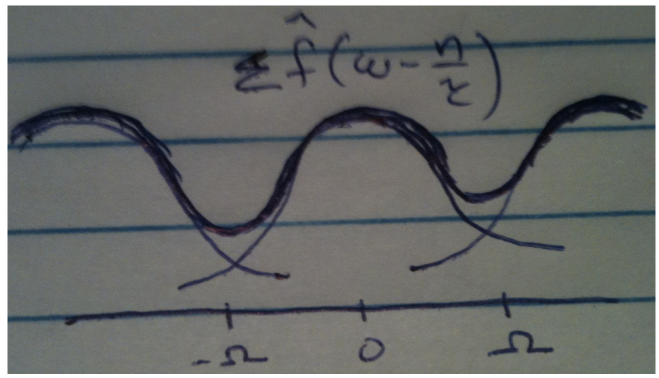


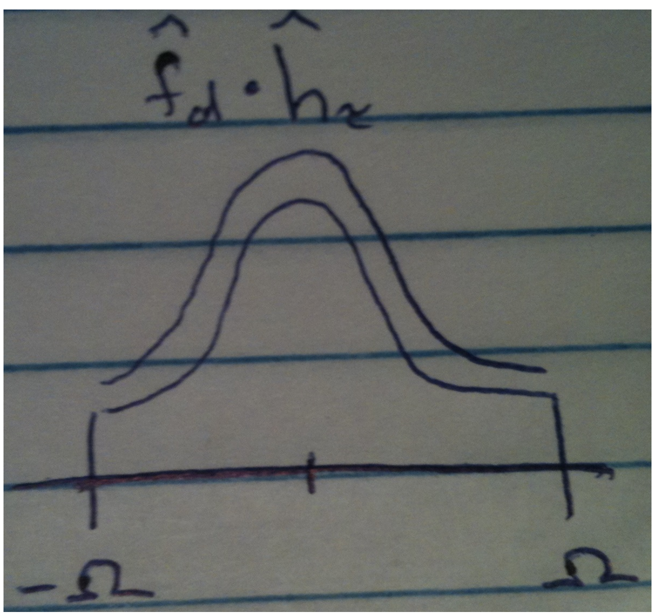


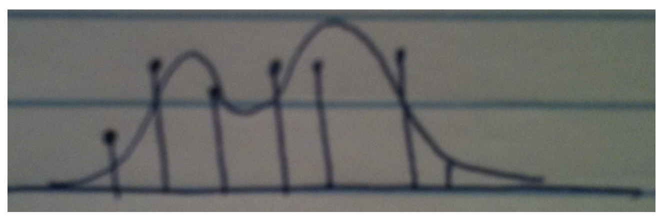


Aliasing:
Assume $\hat{f} \notin \left[-\frac{1}{2\tau}, \frac{1}{2\tau}\right]$, then the support of $\hat{f}\left(\omega - \frac{k}{\tau}\right)$ intersects the support of $\hat{f}(\omega)$. $\Rightarrow$ High-frequency components of f get "folded" $\rightarrow$ *aliasing*.









3 1-19-12

3.1 Shannon's Sampling Theorem (Review)

Let $\Omega = \frac{1}{2} \Rightarrow \tau = 1,\ h_\tau(x) = \text{sinc}(x) = \frac{\sin \pi x}{\pi x}$.

Shannon Sampling Theorem:

$$\begin{aligned}
f(x) &= \sum_{n=-\infty}^{\infty} f(n)\,\text{sinc}(x-n) && \text{(Cardinal series)} \\
f(x) &= (\text{sinc} * f)(x) = \int_{\mathbb{R}} f(y)\,\text{sinc}(x-y)\,dy \\
&= \int_{\mathbb{R}} f(y)\overline{\text{sinc}(y-x)}\,dy && \text{(using that sinc is even and real-valued)} \\
&= \langle f, T_x \,\text{sinc}\rangle
\end{aligned}$$

The sinc function acts as a *reproducing kernel* on the space $B_{1/2}$. That is, it reproduces the function value at the point. So we can define:

Definition 3.1. ***Projection Operator, Low-Pass Filter***

Define the operator $P_\Omega : L^2(\mathbb{R}) \to B_\Omega$ by

$$P_\Omega f = f * \text{sinc} \qquad \left(\Omega = \frac{1}{2}\right)$$

For $f \in B_\Omega$, then $f = f * \text{sinc}$ (because $f = \mathcal{F}'(\hat{f} \cdot \mathbf{1}_{[-\frac{1}{2},\frac{1}{2}]})$) and $P_\Omega^2 = P_\Omega$, $P_\Omega^* = P_\Omega$ (follows from real-valuedness of sinc function). Thus, P_Ω is an orthogonal projection. The sinc function acts as a *low-pass filter* (since the frequencies that survive are in $[-\Omega, \Omega]$). On the other hand, a band-pass filter could allow frequencies from multiple bands (intervals?) to survive.

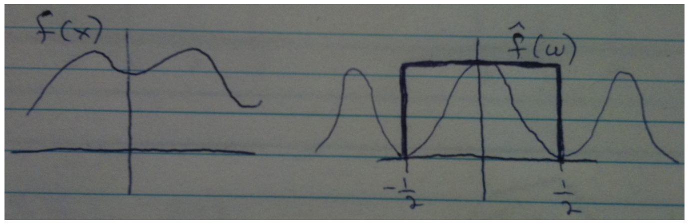
3.2 Limitations of Shannon's Sampling Theorem

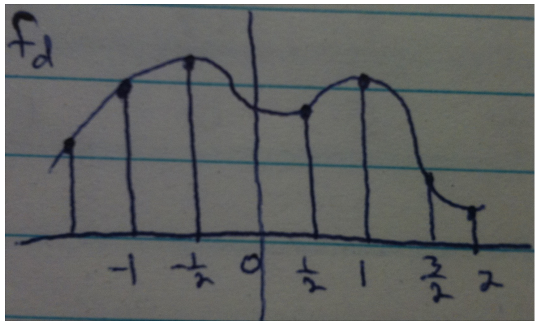
- Aliasing problem: f may not be exactly bandlimited.
- Requires a uniform sampling pattern.
- Truncation error due to the infinite sum.
- Perturbation in the samples, $f(n)$, can cause divergence of the series.

$\Rightarrow$ In practice, you never use the formula $f(x) = \sum_{n=-\infty}^{\infty} f(n)\,\text{sinc}(x-n)$.

The <u>Oversampling Theorem</u> requires that $\tau = \frac{1}{2\Omega}$ (or $\tau < \frac{1}{2\Omega}$). This is called the Nyquist rate.

Let $\Omega = \frac{1}{2}$, $\tau < 1 \Rightarrow$ "oversampling."





Let $\tau = \frac{1}{2} < 1$.

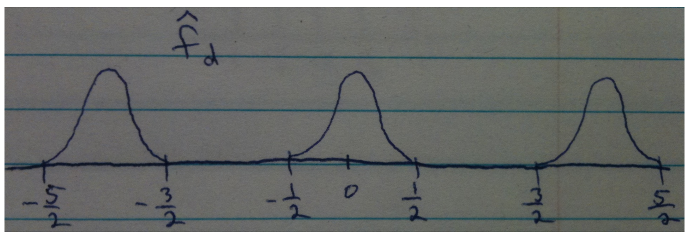
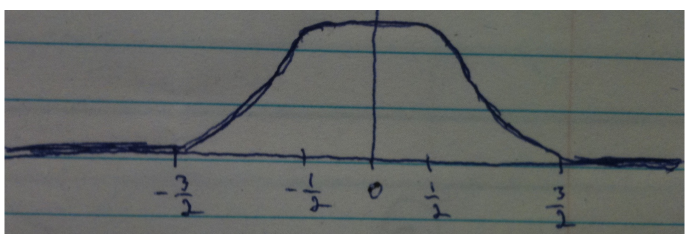


Figure 1: Period is $\frac{1}{\tau} = 2$.



sinc function decay in time: $\mathcal{O}\left(\frac{1}{x}\right)$ causes slow convergence. Instead, we can use a function $g(x)$ such that

$$\hat{g}(\omega) = \begin{cases} 1 & |\omega| < \frac{1}{2} \\ 0 & |\omega| \geq \frac{3}{2} \\ \text{anything} & \text{otherwise} \end{cases}$$

Idea: choose it such that $g(x)$ decays fast.

Definition 3.2. ***Raised Cosine***

$$\hat{g}(\omega) = \begin{cases} 1 & |\omega| < \frac{1}{2} \\ 0 & |\omega| \geq \frac{3}{2} \\ \frac{1}{2}\left(1 + \cos\left(\pi\omega - \frac{\pi}{2}\right)\right) & \frac{1}{2} \leq \omega \leq \frac{3}{2} \\ \frac{1}{2}\left(1 + \cos\left(\pi\omega + \frac{\pi}{2}\right)\right) & -\frac{1}{2} \geq \omega \geq -\frac{3}{2} \end{cases}$$

(See the plot above.) This gives cubic decay of $g(x)$: $\mathcal{O}\left(\frac{1}{x^3}\right)$. (However, we are hiding a constant which does come into play in practice.

This *requires* oversampling because of the transition phase.

Observations/Notes:

- We can design $\hat{g}(\omega)$ such that it is infinitely often differentiable.
- The best possible decay rate for $g(x)$ is of order e^{x^α}, $\alpha < 1$; $\alpha = 1$ is not possible. Typically, we call this subexponential decay. True exponential decay is not possible. (Berding, Molliovin 1962)

Form of uncertainty principle: if $\hat{g}(\omega)$ is compactly supported, then $g(x)$ cannot be compactly supported (and vice versa).

Note: a signal that is limited in time is unlimited in bandwidth (e.g. a phone call).

Proposition 3.3.

If

$$h_\tau(x) = \tau \frac{\sin \frac{\pi x}{\tau}}{\pi x},$$

then $\{T_{n\tau} h_\tau\}_{n\in\mathbb{Z}}$ is an orthonormal basis for B_Ω if $\Omega = \frac{1}{2\tau}$.

Proof. Since $\hat{h}_\tau = \tau \mathbf{1}_{[-\frac{1}{2},\frac{1}{2}]}$, Plancherel implies

$$\begin{aligned} \langle T_{n\tau} h_\tau, T_{m\tau} h_\tau \rangle &= \tau^2 \int_{-\infty}^{\infty} \mathbf{1}_{[-\frac{1}{2},\frac{1}{2}]} e^{-2\pi i (n-m)\tau\omega}\, d\omega \\ &= \tau^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-2\pi i (n-m)\tau\omega}\, d\omega \\ &= \begin{cases} 0 & n \neq m \\ 1 & n = m \end{cases} \end{aligned}$$

So we have proved orthonormality, now we need to prove that it is a basis. Shannon's Sampling Theorem says that any $f \in B_{1/2\tau}$ can be written as a linear combination of $\{T_{n\tau} h_\tau\}_{n\in\mathbb{Z}}$. Thus, $\{T_{n\tau} h_\tau\}_{n\in\mathbb{Z}}$ are a basis. □

Let $\tau = 1$, $\Omega = \frac{1}{2} \Rightarrow \{T_n \operatorname{sinc}\}_{n\in\mathbb{Z}}$ is an orthonormal basis for $B_{1/2}$. Now let $\tau = \frac{1}{2}$, $\Omega = \frac{1}{2}$ (oversampling): $\{T_{n/2} \operatorname{sinc}\}_{n\in\mathbb{Z}}$ is not an orthonormal basis because it has too many functions. This gives rise to "oversampled Shannon":

$$f(x) = \frac{1}{2} \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2}\right) \operatorname{sinc}\left(x - \frac{n}{2}\right).$$

3.3 Frames

Frames are an "overcomplete basis." They preserve almost all of the good properties of bases, and give us some additional properties by virtue of having more functions than are needed.

Frame Theory: Duffin, Schoeffer 1952, in connection with non-uniformly sampled bandlimited sets.

Definition 3.4. ***Orthonormal Basis***

$\{e_k\}_{k\in I}$ is an *orthonormal basis* for a separable Hilbert space $\mathcal{H}$ if

1. $\langle e_k, e_l\rangle = \delta_{kl}$
2. $\{e_k\}_{k\in I}$ spans $\mathcal{H}$, where I is a countable index set (e.g. $I = \mathbb{Z}$)

Properties of Orthonormal Bases:

1. $f = \sum\limits_{k\in I} \langle f, e_k\rangle\, e_k$ if $f \in \mathcal{H}$
2. These coefficients, $\langle f, e_k\rangle$, are unique
3. $\sum\limits_{k} |\langle f, e_k\rangle|^2 = \|f\|_2^2$ (Plancherel)

Definition 3.5. ***Frame, Frame Bounds, Tight Frame***

The sequence of vectors $\{g_k\}_{k\in I}$ is called a *frame* for the Hilbert space $\mathcal{H}$ if there exist constants $A, B > 0$ such that for any $f \in \mathcal{H}$,

$$A\|f\|_2^2 \leq \sum |\langle f, g_k\rangle|^2 \leq B\|f\|_2^2$$

(Typically we want A and B to be similar in size.) A and B are called *frame bounds*. If $A = B$, then $\{g_k\}_{k\in I}$ is called a *tight frame*:

$$\sum_k |\langle f, g_k\rangle|^2 = A\|f\|_2^2$$

If $A = B = 1$ and $\|g_k\| = 1$ for all k, then $\{g_k\}$ is an orthonormal basis.

4 1-24-12

4.1 Frames (Continued)

There exist $A, B > 0$ such that

$$A\|f\|_2^2 \leq \sum |\langle f, g_k\rangle|^2 \leq B\|f\|_2^2 \qquad \text{for all } f \in H_1$$

Motivated by Plancherel:

$$\sum |\langle f, g_k\rangle|^2 = \|f\|_2^2$$

4.2 Two Fundamental Operators

Definition 4.1. ***Synthesis Operator***

$F : \ell^2(I) \to H$,

$$Fc = \sum_{k\in I} c_k g_k \qquad \text{for } c = \{c_k\}_{k\in I} \in \ell^2(I)$$

This is a generalization of the idea of a Fourier series.

Definition 4.2. ***Analysis Operator***

F^*, the adjoint of the synthesis operator (proven next).
$F^* : H \to \ell^2(I)$,

$$F^* f = \{\langle f, g_k\rangle\}_{k\in I}, \qquad f \in H$$

The analysis operator is indeed the adjoint of the synthesis operator.

Proof.

$$\begin{aligned}\underbrace{\langle Fc, f\rangle}_{=\langle c, F^* f\rangle} &= \left\langle \sum c_k g_k, f\right\rangle \\ &= \sum c_k \langle g_k, f\rangle \\ &= \left\langle c, \underbrace{\{\langle g_k, f\rangle\}}_{=F^* f}\right\rangle\end{aligned}$$

□

Definition 4.3. ***Frame Operator***

The *frame operator* is defined as

$$Sf = \sum \langle f, g_k\rangle g_k$$

Note: $S = FF^*$.

Given two self-adjoint operators, P and Q, we denote $P \geq Q$ if

$$\langle f, Pf \rangle \geq \langle f, Qf \rangle \qquad \text{for all } f \in H.$$

S satisfies

$$AI \leq S \leq BI,$$

where I is the identity on H. Since

$$\begin{aligned} Sf = \sum_k \langle f, g_k \rangle g_k \quad \Rightarrow \quad \langle Sf, f \rangle &= \sum \langle f, g_k \rangle \langle g_k, f \rangle \\ &= \sum \langle f, g_k \rangle \overline{\langle f, g_k \rangle} \\ &= \sum | \langle f, g_k \rangle |^2 \end{aligned}$$

Since $A > 0$, S is positive definite on H.

Lower frame bound: $AI \leq S$ implies invertibility of S.
Upper frame bound: $S \leq BI$ implies continuity.

Definition 4.4. ***Dual Frame***

Define $h_k = S^{-1} g_k$ for all $k \in I$. We call $\{h_k\}_{k \in I}$ the *(canonical) dual frame*. $\{h_k\}_{k \in I}$ is a frame for H with frame bounds $\frac{1}{B}$, $\frac{1}{A}$.

From $SS^{-1} = S^{-1}S = I$, we get

$$\begin{aligned} f &= SS^{-1} f \\ &= \sum \left\langle S^{-1} f, g_k \right\rangle g_k \\ &= \sum \left\langle f, S^{-1} g_k \right\rangle g_k \\ &= \sum \left\langle f, h_k \right\rangle g_k \end{aligned}$$

$$\begin{aligned} f &= S^{-1} S f \\ &= \sum \langle Sf, h_k \rangle h_k \\ &= \sum \langle f, Sh_k \rangle h_k \\ &= \sum \langle f, g_k \rangle h_k \end{aligned}$$

$$\begin{aligned} f &= \sum \langle f, h_k \rangle g_k \\ f &= \sum \langle f, g_k \rangle h_k \end{aligned}$$

If $\{g_k\}$ is a frame for H, then any $f \in H$ can be written as

$$f = \sum_k c_k g_k, \qquad \{c_k\} \in \ell^2(I)$$

But $\{c_k\}$ is not unique (unlike in an orthonormal basis). One choice for these coefficients is $c_k = \langle f, h_k \rangle$; these are called the *canonical coefficients*.

Proposition 4.5.

Let $\{g_k\}_{k\in I}$ be a frame for H and $f = \sum c_k g_k$ for $c = \{c_k\} \in \ell^2(I)$. Then

$$\sum_k |c_k|^2 \geq \sum_k |\langle f, h_k\rangle|^2$$

where $h_k = S^{-1} g_k$.

Proof. Set $a_k = \langle f, h_k\rangle$. Then $f = \sum a_k g_k$, and

$$\begin{aligned}
\left\langle f, S^{-1} f\right\rangle &= \sum a_k \left\langle g_k, S^{-1} f\right\rangle = \sum a_k \underbrace{\left\langle S^{-1} g_k, f\right\rangle}_{\overline{a}_k} = \sum |a_k|^2 \\
\left\langle f, S^{-1} f\right\rangle &= \sum c_k \left\langle g_k, S^{-1} f\right\rangle = \sum c_k \overline{a}_k = \langle c, a\rangle \\
\|a\|_2^2 &= \langle c, a\rangle
\end{aligned}$$

Consider

$$\begin{aligned}
\|c\|_2^2 &= \|c - a + a\|_2^2 = \|c - a\|_2^2 + \|a\|_2^2 + \langle c - a, a\rangle + \langle a, c - a\rangle \\
&= \|c - a\|_2^2 + \|a\|_2^2 \qquad \text{(since } \langle c, a\rangle = \langle a, a\rangle) \\
&\geq \|a\|_2^2
\end{aligned}$$

□

In optimization language:
Consider

$$\min \|c_2\| \quad \text{such that} \quad Fc = f \qquad (Fc = \sum c_k g_k)$$

The solution is $c_k = \langle f, h_k\rangle$.

4.3 Tight Frames

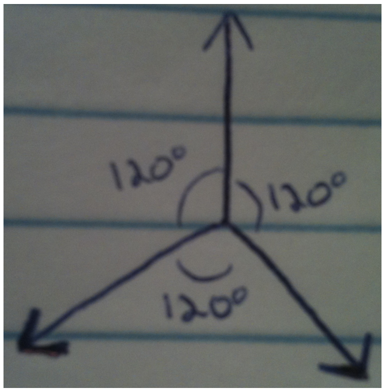
Definition 4.6. ***Tight Frame***

$\{g_k\}$ is called a *tight frame* if $A = B$.

$$AI = S = BI$$

$$\begin{aligned}
S^{-1} &= \frac{1}{A} I \\
h_k &= \frac{1}{A} g_k \\
f &= \frac{1}{A} \sum \langle f, g_k\rangle g_k
\end{aligned}$$

Example 4.7. ***Tight Frames in*** $\mathbb{R}^2$



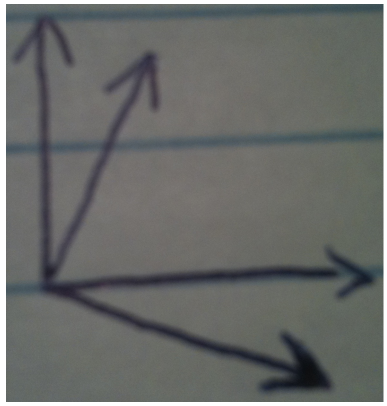


Figure 2: Rotated orthonormal basis.

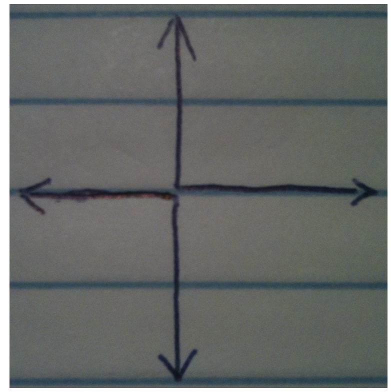


Figure 3: An orthonormal basis included twice.

Example 4.8. ***Oversampling of Bandlimited Function***

Let $\Omega = \frac{1}{2}$.

$\tau = 1 \;\Rightarrow\; \{T_{k\tau}\,\mathrm{sinc}\}_{k\in\mathbb{Z}}$ are ONB for $B_{1/2}$
$\tau = \frac{1}{2} \;\Rightarrow\; \{T_{k/2}\,\mathrm{sinc}\}_{k\in\mathbb{Z}}$ are a tight frame for $B_{1/2}$ with bounds $A = B = 2$.

Definition 4.9. ***Frame Gram Matrix***

A *frame Gram matrix* $R : \ell^2(I) \to \ell^2(I)$ (although it could act only on a subspace of ℓ^2), where

$$R_{kl} = \langle g_l, g_k \rangle \,.$$

Let $c = \ell^2(I)$, then

$$Rc = F^* F c = \left\{ \left\langle \sum_k c_k g_k, g_l \right\rangle \right\}_{l\in I}$$

For an orthonormal basis, $R = I$.
For a frame, R is in general not invertible on $\ell^2(I)$.

When is R invertible?

Definition 4.10. ***Riesz Basis***

(There are many equivalent definitions for a Riesz basis.) $\{g_k\}_{k\in I}$ is a Riesz basis for H if any of the following are true:

1. There exists an orthonormal basis $\{e_k\}$ for H such that

$$Ue_k = g_k \qquad \text{for all } k \in I,\ U \text{ invertible.}$$

2. The Gram matrix R is invertible on $\ell^2(I)$.
3. The coefficients c_k in $f = \sum c_k g_k$ are unique and in $\ell^2(I)$.
4. $\{g_k\}$ is a frame for H, but it fails to be a frame for H if any of the g_k is removed.

5 1-26-12

5.1 ONB, Riesz Basis, and Frames

ONB $\leftrightarrow$ Riesz Basis, Frames $\leftrightarrow$ Tight Frames

Recall: $\{g_k\}_{k\in I}$ is a Riesz basis if

- the Gram matrix R is invertible
- every $f \in H$ can be written as $f = \sum c_k g_k$, where the c_k are unique

How do we find these c_k's?

Definition 5.1. ***Dual Riesz Basis***

For a Riesz basis $\{g_k\}_{k\in I}$, we define the *dual Riesz basis* as $\{h_k\}_{k\in I}$ as $h_k = S^{-1}g_k$.

Definition 5.2. ***Biorthogonal***

For a Riesz basis $\{g_k\}_{k\in I}$, $\{h_k\}_{k\in I}$ are *biorthogonal*:

$$\langle g_k, h_l\rangle = \delta_{kl}$$

Know: the spectrum of frame operator S satisfies $\sigma(S) \subseteq [A, B]$.
Recall:

$$\begin{aligned} S &= FF^* \\ R &= F^*F \\ \Rightarrow \qquad \sigma(R) &\subseteq \{0 \cup [A, B]\} \end{aligned}$$

If $\{g_k\}$ are a Riesz basis, then $\sigma(R) \subseteq [A, B]$.

$\{T_{k\tau}\,\text{sinc}\}_{k\in\mathbb{Z}}$ is a tight frame for $B_{1/2}$ if $\tau \le i$ with $A = B = \frac{1}{2}$. What if the sampling points x_k are not $x_k = k\tau$? When can we recover $f \in B_{1/2}$ from $\{f(x_k)\}_{k\in I}$?

We can recover f whenever $\{T_{x_k}\,\text{sinc}\}_{k\in\mathbb{Z}}$ forms a frame for $B_{1/2}$. Recall:

$$\begin{aligned} f(x_k) &= \langle f, T_{x_k}\,\text{sinc}\rangle \\ f &= \sum \langle f, g_k\rangle \end{aligned}$$

Theorem 5.3. ***Kodec $\frac{1}{4}$ Theorem (1962)***

If the set $\{x_k\}_{k\in\mathbb{Z}}$ satisfies $|x_k - k| < \frac{1}{4}$, then $\{T_{x_k}\,\text{sinc}\}$ is a Riesz basis for $B_{1/2}$. Therefore, we can recover f from $\{f(x_k)\}_{k\in\mathbb{Z}}$.

We need a dual Riesz basis $h_k = S^{-1}g_k$. We have $g_k = T_{x_k}\,\text{sinc}$. $\{g_k\}_{k\in\mathbb{Z}}$ is generated by one function (sinc).

If x_k are not uniformly spaced, then $\{h_k\}_{k\in\mathbb{Z}}$ is not of the form $h_k = T_{x_k}h$. We have to compute $h_k = S^{-1}T_{x_k}\,\text{sinc}$ for each k in order to obtain f:

$$f = \sum_{k\in\mathbb{Z}} f(x_k)h_k.$$

5.2 Finite Frames

Let $H = \mathbb{R}^n$ or $\mathbb{C}^n$.

Consider the set of functions $\{f_k\}_{k=1}^m,\ f_k \in \mathbb{R}^n$.
Collect $\{f_k\}$ in a matrix $F = \begin{bmatrix} f_1 & f_2 & \cdots & f_m \end{bmatrix} \in \mathbb{R}^{n\times m}$ (F is the synthesis operator, $Fc = \sum c_k f_k$).

Note: F typically has more columns than rows. It *cannot* have less columns than rows.

- For an ONB, F is a unitary $n \times n$ matrix: $F^*F = FF^* = I$.
- For a Riesz basis, F is an $n \times n$ invertible matrix, the columns of $(F^{-1})^*$ form the dual Riesz basis.
- For a frame, F is an $n \times m$ of rank n and $m \geq n$.

What are the frame bounds of F?
Let $\sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_n$ be the singular values of F (i.e. the square roots of the eigenvalues of F^*F). Then the lower frame bound is $A = \sigma_1^2$ and the upper frame bound is $B = \sigma_n^2$.

- If F is a tight frame, then $\sigma_1 = \sigma_2 = \cdots = \sigma_n$

From Linear Algebra: the condition number of a matrix is $\frac{\sigma_n}{\sigma_1}$.

Definition 5.4. ***Condition Number***

The *condition number* of a frame is

$$\kappa = \frac{\sigma_n}{\sigma_1} = \frac{B}{A}.$$

5.3 Frames for Coding

Assume we are given a vector $x \in \mathbb{C}^n$ and we want to transmit x from one place to the other. We know with a certain probability up to $\frac{1}{3}$ of the coefficients get lost during transmission. We could re-transmit in case of data loss.

Or we could introduce redundancy: choose a frame $\{f_k\}_{k=1}^m$ and compute $c_k = \langle x, f_k\rangle$ and transmit the c_k's. If $\frac{1}{3}$ of the c_k's are lost, can we still recover x from the remaining c_k's? We need at least $m \geq \frac{3}{2}n$, is this enough?

Example 5.5.

Let $x \in \mathbb{R}^2$. Choose $m \geq \frac{3}{2} \cdot 2 = 3$ elements. Set $F = \begin{bmatrix} f_1 & f_2 & f_3 \end{bmatrix}$. Choose

$$F = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Compute $c_k = \langle x, f_k \rangle$, $k = 1, 2, 3$. If we lose c_1 (or c_3), we need to recover x from c_2, c_3 using f_2, f_3. $\tilde{F} = \begin{bmatrix} f_2 & f_3 \end{bmatrix}$ still spans $\mathbb{R}^2$, so we compute the dual frame to $\tilde{F}$ and get

$$x = c_2 h_2 + c_3 h_3,$$

where $\begin{bmatrix} h_2 & h_3 \end{bmatrix}$ are the dual frame to $\tilde{F}$.

If c_2 is lost, then $\tilde{F} = \begin{bmatrix} f_1 & f_3 \end{bmatrix}$, but $\tilde{F}$ is not a frame for $\mathbb{R}^2$. $\Rightarrow$ cannot recover x from c_1, c_3.

Example 5.6.

Choose instead

$$F = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}$$

We can recover x from any two out of $\{c_k\}_{k=1}^3$.

Example 5.7.

Choose

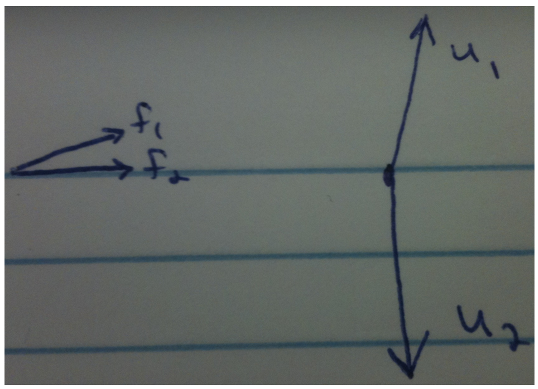
$$F = \begin{bmatrix} 1 & 0.9 & 0 \\ 0 & 0.1 & 1 \end{bmatrix}$$

Assume we lose $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Then

$$\tilde{F} = \begin{bmatrix} 1 & 0.9 \\ 0 & 0.1 \end{bmatrix}$$

The dual frame is

$$\begin{bmatrix} 1 & 0 \\ 9 & -10 \end{bmatrix}$$



The condition number of $\tilde{F}$ is large.

We want redundancy, but we also want that the frame elements are not "too close." $\Rightarrow |\langle f_k, f_l \rangle|$ should be small.

Definition 5.8. ***Coherence***

The *coherence* of a frame $\{f_k\}_{k=1}^m$ is defined as

$$\mu(\{f_k\}_{k=1}^m) = \max_{k \neq l} \frac{|\langle f_k, f_l \rangle|}{\|f_k\| \|f_l\|}.$$

We often consider normalized frames, i.e. $\|f_k\|_2 = 1$ for all k. Then

$$\mu = \max_{k \neq l} |\langle f_k, f_l \rangle| = \max_{k \neq l} |R_{k,l}|$$

So we want frames that

- have small μ
- are normalized
- are tight.

Example 5.9.

If $\{f_k\}$ is an ONB, then $\mu = 0$.
In general, $0 \le \mu \le 1$.
How small can μ be for a given m, n?

Theorem 5.10.

Let $\{f_k\}_{k=1}^m$ be a frame for $\mathbb{R}^n$ or $\mathbb{C}^n$ with $\|f_k\| = 1$ for all k. Then

$$\mu \ge \sqrt{\frac{m-n}{n(m-1)}} \tag{5.1}$$

Equality holds in (5.1) if an only if $\{f_k\}_{k=1}^m$ is an equiangular tight frame.

Definition 5.11. ***Equiangular***

Equiangular means that

$$|\langle f_k, f_l \rangle| = \alpha \quad \text{for } k \ne l$$

6 1-31-12

6.1 A Theorem on Tight Frames

Theorem 6.1.

Let $\{f_k\}_{k=1}^m$ be a frame for $\mathbb{C}^n$ or $\mathbb{R}^n$ with $\|f_k\| = 1,\ k = 1, \ldots, m$. Then

$$\mu \geq \sqrt{\frac{m-n}{n(m-1)}} \tag{6.1}$$

Equality holds if and only if $\{f_k\}_{k=1}^m$ is an equiangular tight frame (ETF).

$$\mu = \max_{k \neq l} |\langle f_k, f_l \rangle|$$
$$|\langle f_k, f_l \rangle| = \alpha$$

Proof. **Idea:** Look at the Gram matrix, $R = [\langle f_k, f_l \rangle]_{k,l=1}^m$. Recall: $\sigma(R) = \{0 \cup [A, B]\}$. Let $\{\lambda_k\}_{k=1}^n$ be the eigenvalues of R that are not zero. Recall:

$$\sum_{k=1}^n \lambda_k = \text{trace}(R) = \sum_{k=1}^m R_{kk} = m = \sum_{k=1}^m \|f_k\|_2^2.$$

Then

$$m^2 = (\text{trace}\, R)^2 = \left(\sum_{k=1}^m \lambda_k\right)^2 = \langle \mathbf{1}, \{\lambda_k\}\rangle^2$$
$$\leq \|\mathbf{1}\|_2^2 \|\{\lambda_k\}\|_2^2 = n \sum_{k=1}^n \lambda_k^2 = n\, \text{trace}(R^2)$$

Recall that

$$\sum_{k=1}^n \lambda_k^2 = \|R\|_F^2$$
$$\|A\|_F \equiv \sqrt{\sum_{k,l=1}^2 |A_{kl}|^2}$$

Then

$$\sum_{k=1}^n \lambda_k^2 = \sum_{k,l=1}^n |\langle f_k, f_l \rangle| \geq \frac{m^2}{n}. \tag{6.2}$$

We have

$$\mu^2 = \max_{k \neq l} |\langle f_k, f_l \rangle| \geq \frac{1}{m^2 - m} \sum_{\substack{k \neq l \\ k,l=1}}^{m} |\langle f_k, f_l \rangle|^2$$

$$= \frac{1}{m^2 - m} \left(\underbrace{\sum_{k,l=1}^{m} |\langle f_k, f_l \rangle|^2}_{\geq \frac{m^2}{n} \text{ by (6.2)}} - m \right)$$

$$\mu^2 \geq \frac{1}{m^2 - m} \cdot \left(\frac{m^2}{n} - m \right) = \frac{m - n}{n(m-1)}$$

which proves (6.1).

Equality in (6.1) implies that

$$|\langle f_k, f_l \rangle| = \frac{m - n}{n(m-1)} \qquad \text{for all } k \neq l$$

Equality in (6.1) also implies that $\lambda_1 = \lambda_2 = \ldots = \lambda_n = \frac{m}{n} \Leftrightarrow$ tight frame. □

Remark 6.2.

This bound cannot be achieved for all n, m.

Corollary 6.3.

Equality in (6.1) can only hold if

- $m \leq \frac{n(n+1)}{2}$ for $\mathbb{R}^n$.
- $m \leq n^2$ for $\mathbb{C}^n$.

Example 6.4. $\mathbb{C}^2$

Let

$$T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad M = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad v = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3 + \sqrt{3}} \\ e^{i\pi/4}\sqrt{3 - \sqrt{3}} \end{bmatrix}.$$

Then $F = [v,\ Tv,\ Mv,\ TMv]$ has $\mu = \frac{1}{\sqrt{3}}$ and frame bounds $A = B = 2$.

Remark 6.5. ***CONJECTURE (Zanner)***

For each n there exists a vector $f \in \mathbb{C}^n$ such that $\{f_{kl}\}_{k,l=1}^n$ is an ETF for $\mathbb{C}^n$, where $f_{kl} = T^k M^l f,\ k,l = 0,\dots,n-1$. Here

$$T = \begin{bmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 1 & & & & 0 \end{bmatrix}, \qquad M = \begin{bmatrix} \omega^0 & & & 0 \\ & \omega^1 & & \\ & & \ddots & \\ 0 & & & \omega^{n-1} \end{bmatrix}, \quad \omega = e^{2\pi i/n}$$

This has been proven for certain n, and verified numerically for $n = 2,\dots,41$.

This is important in Quantum Physics, where it is known as SIC-POVM: "Symmetric Informationally Complete Positive Operator Valued Measure."

A quantum state in finite dimensions can be identified with a positive semidefinite Hermitian matrix. Say $X \in \mathbb{C}^{n\times n}$ and X is hpd (hermitian positive semidefinit). We want to measure X. This means that we take measurements of the form $\langle X, F_k\rangle := \text{trace}(XF_k)$, $k = 1,\dots,m$. The F_k's are hpd and are often chosen to be rank-one (i.e. pure quantum state). If F_k is rank one, then $F_k = f_k f_k^*$. Given $\langle X, F_k\rangle$, can I recover X? (X does not have to be rank one.) If X is rank n,then we need at least $m = n^2$ measurements. In practice, we often want additional properties for the F_k.

Any ETF $\{f_{kl}\}_{k,l=1}^m$ gives a SIC-POVM $\{F_{kl}\}_{k,l=1}^n$, where $F_{kl} = f_{kl} f_{kl}^*$.

6.2 Mutually Unbiased Bases

Assume we have two orthonormal bases $U = [u_1,\ \dots,\ u_n]$, $V = [v_1,\ \dots,\ v_n]$ in $\mathbb{C}^n$. Let $F = [U\ V]$. How small can $\mu(F)$ be?

Example 6.6.

Let $U = I$, $V =$ Discrete-Fourier Transform matrix: $V_{kl} = \frac{1}{\sqrt{n}} e^{2\pi ikl/n},\ k,l = 0,\dots,n-1$. We can show $\mu = \frac{1}{\sqrt{n}}$.

$$R = F^*F = \begin{bmatrix} U^* \\ V^* \end{bmatrix} \begin{bmatrix} U & V \end{bmatrix} = \begin{bmatrix} I & DFT \\ DFT^* & I \end{bmatrix}$$

$$\Rightarrow \qquad \max_{k\neq l} |R_{kl}| = \frac{1}{\sqrt{n}}$$

$\frac{1}{\sqrt{n}}$ is the smallest coherence between two orthonormal bases. Call such U, V *mutually unbiased bases* (MUB).

[Quantum Physics, Schwinger 1965]

Theorem 6.7.

In $\mathbb{C}^n$, we can have at most $n+1$ MUB's (this means that the frame $F = [U_1,\ \ldots,\ U_{n+1}]$ has $\mu = \frac{1}{\sqrt{n}}$.)

In $\mathbb{R}^n$, it is at most $\left\lfloor \frac{n}{2} \right\rfloor + 1$.

Example 6.8.

Consider $\mathbb{C}^n$, where n is a prime ≥ 5. Let $f \in \mathbb{C}^n$ be defined by

$$f(k) = \frac{1}{\sqrt{n}} e^{2\pi i k^3/n}, \qquad k = 0, \ldots, n-1.$$

Let

$$f_{j,l} = T^j M^l f, \qquad j, l = 0, \ldots, n-1.$$

Then $\{f_{jl}\}_{l=0}^{n-1}$ is an orthonormal basis $\Rightarrow$ call it U_j.

$$\mu\left([U_j\ U_i]\right) = \frac{1}{\sqrt{n}}, \qquad j, i = 0, \ldots, n-1.$$

Choose $U_n = I$.

7 2-2-12

7.1 DFT and FFT

Let $x \in \mathbb{C}^n$. We want to compute the Fourier transform $\hat{x}$.

Definition 7.1. ***Discrete Fourier Transform***

The *Discrete Fourier Transform (DFT)*:

$$\hat{x}(k) = \frac{1}{\sqrt{n}} \sum_{l=0}^{n-1} x(l) e^{-2\pi i k l/n}, \qquad k = 0, 1, \ldots, n-1$$

The matrix

$$F = \begin{bmatrix} \omega_n^{0\cdot 0} & \omega_n^{0\cdot 1} & \ldots & \omega_n^{0\cdot(n-1)} \\ \omega_n^{1\cdot 0} & \omega_n^{1\cdot 1} & \ldots & \omega_n^{1\cdot(n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_n^{(n-1)\cdot 0} & \omega_n^{(n-1)\cdot 1} & \ldots & \omega_n^{(n-1)\cdot(n-1)} \end{bmatrix}$$

where $\omega = e^{2\pi i/n}$ is the "$n$the root of unity." F is unitary: $F^* = F^{-1}$.

A naive implementation of $\hat{x}$ takes $O(n^2)$ operations.

7.1.1 Fast Fourier Transform (FFT)

Cooley-Tuckey (1965), Gauss (1805)

Consider

$$\hat{x}(k) = \sum_{l=0}^{n-1} x(l) e^{-2\pi i k/n}, \qquad k = 0, 1, \ldots, n-1.$$

When the frequency index is even, we group the terms with index l and $l + \frac{n}{2}$:

$$\hat{x}(2k) = \sum_{l=0}^{\frac{n}{2}-1} \left(x(l) + x\left(l + \frac{n}{2}\right)\right) e^{\overbrace{-2\pi i k l/\frac{n}{2}}^{=-\pi i k l/n}}.$$

When the frequency index is odd, we get

$$\hat{x}(2k+1) = \sum_{l=0}^{\frac{n}{2}-1} e^{-2\pi i l/n} \left(x(l) - x\left(l + \frac{n}{2}\right)\right) e^{-2\pi i k l/\frac{n}{2}}$$

Key observation:
Even frequencies can be computed by computing the DFT of a $\frac{n}{2}$-length signal.

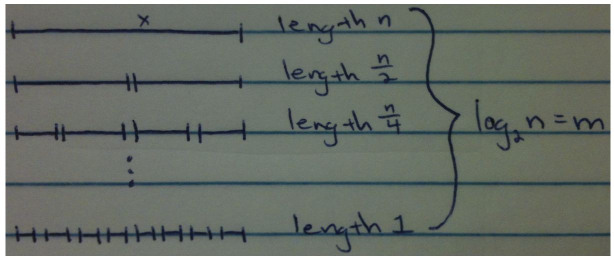
$$x_{\text{even}}(l) = x(l) + x\left(l + \frac{n}{2}\right), \qquad l = 0, \ldots, \frac{n}{2} - 1.$$

Odd frequencies can be computed by the DFT of a $\frac{n}{2}$-length signal.

$$x_{\text{odd}}(l) = e^{-2\pi i l/n} \left(x(l) - x\left(l + \frac{n}{2}\right)\right)$$

$\Rightarrow$ A DFT of length n can be computed by 2 DFT's of length $\frac{n}{2}$, plus n operations (for forming x_{even} and x_{odd}).

Let $n = 2^m$.



Splittings can be organized such that the total effort for computing $\hat{x}$ is $O(n \log n)$ operations ($= 3n \log n$).
$\Rightarrow$ FFT

We can extend this idea to signals of length $3^m, 5^m, 2^{m_1} \cdot 3^{m_2}, \ldots$. This is still fast as long as the prime factors are small.

In Matlab: `FFT`, `IFFT` (inverse)

7.1.2 Matrix Interpretation of FFT

Observation: We can write the DFT matrix F_n as

$$F_n = P_n \begin{bmatrix} F_{n/2} & \mathbf{0} \\ \mathbf{0} & F_{n/2} \end{bmatrix} Q_n$$

where Q_n is a permutation matrix ($\Rightarrow$ doesn't count towards operations); it is called the perfect shuffle permutation.

$$P_n = \begin{bmatrix} I_{n/2} & D_{n/2} \\ I_{n/2} & -D_{n/2} \end{bmatrix}$$

where $I_{n/2}$ is the identity matrix and $D_{n/2}$ is a diagonal matrix with powers of ω in its diagonal.

Example 7.2. ***FFT vs. DFT***

Say $n = 8192 = 2^{13}$. FFT is about 1000 times faster than DFT.

7.1.3 (Fast) Convolution of Finite Signals

Let $f, g \in \mathbb{C}^n$. Convolution:

$$(f * g)(k) = \sum_{l=?}^{?} f(l) g(k-l), \qquad k = 0, \ldots, n-1$$

If $k = 0, l = 1$, we need $g(-1)$. We need to define $g(k-l)$ for $k - l < 0$ or $k - l > n - 1$. What do we do at the signal boundaries?

<u>Two standard cases:</u>

1. "Zero padding:" Define

$$\left.\begin{array}{l} g(k)=0 \\ f(k)=0 \end{array}\right\} \quad \text{for } k<0,\ k\geq n$$

   Drawbacks:

   (a) If f, g have length n, then $f * g$ has $2n-1$ nonzero entries.

   (b) It can introduce an artificial "jump."

2. Periodic extension:

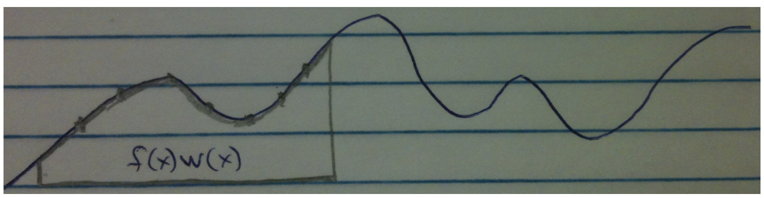
$$g(k)=g(k \bmod n)$$

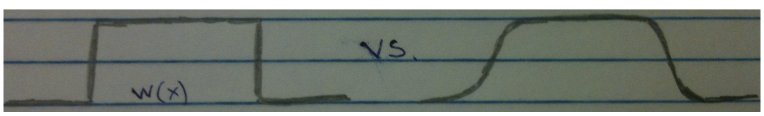
   If f and g are n-periodic, then $f * g$ is also n-periodic. Also, the algebraic properties of the Fourier Transform are preserved.

$$(f * g)=F^{-1}(\hat{f}\cdot\hat{g})$$

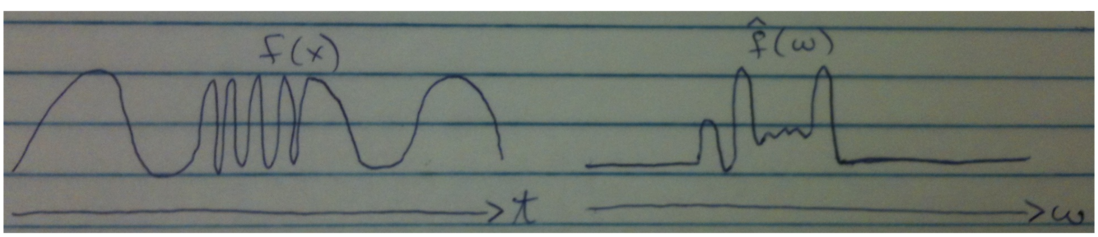
   In Matlab: `fft(fft(f)×fft(g))`

Let $f, g \in \mathbb{C}^n$.





Why do we take a Fourier Transform?
To understand the frequency behavior of $f(x)$.



Information about <u>when</u> frequencies dominate in f is hidden in $\hat{f}(\omega)$. For example, musical score is a joint time-frequency representation of music. Playing music $\leftrightarrow$ synthesis operator, $f=\sum_{k,l} c_{k,l} f_{k,l}$. Writing down musical score $\leftrightarrow$ analysis operator $\langle f, f_{k,l}\rangle = c_{k,l}$.

8 2-7-12

8.1 Uncertainty Principle

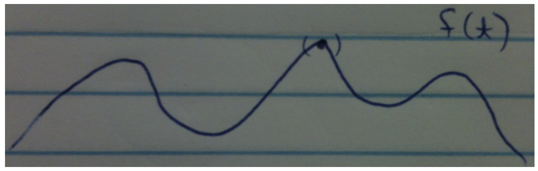
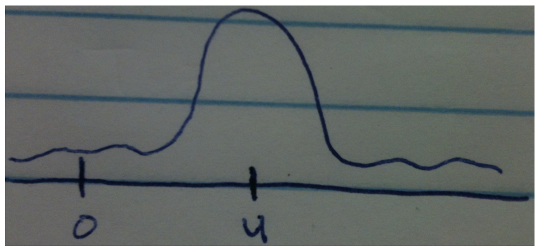


Figure 4: The idea of "instantaneous frequency."

We need some notion of "localization" or "concentration of energy."



f is concentrated at a point u if

$$\int_{\mathbb{R}} |f(x)|^2 (x-u)^2 \, dx$$

is small compared to $\|f\|_2^2$. Similarly, $\hat{f}(\omega)$ is concentrated at ξ if

$$\int_{\omega} |\hat{f}(\omega)|^2 (\omega-\xi)^2 \, d\omega$$

is small relative to $\|\hat{f}\|_2^2$.

Denote

$$\sigma_x^2 = \frac{1}{\|f\|_2^2} \int_{\mathbb{R}} (x-u)^2 |f(x)|^2 \, dx$$
$$\sigma_\omega^2 = \frac{1}{\|\hat{f}\|_2^2} \int_{\mathbb{R}} (\omega-\xi)^2 |\hat{f}(\omega)|^2 \, d\omega$$

Theorem 8.1. ***Heisenberg-Pauli-Weyl Inequality***

(Heisenberg 1927, Wiener 1925)

Let $f \in L^2(\mathbb{R})$. Then

$$\sigma_x \sigma_\omega \geq \frac{1}{4\pi}$$

with equality if and only if

$$f(x) = a e^{2\pi i \xi x} e^{-b(x-u)^2}, \qquad a, b, \xi \in \mathbb{R},\ a, b > 0$$

Proof. We assume that $\lim_{x\to\infty} \sqrt{x} f(x) = 0$ and $\|f\|_2 = 1$ (although the theorem holds for all $f \in L^2(\mathbb{R})$). If $f(x)$ is localized around u and $\hat{f}(\omega)$ is localized around ξ, then $f(x+u)e^{-2\pi i x \xi}$ is localized around $(0,0)$.

Thus, it is sufficient to consider $u = 0$, $\xi = 0$ (otherwise make a change of variables).

$$\sigma_x^2 \sigma_\omega^2 = \left(\int_{\mathbb{R}} |xf(x)|^2 \, dx \right) \left(\int_{\mathbb{R}} |\omega \hat{f}(\omega)|^2 \, d\omega \right)$$

Since $i\omega \hat{f}(\omega) = \frac{1}{2\pi} (\widehat{f'})(\omega)$, we get

$$\sigma_x^2 \sigma_\omega^2 = \frac{1}{4\pi^2} \int |xf(x)|^2 \, dx \int |f'(x)|^2 \, dx,$$

where we have used Plancherel for $\widehat{f'}$. Cauchy-Schwarz gives us

$$\begin{aligned} \sigma_x^2 \sigma_\omega^2 &\geq \frac{1}{4\pi^2} \left(\int |xf(x)f'(x)| \, dx \right) \\ &= \frac{1}{4\pi^2} \left(\int \frac{x}{2} |2f'(x)f(x)| \, dx \right)^2 \\ &= \frac{1}{16\pi^2} \left(\int x(|f(x)|^2)' \, dx \right)^2 \end{aligned}$$

Since $\sqrt{x} f(x) \to 0$, integration by parts gives

$$\begin{aligned} \sigma_x^2 \sigma_\omega^2 &\geq \frac{1}{16\pi^2} \left(\int |f(x)|^2 \, dx \right)^2 \\ &= \frac{1}{16\pi^2} \|f\|_2^2 \\ &= \frac{1}{16\pi^2} \end{aligned}$$

Equality: The only inequality step is when we use Cauchy-Schwarz. Cauchy-Schwarz is an equality if the two functions are linearly independent:

$$f'(x) = cxf(x) \quad \Rightarrow \quad f(x) = ae^{-bx^2}$$

gives $\sigma_x \sigma_\omega = \frac{1}{4\pi}$, $a, b > 0$. If $n \neq 0$, $\xi \neq 0$, then we get equality if

$$f(x) = ae^{2\pi i \xi x} e^{-b(x-u)^2}.$$

□

8.2 Uncertainty Principle in Quantum Mechanics

Classical mechanics: the state of a system is completely determined by the position x and momentum ω of the particle.

Quantum mechanics:

- Observables for position: multiplication operator $(Xf)(x) = xf(x)$
- Observables for momentum: momentum operator $Pf = \frac{1}{2\pi i} \frac{df}{dx}$ (where we have ignored Planck's constant)

If a particle is in the state $f \in L^2(\mathbb{R})$, $\|f\|_2 = 1$, then the expected position is

$$\langle Xf, f \rangle = \int x|f(x)|^2 \, dx$$

where $|f(x)|^2$ is the probability density for random variable x. Position uncertainty is the standard deviation of x:

$$\sigma_x = \left(\int (x-q)^2 |f(x)|^2 \, dx \right)^{1/2}$$

The particle is "most" likely located in the interval $[q - \sigma_x x, q + \sigma_x x]$, where σ_x depends on f!

Expected momentum is $\langle Pf, f \rangle = \left\langle X\hat{f}, \hat{f} \right\rangle = \int \omega |\hat{f}(\omega)|^2 \, d\omega$, since $P = \mathcal{F}^{-1} X \mathcal{F}$ and by Parseval. The momentum uncertainty is

$$\sigma_\omega = \langle (P-p)f, f \rangle^{1/2} = \left(\int (\omega - p)^2 |\hat{f}(\omega)|^2 \, d\omega \right)^{1/2}$$

$$\sigma_x \sigma_\omega \geq \frac{1}{4\pi}$$

(x, ω) can be considered as a point in $\mathbb{R} \times \hat{\mathbb{R}}$, which is called *phase space* or the *time-frequency plane*. (Recall: $\mathbb{R} \leftrightarrow \hat{\mathbb{R}}$, $\mathbb{Z} \leftrightarrow \mathbb{T}$.)

We cannot assign a unique point (x, ω) in phase space to a particle.

Definition 8.2. ***Commutator***

For two operators A, B, we define the *commutator*

$$[A, B] := AB - BA.$$

If the operators commute, the commutator is zero.

$$[X, P] = XP - PX = -\frac{1}{2\pi i} I \tag{8.1}$$

$\Rightarrow$ X and P do not commute. We could derive the uncertainty principle from (8.1).

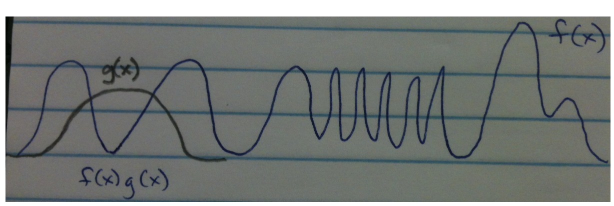
Note:

$$\left. \frac{d}{d\omega} M_\omega f \right|_{\omega=0} = 2\pi i X f$$

$$\left. \frac{d}{dx} T_x f \right|_{x=0} = -2\pi i P f$$

8.3 Short-Time Fourier Transform (STFT)

Given a function $f \in L^2(\mathbb{R})$, we want a "local" time-frequency representation of f.



Definition 8.3. ***Short-Time Fourier Transform (STFT)***

Fix a function $g \neq 0$ (the window, atom, etc.). The *short-time Fourier transform* of a function f w.r.t. g is

$$(V_g f)(x, \omega) = \int_{\mathbb{R}} f(t)\overline{g(t-x)}e^{-2\pi i \omega t}\, dt$$

for $x, \omega \in \mathbb{R} \times \mathbb{R}$. (The conjugate will allow us to interpret this as an inner product.)

V_g maps f on $\mathbb{R}$ into a function on $\mathbb{R} \times \mathbb{R}$ with time and frequency as coordinates (joint representation). The properties of $V_g f$ depend crucially on the choice of g.

9 2-9-12

Next Tuesday, 2/14: Office Hours are 12-1 (instead of 3-4)

9.1 Short-Time Fourier Transform (Continued)

Recall:

$$V_g f(x,\omega) = \int_{\mathbb{R}} f(t)\overline{g(t-x)}e^{-2\pi i\omega t}\,dt,$$

$(w,\omega) \in \mathbb{R}^2$.

Lemma 9.1.

If $f, g \in L^2(\mathbb{R})$, then

$$\begin{aligned}(V_g f)(x,\omega) &= \widehat{f \cdot T_x \overline{g}}(\omega) \\ &= \langle f, M_\omega T_x g\rangle \\ &= \left\langle \hat{f}, T_\omega M_{-x}\hat{g}\right\rangle \\ &= e^{-2\pi i x\omega} V_{\hat{g}}\hat{f}(\omega,-x)\end{aligned}$$

Example 9.2.

For $g(x) = f(x) = e^{-\pi x^2}$

$$\begin{aligned} \langle \phi, M_\omega T_x \phi\rangle &= e^{-\pi i x\omega}\phi(x)\underbrace{\hat{\phi}(\omega)}_{=\phi(\omega)} \\ \Rightarrow \qquad V_\phi\phi(w,\omega) &= e^{-\pi i x\omega}e^{-\pi(x^2+\omega^2)}\end{aligned}$$

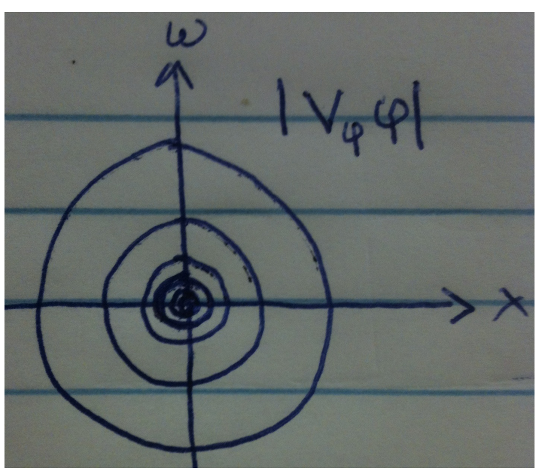


Figure 5: $V_\phi\phi$ is a 2-D Gaussian. It is focused at the origin and decays rapidly.

Example 9.3.

If $g = S,\ f \in \mathcal{S}$ (Schwartz class), then

$$V_g f(x,\omega) = e^{-2\pi i x\omega} f(x)$$

Example 9.4.

If $g = 1,\ f \in L^1$, then

$$V_g f(x,\omega) = \hat{f}(\omega)$$

9.2 STFT Orthogonality Relations

Let $f_1, f_2, g_1, g_2 \in L^2$. Then

$$\langle V_{g_1} f_1, V_{g_2} f_2 \rangle = \langle f_1, f_2 \rangle\, \overline{\langle g_1, g_2 \rangle}$$

$\Rightarrow$ This implies the inversion of the STFT:

Let $g, \gamma \in L^2(\mathbb{R})$ and $\langle g, \gamma \rangle \neq 0$. Then for all $f \in L^2(\mathbb{R})$:

$$f = \frac{1}{\langle g, \gamma \rangle} \int_{\mathbb{R}} \int_{\mathbb{R}} V_g f(x,\omega) M_\omega T_x \gamma \, d\omega \, dx$$

in the weak sense:

$$\langle f, h \rangle = \frac{1}{\langle g, \gamma \rangle} \int \int V_g f(x,\omega)\, \langle M_\omega T_x \gamma, h \rangle \; dx \, d\omega$$

- g: analysis window
- γ: synthesis window
  - Can choose $\gamma = g$

Remark 9.5. ***Isometry***

If $g \in L^2(\mathbb{R})$ and $\|g\|_2 = 1$, then $\|f\|_2 = \|V_g f\|_2$.

$\Rightarrow$ The STFT is an isometry from $L^2(\mathbb{R})$ into $L^2(\mathbb{R})$.

9.3 Quadratic Time-Frequency Representations

Definition 9.6. ***Spectrogram***

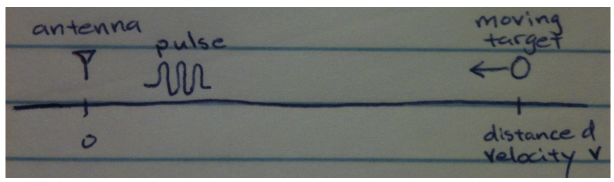
$|V_g f(x,\omega)|^2$ is known as a *spectrogram* (used in audio processing and music).

Definition 9.7. ***Ambiguity Function***

$$Af(x,\omega) = \int f\left(t+\frac{x}{2}\right)\overline{f\left(t-\frac{x}{2}\right)}e^{-2\pi it\omega}\,dt$$
$$= e^{\pi ix\omega}V_f f(x,\omega)$$

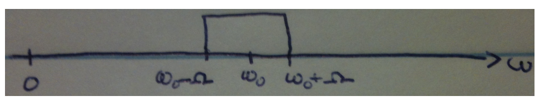
Af does not depend on a window function g.

9.3.1 Ambiguity Function and Radar



We want to determine d and v. We send a pulse $f(t) = g(t)e^{2\pi it\omega_0}$, where ω_0 is the carrier frequency and $g(t)$ is the envelope signal. If $\operatorname{supp}\hat{g} = [-\Omega,\Omega]$, then $\operatorname{supp}\hat{f} \approx [-\Omega+\omega_0, \Omega+\omega_0]$.

Narrowband assumption: $2\Omega \ll \omega_0$.



The antenna receives the echo with time delay $\Delta t = \frac{2d}{c}$, where c is the speed of light. The motion of the target induces *Doppler shift*: at each frequency $\omega \in [\omega_0-\Omega, \omega_0+\Omega]$ is shifted by $\Delta\omega = -\frac{2\pi iv\omega}{c} \approx -\frac{2\pi iv\omega_0}{c}$, since $\Omega \ll \omega_0$.

$\Rightarrow$ The received signal is
$$s(t) = \alpha f(t-\Delta t)e^{2\pi it\Delta\omega},$$
where α is the attenuation. From $s(t)$ we need to estimate $\Delta t, \Delta\omega$. At the receiver, compute $|\langle s, M_\omega T_x f\rangle|$ for all $x,\omega \in \mathbb{R}$.

Covariance property of the STFT (proved in HW):
$$|\langle s, M_\omega T_x f\rangle| = |V_f f(t-\Delta t, \omega-\Delta\omega)| = |Af(t-\Delta t, \omega-\Delta\omega)|$$

Lemma 9.8.

If $f \in L^2(\mathbb{R})$, $f \neq 0$, then
$$|Af(x,\omega)| < Af(0,0) = \|f\|_2^2$$
for all $(x,\omega) \neq (0,0)$.

This follows from

$$|Af(x,\omega)| = |\langle f, M_\omega T_x f\rangle| \le \|f\|\|f\|$$

$\Rightarrow$ Simply compute $\max |Af(x,\omega)|$ to find $\Delta t, \Delta\omega$.

9.3.2 Wigner Distribution

(Wigner, 1932)

$$\begin{aligned} Wf(x,\omega) &= \int f\left(x+\frac{t}{2}\right)\overline{f\left(x-\frac{t}{2}\right)}e^{-2\pi i\omega t}\,dt \\ &= 2e^{4\pi i\omega x}V_{If}f(2x,2\omega), \qquad \text{where } If(x)=f(-x) \end{aligned}$$

9.4 Gabor Frames

The STFT is highly redundant. Do we have to take all $\langle f, M_\omega T_x g\rangle$ for $(x,\omega)\in\mathbb{R}^2$?

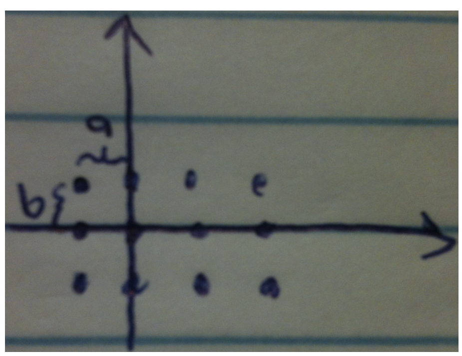
By the inversion formula, we can write f as (assuming $\langle g,\gamma\rangle = 1$)

$$f = \iint_{\mathbb{R}^2} \langle f, M_\omega T_x g\rangle\, M_\omega T_x \gamma\, d\omega\, dx$$

1. We could try to replace the integrals by Riemann sums over some lattice in $\mathbb{R}\times\mathbb{R}$

$$f \stackrel{?}{=} \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} \langle f, M_{lb}T_{ka}g\rangle\, M_{lb}T_{ka}\gamma$$

   for $a, b > 0$; we need to find γ.



2. Or we could try to express f for a given γ as

$$f = \sum_{k,l\in\mathbb{Z}} c_{kl} M_{lb}T_{ka}\gamma,$$

   where we need to find $\{c_{kl}\}$.

3. Recall that $\|V_g f\| = \|f\|$ for $\|g\|_2 = 1$. Try to sample $V_g f(x,\omega)$ dense enough so that

$$A\|f\|_2^2 \le \sum |V_g f(ka, lb)|^2 \le B\|f\|_2^2$$

   for some $A, B > 0$.

All 3 problems are equivalent and solved via frame theory.

Denis Gabor (1948) wanted to represent a function $f \in L^2(\mathbb{R})$ as

$$f = \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} c_{kl} T_{ka} M_{lb} g$$

where $g(x) = e^{-\pi x^2},\ a = b = 1$.

Definition 9.9. ***Gabor System***

Let

$$g_{kl} := T_{ka} M_{lb} g.$$

Then $\{g_{kl}\}_{k,l\in\mathbb{Z}}$ is called a *Gabor system*.

This is also known as a Weyl-Heisenberg system. In quantum physics, for g = Gaussian and $a = b = 1$, then $\{g_{kl}\}_{k,l\in\mathbb{Z}}$ are known as (canonical) coherent states.

10 2-14-12

10.1 Gabor Systems (Continued)

$$g_{kl} = T_{ka}M_{lb}g, \qquad k, l \in \mathbb{Z}$$

Definition 10.1. ***Gabor Frame***

If the set $\{g_{kl}\}_{k,l\in\mathbb{Z}}$ is a frame, then it is called a *Gabor frame*.

Theorem 10.2.

If $\{g_{kl}\}$ is a (Gabor) frame for $L^2(\mathbb{R})$, then there exists a "dual window" $\gamma \in L^2(\mathbb{R})$ such that the dual frame is $\{\gamma_{kl}\}$, where $\gamma_{kl} = T_{ka}M_{lb}\gamma$. Hence, every $f \in L^2(\mathbb{R})$ can be written as

$$f = \sum_{k,l} \langle f, T_{ka}M_{lb}g\rangle\, T_{ka}M_{lb}\gamma$$

and

$$f = \sum_{k,l} \langle f, T_{ka}M_{lb}\gamma\rangle\, T_{ka}M_{lb}g.$$

Proof. Show that the frame operator S commutes with $T_{ka}M_{lb}$.

$$(T_{ka}M_{lb})^{-1}S(T_{ka}M_{lb}) = \sum_{m,n} \langle T_{ka}M_{lb}f, T_{ma}M_{nb}g\rangle\, (T_{ka}M_{lb})^{-1}T_{ma}M_{nb}g \tag{10.1}$$

Recall:

$$(T_{ka}M_{lb})^{-1}(T_{ma}K_{nb}) = e^{-2\pi iab(m-k)l}T_{a(m-k)}M_{b(n-l)}$$

The phase factor $e^{-2\pi iab(m-k)l}$ cancels in (10.1) and we get

$$\begin{aligned}(T_{ka}M_{lb})^{-1}S(T_{ka}M_{lb}) &= \sum_{m,n} \left\langle f, T_{a(m-k)}M_{b(n-l)}g\right\rangle T_{a(m-k)}M_{b(n-l)}g \\ &= Sf \qquad \text{(by change of variables)}\end{aligned}$$

$\Rightarrow$ So S^{-1} also commutes with $T_{ka}M_{lb}$. What does this mean?

$$\underbrace{S^{-1}(T_{ka}M_{lb}g)}_{\gamma_{kl}} = T_{ka}M_{lb}\underbrace{S^{-1}g}_{\gamma}$$

$\Rightarrow$ This means that the dual frame is $\gamma_{kl} = T_{ka}M_{lb}\gamma$. The rest of the proof follows from the frame property. ∎

The significance is that we need to compute only one dual window, e.g. $\gamma = S^{-1}g$.

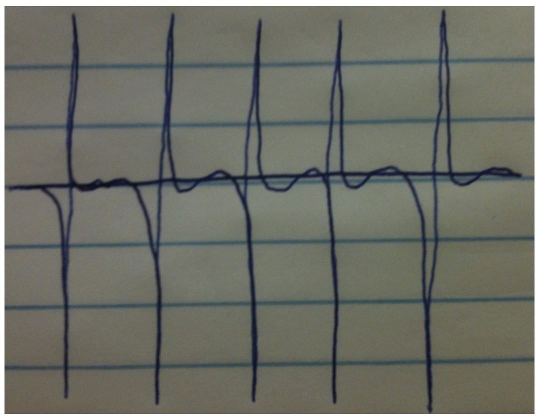
We call the discretized STFT the *Gabor transform*.

What about Gabor's choice: $g(x) = e^{-\pi x^2}$, $a = b = 1$? Is this a Gabor frame for $L^2(\mathbb{R})$?

von Neumann conjectured that $\{g_{kl}\}$ with $g(x) = e^{-\pi x^2}$, $a = b = 1$ is complete in $L^2(\mathbb{R})$. (That is, the span of $\{g_{kl}\}$ is dense in L^2.)

Theorem 10.3.

If $g(x) = e^{-\pi x^2}$ and $a = b = 1$, then $\{g_{kl}\}_{k,l\in\mathbb{Z}}$ is complete in $L^2(\mathbb{R})$. It stays complete complete if we remove an arbitrary function from $\{g_{kl}\}$, but it is incomplete if we remove any two functions from $\{g_{kl}\}$. However, it is <u>not</u> a frame for $L^2(\mathbb{R})$, since the lower frame bound $A = 0$. There exists a dual system (not a frame) $\gamma_{kl} = T_k M_l \gamma$, $\gamma \notin L^2$, where γ looks like:



Theorem 10.4.

If $g(x) = e^{-\pi x^2}$, $ab < 1$, then $\{g_{kl}\}$ is a frame for $L^2(\mathbb{R})$. If $ab > 1$, there exists <u>no</u> g such that $\{g_{kl}\}$ is a frame for $L^2(\mathbb{R})$.

(Rieffel, H. Landau, Daubechies)

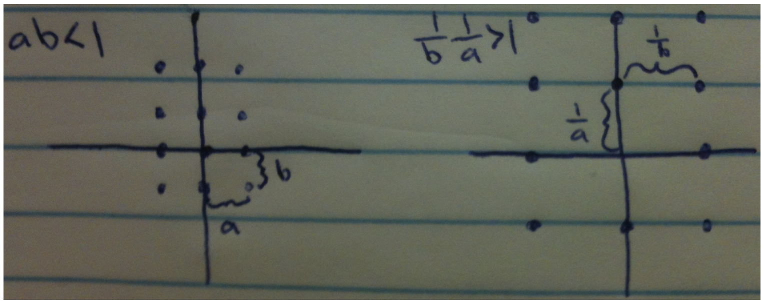
Definition 10.5. ***Redundancy***

We call $\rho = \frac{1}{ab}$ the *redundancy* of the Gabor system.

- $ab = 1$: "critical sampling" of the phase space
- $ab < 1$: oversampling, "nice" Gabor frames for $L^2(\mathbb{R})$ exist
- $ab > 1$: undersampling, no Gabor frames for $L^2(\mathbb{R})$ exist

10.2 Gabor Duality Conditions

Consider the Gabor system $\{g_{kl}\}$ generated by (g, a, b) and another Gabor system $\{g_{mn}\}$ generated by $(g, \frac{1}{b}, \frac{1}{a})$. We have two time-frequency lattices: $(ka, lb)_{k,l\in\mathbb{Z}}$ and $(\frac{m}{b}, \frac{n}{a})_{m,n\in\mathbb{Z}}$.



Theorem 10.6. ***(Jenssen, Daubechies)***

Let $g \in L^2(\mathbb{R})$ and $a, b > 0$. Then the Gabor system $\{g_{kl}\}$ generated by (g, a, b) is a frame for $L^2(\mathbb{R})$ if and only if $\{g_{mn}\}$ generated by $(g, \frac{1}{b}, \frac{1}{a})$ is a Riesz basis for a subspace of $L^2(\mathbb{R})$.

In particular, (g, a, b) generates a tight frame for $L^2(\mathbb{R}) \Leftrightarrow (g, \frac{1}{b}, \frac{1}{a})$ generates an ONB for a subspace of $L^2(\mathbb{R})$.

11 2-16-12

11.1 Gabor Duality Conditions (Continued)

Gabor frames for $\ell^2(\mathbb{Z})$:
$g \in \ell^2(\mathbb{Z})$,

$$g_{kl}(n) = g(n-ka)e^{2\pi inlb},$$

$b = \frac{1}{M},\ M \in \mathbb{N}_+,\ a \in \mathbb{N}_+,\ k \in \mathbb{Z},\ l = 0, \dots, M-1.$ $(\mathbb{Z} \xrightarrow{\text{FT}} \mathbb{T}.)$

Gabor frames on $\mathbb{C}^N$:
$g \in \mathbb{C}^N$,

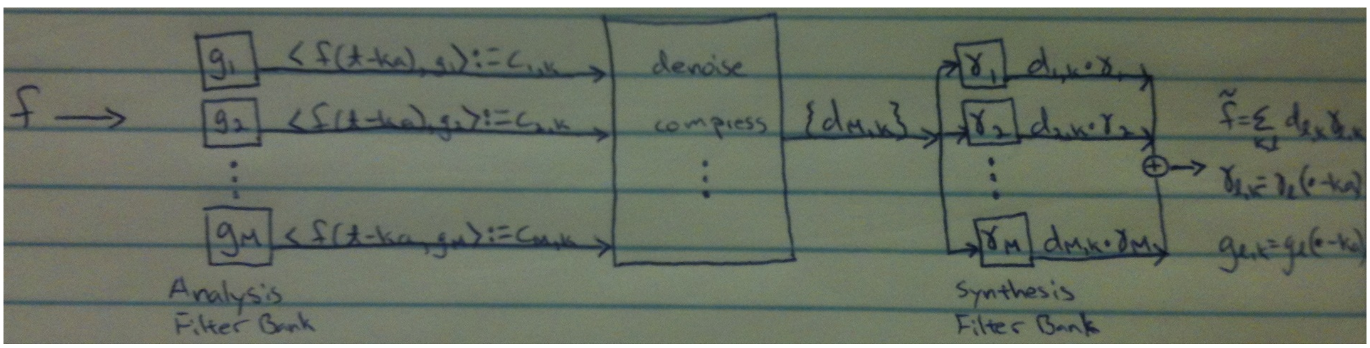
$$g_{kl}(n) = g(n-ka)e^{2\pi inlb},$$

a, b divide N.

For example: $N = 128,\ a = 8,\ b = 8,\ \frac{N}{a} = 16$ time shifts, $\frac{N}{b} = 16$ frequency shifts. $k = 0, \dots, \frac{N}{a} - 1$, $l = 0, \dots, \frac{N}{b} - 1$. $\Rightarrow$ We have 256 elements g_{kl}.

11.1.1 Application: Filter Banks

Given a (sampled) function $f \in \ell^2(\mathbb{Z})$, we want to denoise, store, compress, etc... the signal.



Perfect reconstruction condition:
If $d_{l,k} = c_{l,k}$, then we want that $\hat{f} \equiv f$.
We can ensure perfect reconstruction if $\{g_{lk}\}$ is a frame for $\ell^2(\mathbb{Z})$ (or $L^2(\mathbb{R})$) and $\{\gamma_{lk}\}$ is a dual frame.

- Linear independence of $\{g_{lk}\}$ is not necessary, but (over) completeness is.
- Modulated Filter Bank: $g_{l,k}(t) = g(t-ka)e^{2\pi itlb}$, $b = \frac{1}{M}$.
  - If $g = \mathbf{1}_{[0,a]}$, then this is called a DFT-FB (DFT Filter Bank).

11.1.2 Application: Data Transmission

Given a discrete (possibly binary) sequence of coefficients that we need to transmit. Group the data into blocks of size N: $\{c_{m,n}\}_{n=0}^{N-1}$, m-time index. Form a continuous-time signal $f(t) = \sum_{m,n} c_{m,n} g_n(t - ma)$.

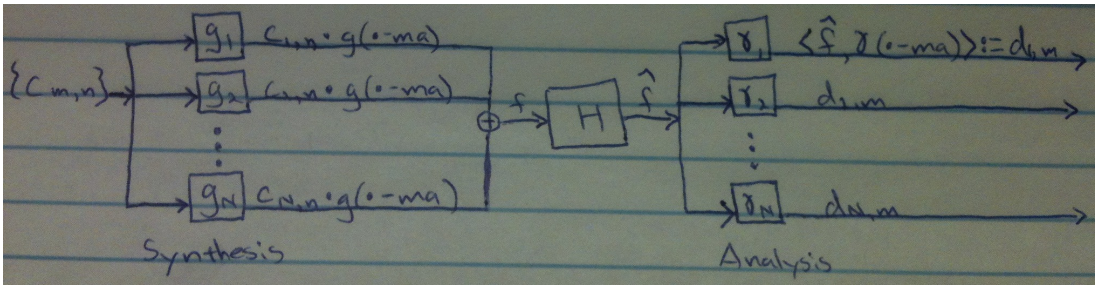


Figure 6: Operator H models the communication channel (e.g. phone line).

Perfect reconstruction condition: $d_{n,m} = c_{n,m}$. Need: $\{g_{n,m}\}$ are *linearly independent* (no frames); we can have incompleteness.

In DSL, we choose

$$g_{n,m}(t) = g(t - ma)e^{2\pi itlb},$$

$g(t) = \mathbf{1}_{[0,c]}$, $c < a$.

OFDM (orthogonal frequency division multiplexing) uses Gabor-type systems.

Know: using $g(x) = e^{-\pi x^2}$, $a = b = 1$ "does not work" in practice; the dual is not in $L^2(\mathbb{R})$ (it is a distribution).

Can use: $g(x) = \mathbf{1}_{[0,1]}$, $a = b = 1$. $\{g_{k,l}\}$ is a Gabor ONB for $L^2(\mathbb{R})$. But $\hat{g}(\omega) = \frac{\sin \pi\omega}{\pi\omega} \Rightarrow$ no good frequency localization.

Can we modify $g = \mathbf{1}_{[0,1]}$ to get a time-frequency well-localized g such that $\{g_{kl}\}$ is an ONB for $L^2(\mathbb{R})$?
Answer: No.
Can we get a Riesz basis?
Answer: Still no.

Theorem 11.1. ***Balian-Low Theorem (also Coifamn, Daubechies,...)***

If $\{g_{kl}\}_{k,l\in\mathbb{Z}}$ with $a = b = 1$ is an ONB for $L^2(\mathbb{R})$, then either

$$\int |xg(x)|^2 \, dx = \infty \qquad \text{OR} \qquad \int |\omega\hat{g}(\omega)|^2 \, d\omega = \infty.$$

See Proposition (13.2).

Proof. Let X be the position operator $Xf(x) = xf(x)$, and let P be the momentum operator $Pf(x) = \frac{1}{2\pi i}f'(x)$.

Recall:

1. $(PX - XP)f = \frac{1}{2\pi i}f$ for $f \in$ domain$(XP) \cap$ domain(PX).
2. $\widehat{Pf} = X\hat{f}$

The proof is done via contradiction. Assume that $\{g_{kl}\}$ is an ONB for $L^2(\mathbb{R})$, that $Xg \in L^2(\mathbb{R})$, and that $Pg \in L^2(\mathbb{R})$. Then

$$\langle Xg, Pg\rangle = \sum_{k,l} \langle Xg, T_k M_l g\rangle \langle T_k M_l g, Pg\rangle$$

Since

$$\begin{aligned} XT_k M_l g(x) &= (k + x - k)e^{2\pi i l x} g(x - k) \\ &= kT_k M_l g(x) + T_k M_l X g(x) \end{aligned}$$

So

$$\begin{aligned} \langle Xg, T_k M_l g\rangle &= \langle g, XT_k M_l g\rangle \\ &= \underbrace{k \overbrace{\langle g, T_k M_l g\rangle}^{=0 \text{ if } (k,l)\neq(0,0)}}_{=0} + \langle g, T_k M_l X g\rangle \\ &= \langle T_{-k} M_{-l} g, Xg\rangle \qquad\qquad \text{(because } k, l \in \mathbb{Z}) \end{aligned}$$

We also have

$$\begin{aligned} \langle T_k M_l g, Pg\rangle &= \langle PT_k M_l g, g\rangle \\ &= \langle XM_{-k} T_l \hat{g}, \hat{g}\rangle \\ &= \underbrace{l \langle M_{-k} T_l \hat{g}, \hat{g}\rangle}_{=0} + \langle M_{-k} T_l X \hat{g}, \hat{g}\rangle \\ &= \langle Pg, T_{-k} M_{-l} g\rangle \end{aligned}$$

Using these formulas, we get

$$\begin{aligned} \langle Xg, Pg\rangle &= \sum_{k,l} \langle Pg, T_{-k} M_{-l} g\rangle \langle T_{-k} M_{-l} g, Xg\rangle \\ &= \langle Pg, Xg\rangle \end{aligned}$$

$\Rightarrow$ For $g \in \operatorname{dom}(XP) \cap \operatorname{dom}(PX)$, it follows that

$$\begin{aligned} \langle Xg, Pg\rangle - \langle Pg, Xg\rangle &= 0 \\ &= \langle PXg, g\rangle - \langle XPg, g\rangle \\ &= \langle (PX - XP)g, g\rangle \\ \text{(uncertainty principle)} \qquad &= \frac{1}{2\pi i} \langle g, g\rangle \neq 0 \quad \text{if } g \neq 0 \qquad\qquad \Rightarrow\Leftarrow \end{aligned}$$

Now we need to get rid of our assumption that $g \in \operatorname{dom}(XP) \cap \operatorname{dom}(PX)$. If $g \notin \operatorname{dom}(XP) \cap \operatorname{dom}(PX)$, we can choose a sequence $f_n \in \mathcal{S}(\mathbb{R})$ such that $\|f_n - g\|_2 \to 0$, $\|Xf_n - Xg\|_2 \to 0$, and $\|Pf_n - Pg\|_2 \to 0$. Then

$$\lim_{n\to\infty} (\langle Xf_n, Pf_n\rangle - \langle Pf_n, Xf_n\rangle) = \langle Xg, Pg\rangle - \langle Pg, Xg\rangle = 0.$$

But since $f_n \in \mathcal{S}(\mathbb{R}) \in \operatorname{dom}(XP) \cap \operatorname{dom}(PX)$, we also have

$$\begin{aligned} \lim_{n\to\infty} \langle (PX - XP)f_n, f_n\rangle &= \frac{1}{2\pi i} \lim_{n\to\infty} \|f_n\|_2^2 \\ &= \frac{1}{2\pi i} \|g\|_2^2 \\ &\neq 0 \qquad\qquad \Rightarrow\Leftarrow \end{aligned}$$

□

This theorem can be generalized to Riesz bases, and to $L^2(\mathbb{R}^d)$.

12 2-21-12

12.1 Wavelets

Problems of STFT and Gabor Systems:

- If a window is too large, then it cannot localize sharp transitions in signal (for a fixed window size)
- If a window is too small, then it cannot detect low frequency oscillators (for a fixed window size)
- The Balian-Low Theorem: there do not exist "nice" Gabor ONB's.

Key idea of waveletes: use *translations* and *dilations* (no modulations) of a *single function* to analyze a signal at different "resolutions."

Definition 12.1. *Wavelet*

A *wavelet* is a function (often called "*mother wavelet*") $\Psi \in L^2(\mathbb{R})$ with

$$\int_{-\infty}^{\infty} \Psi(x)\,dx = 0,$$

which is normalized to have $\|\Psi\|_2 = 1$, and it is centered around $x = 0$.

Let's generate a family of time-frequency atoms (= window functions):

$$\begin{aligned}\Psi_{a,b}(x) &= \frac{1}{\sqrt{a}}\Psi\left(\frac{x-b}{a}\right), \qquad a > 0,\ b \in \mathbb{R}\\ &= T_b D_a \Psi\end{aligned}$$

Note: $\|\Psi_{a,b}\|_2 = 1$.

Also note:

$$\begin{aligned}T_b D_a \psi(x) &= D_a T_{b/a} \psi(x)\\ D_a T_b \psi(x) &= T_{ba} D_a \psi(x)\end{aligned}$$

Definition 12.2. ***Wavelet Transform***

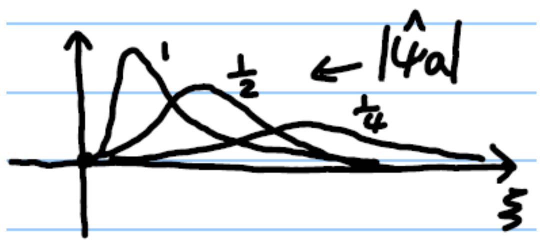
The *wavelet transform* (often called "*continuous wavelet transform*") of $f \in L^2(\mathbb{R})$ is defined as

$$\begin{aligned} Wf(a,b) = W_\Psi f(a,b) &:= \langle f, \Psi_{a,b} \rangle \\ &= \int_{-\infty}^{\infty} f(x) \overline{T_b D_a \Psi(x)}\, dx \\ &= \int_{-\infty}^{\infty} f(x) \frac{1}{\sqrt{a}} \overline{\Psi\left(\frac{x-b}{a}\right)}\, dx \end{aligned}$$

This can be viewed as a linear filtering:

$$\int_{-\infty}^{\infty} f(x) \frac{1}{\sqrt{a}} \overline{\Psi\left(\frac{x-b}{a}\right)}\, dx = f * \overline{\tilde{\Psi}_a}(b)$$

$$\tilde{\Psi}_a(x) := \frac{1}{\sqrt{a}} \Psi\left(-\frac{x}{a}\right) \quad \overset{\mathcal{F}}{\mapsto} \quad \widehat{\tilde{\Psi}_a}(\xi) = \sqrt{a}\,\overline{\hat{\Psi}(a\xi)} = D_{1/a} \overline{\hat{\Psi}(\xi)}$$



Types of wavelets:

- Real waveletes → good for edges
- Analytic (or complex) wavelets → can detect phases of a signal

For the time being, let's focus on real wavelets.

Example 12.3. ***Mexican Hat Function, or Laplacian of Gaussian (LoG)***

$$\Psi(x) = \frac{2}{\pi^{1/4}\sqrt{3\sigma}} \left(1 - \frac{x^2}{\sigma^2}\right) e^{-x^2/2\sigma^2} = -\frac{d^2}{dx^2} \left(\frac{2\sigma^2}{\pi^{1/4}\sqrt{3\sigma}} \cdot e^{-x^2/2\sigma^2} \right)$$

$$\hat{\Psi}(\xi) = 8\sqrt{\frac{2}{3}} \pi^{9/4} \sigma^{5/2} \xi^2 e^{-2\pi^2\sigma^2\xi^2}$$

$\hat{\Psi}(0) = 0$, $\hat{\Psi}(\xi) \sim \xi^2$ around $\xi = 0$ → approximates $\frac{d^2}{dx^2}$.

Theorem 12.4. ***Calderón-Grossman-Morlet***

Let $\Psi \in L^2(\mathbb{R})$, $\Psi \in \mathbb{R}$ such that

$$C_\Psi := \int_0^\infty \frac{|\hat{\Psi}(\xi)|^2}{\xi}\, d\xi < \infty.$$

Then any $f \in L^2(\mathbb{R})$ satisfies

$$f(x) = \frac{1}{C_\Psi}\int_0^\infty \int_{-\infty}^\infty Wf(a,b)\Psi_{a,b}(x)\, db\, \frac{da}{a^2} \tag{12.1}$$
$$\|f\|_2^2 = \frac{1}{C_\Psi}\int_0^\infty \int_{-\infty}^\infty |Wf(a,b)|^2\, db\, \frac{da}{a^2}$$

Proof.

$$\begin{aligned}
Wf(a,b) &= f * \tilde{\Psi}_a(b)\\
\text{(RHS of 12.1)} &= \frac{1}{C_\Psi}\int_0^\infty (Wf(a,\cdot) * \Psi_{a,\cdot})(x)\, \frac{da}{a^2}\\
&= \frac{1}{C_\Psi}\int_0^\infty (f * \tilde{\Psi}_a * \Psi_a)(x)\, \frac{da}{a^2}\\
\text{(by } \mathcal{F}) \qquad &= \frac{1}{C_\Psi}\int_0^\infty \hat{f}(\xi)\sqrt{a}\overline{\hat{\Psi}(a\xi)}\sqrt{a}\hat{\Psi}(a\xi)\, \frac{da}{a^2}\\
&= \frac{\hat{f}(\xi)}{C_\Psi}\int_0^\infty \frac{|\hat{\Psi}(a\xi)|^2}{a}\, da\\
&= \frac{\hat{f}(\xi)}{C_\Psi}\int_0^\infty \frac{|\hat{\Psi}(\eta)|^2}{\eta}\, d\eta \qquad\qquad (\eta = a\xi)
\end{aligned}$$

□

$C_\Psi < \infty$ is called the *admissibility condition*, and (12.1) is called *Calderón's reproducing formula*.

$$f(x) = \frac{1}{C_\Psi}\int_0^\infty f * \tilde{\Psi}_a * \Psi_a(x) \frac{da}{a^2}$$

This is also called the *resolution of the identity*.

To guarantee $C_\Psi < \infty$, we need

$$\hat{\Psi}(0) = 0 \quad \Leftrightarrow \quad \int_{-\infty}^\infty \Psi(x)\, dx = 0$$

So, Ψ must be oscillatory with $\pm$ values. We also need decay on Ψ, e.g.

$$\int_{-\infty}^\infty (1 + |x|)\Psi(x)\, dx < \infty.$$

12.2 Reproducing Kernel

CWT (Continuous Wavelet Transform) = a redundant representation.

$$Wf(a,b) = \int_{-\infty}^{\infty} \underbrace{\left(\frac{1}{C_\Psi} \int_0^\infty \int_{-\infty}^{\infty} Wf(a',b')\Psi_{a',b'}(x)\, db' \, \frac{da'}{a'^2} \right)}_{=f(x)} \cdot \overline{\Psi_{a,b}(x)}\, dx \tag{12.2}$$

$$= \frac{1}{C_\Psi} \int_0^\infty \int_{-\infty}^{\infty} K(a,a',b,b') Wf(a',b')\, db' \, \frac{da'}{a'^2}$$

where $K(a,a'b,b') := \left\langle \Psi_{a,b}, \Psi_{a',b'} \right\rangle$, which measures the correlation between $\Psi_{a,b}$ and $\Psi_{a',b'}$. If $K(a,a',b,b') = \delta(a-a')\delta(b-b')$, then there is no redundancy.

Proposition 12.5.

A function $\Phi(a,b) \in L^2(\mathbb{R}_+ \times \mathbb{R})$ is a wavelet transform of some $f \in L^2(\mathbb{R}) \Leftrightarrow \Phi(a,b)$ satisfies (12.2).

12.3 Scaling Function (Father Wavelet)

The reconstruction formula requires all values of scales $0 < a < \infty$. If we only know $Wf(a,b)$ for $a < a_0$, then we need complementary info for $a > a_0$ provided by the *scaling function* (*father wavelet*).

$$|\hat{\phi}(\xi)|^2 := \int_1^\infty |\hat{\Psi}(a\xi)|^2 \frac{da}{a}$$

$$= \int_\xi^\infty \frac{|\hat{\Psi}(\eta)|^2}{\eta}\, d\eta$$

The phase of ϕ can be arbitrarily chosen.

- $\lim_{\xi \to 0} |\hat{\phi}(\xi)|^2 = C_\Psi$
- $\|\phi\|_2 = 1 \leftarrow$ Exercise: use the definition. So the low frequency approximation of f at scale a can be written as

$$Lf(a,x) := \left\langle f, \underbrace{D_a\phi}_{=\phi_a} \right\rangle = f * \tilde{\phi}_a(x)$$

$$f(x) = \frac{1}{C_\Psi} \int_0^{a_0} (Wf(a,\cdot) * \Psi_a)(x) \frac{da}{a^2} + \frac{1}{C_\Psi a_0} \left(Lf(a_0,\cdot) * \phi_{a_0} \right)(x)$$

Example 12.6. ***Mexican Hat Function***

$$
\begin{aligned}
\Psi(x) &= \frac{2}{\pi^{1/4}\sqrt{3\sigma}}\left(1-\frac{x^2}{\sigma^2}\right)e^{-x^2/2\sigma^2} \\
\hat{\Psi}(\xi) &= 8\sqrt{\frac{2}{3}}\pi^{9/4}\sigma^{5/2}\xi^2 e^{-2\pi^2\sigma^2\xi^2} \\
\Rightarrow \qquad |\hat{\phi}(\xi)|^2 &= \frac{4\sigma}{3\sqrt{\pi}}\left(1+4\pi^2\sigma^2\xi^2\right)e^{-4\pi^2\sigma^2\xi^2} \\
\Rightarrow \qquad \hat{\phi}(\xi) &= 2\sqrt{\frac{\sigma}{3\sqrt{\pi}}}\sqrt{1+4\pi^2\sigma^2\xi^2}e^{-2\pi^2\sigma^2\xi^2}
\end{aligned}
$$

→ Choose a simple phase factor.

12.4 Discrete Wavelet Transforms

How to sample $Wf(a,b)$?

Another great insight by J. Morlet: "regular hyperbolic grid"

$$(a,b) = (a_0^m, n a_0^m b_0), \qquad m,n \in \mathbb{Z}$$

Theorem 12.7. ***Regular Sampling Theorem (Daubechies, 199)***

Let Ψ be a real-valued L^2-function. For fixed a_0, b_0, define

$$\begin{aligned}\Psi_{m,n}(x) &:= a_0^{-m/2}\Psi(a_0^{-m}x - nb_0) \\ &= \frac{1}{\sqrt{a_0^m}}\Psi\left(\frac{x - na_0^m b_0}{a_0^m}\right) \\ &= T_{na_0^m b_0} D_{a_0^m}\Psi(x), \qquad m, n \in \mathbb{Z}\end{aligned}$$

1. If $\{\Psi_{m,n}\}_{(m,n)\in\mathbb{Z}^2}$ is a frame in $L^2(\mathbb{R})$ with frame bounds A, B, then we must have

$$A \leq \frac{1}{b_0}\sum_{-\infty}^{\infty}|\hat{\Psi}(a_0^m\xi)|^2 \leq B$$

   for $\xi \in \mathbb{R}$ a.e. In particular, Ψ satisfies the admissibility condition:

$$C_\Psi = \int_0^\infty |\hat{\Psi}(\xi)|^2\frac{d\xi}{\xi} < \infty$$

2. If for some $\epsilon > 0$, Ψ satisfies $|x|^{\frac{1}{2}+\epsilon}\Psi \in L^2$, $|\xi|^\epsilon\hat{\Psi} \in L^2$, and $\int \Psi(x)\, dx = 0$, then Ψ satisfies

$$\begin{cases} \operatorname{ess\,inf} \sum_{m\in\mathbb{Z}} |\hat{\Psi}(a_0^m\xi)|^2 > 0 \\ \operatorname{ess\,sup} \sum_{m\in\mathbb{Z}} |\hat{\Psi}(a_0^m\xi)|^2 < \infty \end{cases} \tag{12.3}$$

   for any a_0 close enough to 1. (i.e., there exists $\alpha = \alpha(\Psi) > 1$ such that (12.3) is satisfied for any $a_0 \in (1, \alpha)$.) Moreover, if b_0 is close enough to 0 (i.e., there exists $\beta = \beta(a_0, \Psi)$ for a_0 satisfying (12.3) such that $b_0 \in (0, \beta)$), then $\{\Psi_{m,n}\}_{(m,n)\in\mathbb{Z}^2}$ constitute a frame.

Example 12.8.

Let $\Psi(x)$ be the Mexican hat function, $a_0 = 2, b_0 = \frac{1}{4}$. Then $\{\Psi_{m,n}\}_{(m,n)\in\mathbb{Z}^2}$ forms a frame (called a wavelet frame). $A = 13.09,\ B = 14.18$, i.e., it is almost tight!

Definition 12.9. ***Dual Frame***

The wavelet frame operator S commutes with dilations $D_{a_0^m}$, but <u>not</u> with translations $T_{na_0^m b_0}$.

$\Rightarrow$ The dual frame, $S^{-1}(T_{na_0^m b_0} D_{a_0^m}\Psi)$ is in general <u>not</u> a wavelet system (unlike for Gabor frames).

13 2-28-12

13.1 Wavelet ONB for $L^2(\mathbb{R})$

Definition 13.1. ***Haar Basis (Haar, 1909)***

$$\psi(x) = \begin{cases} 1 & 0 \leq x < \frac{1}{2} \\ -1 & \frac{1}{2} \leq x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \psi_{m,n}(x) &= 2^{-m/2}\psi(2^{-m}x - n) \\ &= T_{2^m n} D_{2^m} \psi(x), \qquad m \in \mathbb{N},\ n \in \mathbb{Z} \end{aligned}$$

$\Rightarrow \{\psi_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$.

But the Haar Wavelet is not continuous, and therefore $\hat{\psi}$ is not localized.

Proposition 13.2. ***Meyer (1985), Stromberg (1982)***

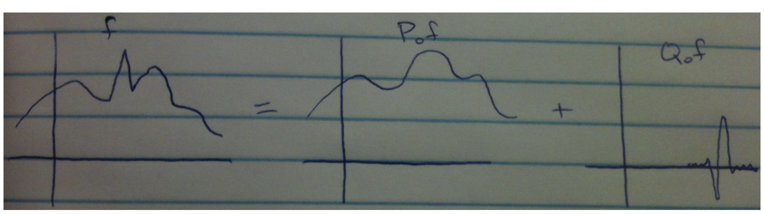
There exists a wavelet ONB for $L^2(\mathbb{R})$, where ψ can have exponential decay and $\psi \in \mathbb{C}^k,\ k < \infty$.

$\Rightarrow$ There is no Balian-Low Theorem (11.1) for wavelets.

13.2 Multiresolution Analysis and Wavelet ONB's

Let V_{-1} be a subspace of $L^2(\mathbb{R})$ and $f \in V_{-1}$. We want to decompress f into a "smooth" part (low-frequency) and a "rough" part (or coarse, or edgy) (high frequency). We do this by projecting f onto $V_0 \subseteq V_{-1}$ containing the smooth part of f, $P_0 f$, and project f onto $W_0 := V_0^{\perp}$ containing the rough part of f, $Q_0 f$.

$$f = P_0 f + Q_0 f, \qquad V_{-1} = V_0 \oplus W_0$$



Leave $Q_0 f$ and proceed with $P_0 f$ by representing V_0 as

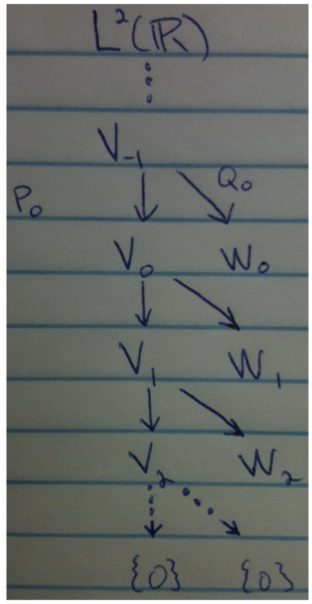
$$V_0 = \underbrace{V_1}_{\text{"smooth"}} \oplus \underbrace{W_1}_{\text{"rough"}}$$

- Project onto V_1: P_1
- Project onto W_1: Q_1

Since $P_1 P_0 f = P_1 f$ and $Q_1 Q_0 f = Q_1 f$, we get

$$\begin{aligned} P_0 f &= P_1 f + Q_1 f \\ f &= P_1 f + Q_1 f + Q_0 f \end{aligned}$$

We can split $P_1 f$ into $P_2 f$ and $Q_2 f$, ...



Definition 13.3. ***Multiresolution Analysis, Scaling Function***

A *multiresolution analysis (MRA)* of $L^2(\mathbb{R})$ is an increasing sequence of closed subspaces $V_m \in L^2(\mathbb{R})$,

$$\{0\} \subset \cdots \subset V_2 \subset V_1 \subset V_0 \subset \cdots \subset L_2(\mathbb{R}),$$

such that

1. $\overline{\cup_{m\in\mathbb{Z}} V_m} = L^2(\mathbb{R})$
2. $\cap_{m\in\mathbb{Z}} V_m = \{0\}$
3. $f(\cdot) \in V_m \;\Leftrightarrow\; f(2^m \cdot) \in V_0$
4. There exists a function $\phi \in L^2(\mathbb{R})$ (not a wavelet) whose integer translates $T_k\phi,\ k \in \mathbb{Z}$, form a Riesz basis for V_0. ϕ is called the *scaling function*.

MRA is key to

- Constructing a wavelet ONB
- Fast Wavelet Transform

Remarks:

- V_0 is invariant under integer translations: $f \in V_0 \;\Leftrightarrow\; T_k f \in V_0$
- MRA condition #3 implies that $f \in V_m \;\Leftrightarrow\; T_{2^m k} f \in V_m \;\forall\, k \in \mathbb{Z}$

- MRA conditions #3 and #4 imply that V_m is spanned by the following functions:

$$\begin{aligned} \phi_{m,k}(x) &:= 2^{-m/2}\phi(2^{-m}x - k) \\ &= T_{2^m} D_{2^m} \phi(x). \end{aligned}$$

Lemma 13.4.

The scaling function ϕ satisfies a "scaling equation." There exists a sequence $\{\alpha_k\}_{k\in\mathbb{Z}}$ of real numbers with the following property:

$$\begin{aligned} \phi(x) &= \sqrt{2}\sum_{k\in\mathbb{Z}} \alpha_k \phi(2x - k) \\ &= \sum_{k\in\mathbb{Z}} \alpha_k T_{k/2} D_{1/2} \phi(x). \end{aligned}$$

Proof. This result follows from the fact that $\phi \in V_0 \subset V_{-1} = \text{span}\left\{\sqrt{2}\phi(2x-k),\ k \in \mathbb{Z}\right\} = \text{span}\left\{T_{k/2}D_{1/2}\phi(x)\right\}$. □

Note:

- $V_{m-1} = V_m \oplus W_m$, $V_m \perp W_m$
- $P_{m-1} = P_m + Q_m$, $Q_m = P_{m-1} - P_m$
  - Averages lead to smoothing, differences lead to fine details.
- We have $V_m = \oplus_{j\geq m+1} W_j$, and so $L^2(\mathbb{R}) = \oplus_{j\in\mathbb{Z}} W_j$
- Also, $f(\cdot) \in W_m \;\Leftrightarrow\; f(2^m \cdot) \in W_0$

$f \in L^2(\mathbb{R})$ can be decomposed as

$$\begin{aligned} f &= \sum_{j\in\mathbb{Z}} Q_j f = \sum_{j\geq m+1} Q_j f + \sum_{j<m+1} Q_j f \\ &= P_m f + \sum_{j<m+1} Q_j f. \end{aligned}$$

$Q_j f$ contains the details of f which distinguish P_{j-1} from P_j, since $Q_j = P_{j-1} - P_j$. We will see that for every MRA there exists a wavelet ψ such that $\psi_{m,k}(x) = 2^{-m/2}\psi(2^{-m}x - k)$, where $\{\psi_{m,k}\}_{k\in\mathbb{Z}}$ is an ONB for W_m for fixed m. ψ can be explicitly constructed from ϕ.

14 3-1-12

14.1 MRA (Continued)

(Recall) MRA:

1. $\cup V_m = L^2(\mathbb{R})$
2. $\cap V_m = \{0\}$
3. $f(\cdot) \in V_m \;\Leftrightarrow\; f(2^m \cdot) \in V_0$
4. $\{T_k \phi\}$ is a Riesz basis for V_0

ϕ is the scaling function:

$$\phi(x) = \sqrt{2} \sum \phi_k \phi(2x - k)$$

Example 14.1. ***Haar Wavelet ONB***

Set

$$\phi(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

V_0 consists of functions that are constant on $[k, k+1],\; k \in \mathbb{Z}$.

For general m:

$$\begin{aligned} V_m &= \operatorname{span}\{\phi_{m,k},\; k \in \mathbb{Z}\} \\ &= \{f \in L^2(\mathbb{R}) \mid f \text{ is constant on } [2^m k, 2^m(k+1)],\; k \in \mathbb{Z}\} \end{aligned}$$

The family V_m generates a MRA.

In this example (but not in general), $\{\phi_{m,k}\}_{k \in \mathbb{Z}}$ is an ONB for V_m.

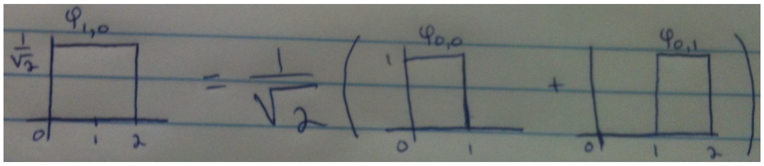
$$\begin{aligned} P_m f &= \sum_k \underbrace{\langle P_m f, \phi_{m,k} \rangle}_{=: c_k^m} \phi_{m,k} \\ c_m^k &= \langle P_m f, \phi_{m,k} \rangle = \langle f, P_m \phi_{m,k} \rangle = \langle f, \phi_{m,k} \rangle \\ &= 2^{-m/2} \int_{2^m k}^{2^m(k+1)} f(x)\, dx \end{aligned}$$

This is a local average. What is the difference between $P_m f$ and the next coarser level, $P_{m+1} f$? The scaling equation in our example is

$$\phi(x) = \sqrt{2} \left(\frac{1}{\sqrt{2}} \phi(2x) + \frac{1}{\sqrt{2}} \phi(2x - 1) \right).$$

More generally:

$$\phi_{m+1,k} = \frac{1}{\sqrt{2}} (\phi_{m,2k} + \phi_{m,2k+1})$$



Thus,

$$c_k^{m+1} = \langle f, \phi_{m+1,k}\rangle = \frac{1}{\sqrt{2}}(c_{2k}^m + c_{2k+1}^m)$$

$P_{m+1}f$ is an averaged version of $P_m f$.
Difference:

$$\begin{aligned}
P_m f - P_{m+1} f &= \sum_k c_k^m \phi_{m,k} - \sum_k c_k^{m+1}\phi_{m+1,k} \\
&= \underbrace{\sum_k c_k^m \phi_{m,k}}_{=\sum_k c_{2k}^m \phi_{m,2k} + \sum_k c_{2k+1}^m \phi_{m,2k+1}} - \frac{1}{2}\sum_k (c_{2k}^m + c_{2k+1}^m)(\phi_{m,2k} + \phi_{m,2k+1}) \\
&= \frac{1}{2}\sum_k (c_{2k}^m - c_{2k+1}^m)(\phi_{m,2k} - \phi_{m,2k+1}) \qquad (14.1)
\end{aligned}$$

The difference $\phi_{m,2k} - \phi_{m,2k+1}$ is:

$$\frac{1}{\sqrt{2}}(\phi_{m,2k} - \phi_{m,2k+1}) =: \psi_{m+1,k}$$

$$\text{where} \qquad \psi(x) = \phi(2x) - \phi(2x-1) = \begin{cases} 1 & 0 \le x < \frac{1}{2} \\ -1 & \frac{1}{2} \le x < 1 \\ 0 & \text{otherwise} \end{cases}$$

The system of functions $\{\psi_{m,k}\}_{k\in\mathbb{Z}}$ is an ONB for W_m. The projectors Q_m fulfill

$$Q_{m+1} f = P_{m+1} f - P_m f = \sum d_k^{m+1}\psi_{m+1,k}$$

$$\text{with} \qquad d_k^{m+1} = \underbrace{\langle f, \psi_{m+1,k}\rangle} = \frac{1}{\sqrt{2}}(c_{2k}^m - c_{2k+1}^m)$$

where the underlined term gives the coefficients of the discrete wavelet transform of f with respect to $\{\psi_{m+1,k}\}_k$ and scale $m+1$.

Theorem 14.2.

Let $\{V_m\}$ be a MRA generated by the scaling function $\phi \in V_0$, where $\{T_k\phi\}_{k\in\mathbb{Z}}$ is an ONB for V_0. The function $\psi \in V_{-1}$ defined by

$$\psi(x) = \sqrt{2}\sum_k b_k \phi(2x-k) = \sum b_k \phi_{-1,k}(x),$$

where $b_k = (-1)^k a_k$ and the a_k are scaling coefficients, has the following properties:

1. $\{\psi_{m,k}(\cdot) = 2^{-m/2}\psi(2^{-m/2}\cdot - k),\ k \in \mathbb{Z}\}$ forms an ONB for W_m.
2. $\{\psi_{m,k}\}_{m,k\in\mathbb{Z}}$ is an ONB for $L^2(\mathbb{R})$.
3. ψ is a wavelet with $C_\psi = 2\ln 2$.

Note: if $\{T_k h\}_{k\in\mathbb{Z}}$ is a Riesz basis for V_0, then we can define a function $\phi = S^{-1/2}h$, where S is the frame operator, and $\{T_k\phi\}_{k\in\mathbb{Z}}$ is an ONB for V_0.

14.2 Fast Wavelet Transform

Given a wavelet ONB $\{\psi_{m,k}\}$ for $L^2(\mathbb{R})$, the discrete wavelet transform (DWT) is $Wf = \{\langle f, \psi_{m,n}\rangle\}_{m,n\in\mathbb{Z}}$. We want to the compute the DWT with a fast algorithm.

Consider a function $f \in V_0$. We know that

$$f(x) = \sum_k c_k^0 \phi(x-k), \qquad \text{where } c_k^0 = \langle f, T_k\phi\rangle$$

Assume we have computed $\{c_k^0\}$. (It is often assumed that c_k^0 "=" $f(k)$, although $c_k^0 = \langle f, T_k\phi\rangle$ does not imply in general that $c_k^0 = f(k)$.)

We use the following notation:

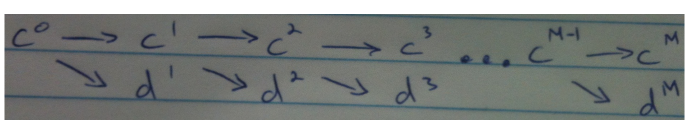
$$\begin{aligned} d_k^m &= \langle f, \psi_{m,k}\rangle\,, \qquad & d^m &= \{d_k^m\}_{k\in\mathbb{Z}} \\ c_k^m &= \langle f, \phi_{m,k}\rangle\,, \qquad & c^m &= \{c_k^m\}_{k\in\mathbb{Z}} \end{aligned}$$

We get via $\phi(x) = \sqrt{2}\sum_l a_l\phi(2x-l)$ and $\psi(x) = \sum_l b_l\phi_{-1,l}(x)$ that

$$\begin{aligned} d_k^m = \langle f, \psi_{m,k}\rangle &= \sum_l b_l \underbrace{\langle f, \phi_{m-1,2k+l}\rangle}_{c_{2k+l}^{m-1}} \\ &= \sum_l b_{l-2k} c_l^{m-1} \\ c_k^m = \langle f, \phi_{m,k}\rangle &= \sum_l a_l \underbrace{f, \phi_{m-1,2k+l}}_{c_{2k+l}^{m-1}} \\ &= \sum_l a_{l-2k} c_l^{m-1} \end{aligned}$$

$\Rightarrow$ We can compute c^m, d^m recursively from c^{m-1}.

Input: $\{f(x)\}$ "=" c^0, M levels of decomposition.



Output: $d^1, d^2, d^3, \ldots, d^M, c^M$

14.2.1 Complexity of the Fast Wavelet Transform

Let c^0 be a signal of length $N = 2^p$, $p \in \mathbb{N}$. Recall:

$$\begin{aligned} d_k^m &= \sum_l b_{l-2k} c_l^{m-1} \\ c_k^m &= \sum_l a_{l-2k} c_l^{m-1} \end{aligned}$$

Assume that $\{a_k\}, \{b_k\}$ have L nonzero coefficients. ($L = 2$ for a Haar basis.) With appropriate boundary conditions (e.g. periodization, zero padding, etc.), each c^m and d^n has $2^{-m}N$ samples. Then c^{m+1}, d^{m+1} can be computed from c^m with $2^{-m}NL$ operations. $\Rightarrow$ The total number of operations is $O(NL)$.

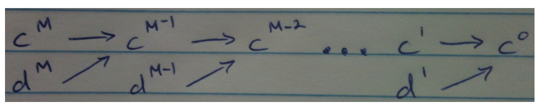
14.3 Inverse Wavelet Transform

Goal: To reconstruct c^0 from $\{d^1, d^2, \ldots, d^M, c^M\}$.

Assume we already have computed c^1, d^1 and we need to compute c^0. Using $V_0 = V_1 \oplus W_1$, we have

$$
\begin{aligned}
f &= \sum_k c_k^0 \phi_{0,k} = \sum_k c_k^1 \phi_{1,k} + \sum_k d_k^1 \psi_{1,k} \\
&= \sum_k c_k^1 \sum_l a_l \phi_{0,2k+l} + \sum_k d_k^1 \sum_l b_l \phi_{0,2k+l} \\
c_j^0 &= \langle f, \phi_{0,j} \rangle = \sum_l c_k^1 \langle \phi_{1,k}, \phi_{0,j} \rangle + \sum_k d_k^1 \langle \psi_{1,k}, \phi_{0,j} \rangle \\
&= \sum_k c_k^1 a_{j-2k} + \sum_k d_k^1 b_{j-2k}
\end{aligned}
$$

Similarly, we can compute c^{m-1} from c^m, d^m:



15 3-6-12

15.1 Wavelets in 2-D

There are two common approaches:

1. via tensor products. This leads to a separable MRA on $L^2(\mathbb{R})$. Given a 1-D MRA $\{V_j\}$, we get a 2-D MRA via $V_j \otimes V_j$. $\Rightarrow$ Wavelets

$$\psi \underbrace{j_1, j_2}_{\text{translations}} , \underbrace{n_1, n_2}_{\text{dilations}} (x_1, x_2) = \psi_{j_1,n_1}(x_1) \cdot \psi_{j_2,n_2}(x_2)$$

2. Three different wavelets:

   - ψ_H has horizontal orientation
   - ψ_V has vertical orientation
   - ψ_D has diagonal orientation

   $\Rightarrow$

$$(V_{j-1} \otimes V_{j-1}) = \underbrace{(V_j \otimes V_j)}_{\text{coarse approx.}} \oplus \underbrace{(V_j \oplus W_j)}_{\text{vertical details}} \oplus \underbrace{(W_j \otimes V_j)}_{\text{horizontal details}} \oplus \underbrace{(W_j \otimes W_j)}_{\text{diagonal details}}$$

"Waveletes are good for representing point singularities, but not so good for line singularities."
$\Rightarrow$ Wavelets are not that great for representing edges in images.

New constructions for 2-D: Curvelets, Shearlets

15.2 Comparison of Wavelets and Gabor Systems

Group-theoretical viewpoint:
Gabor: $(T_x M_\omega)(T_{x'} M_{\omega'}) = e^{2\pi i x' - \omega} T_{x+x'} M_{\omega+\omega'}$
$\Rightarrow$ (reduced) Heisenberg group with group multiplication:

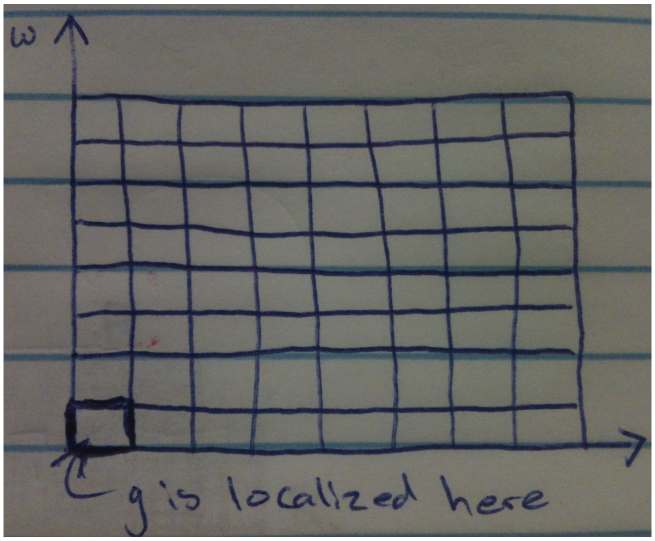
$$(x, \omega, e^{2\pi i \tau}) \cdot (x', \omega', e^{2\pi i \tau'}) = (x + x', \omega + \omega', e^{2\pi i (\tau + \tau')} \underbrace{e^{\pi i (x'\omega - x\omega')}}_{\text{symplectic form}})$$

Wavelets: $(T_b D_a)(T_x D_s) = T_{ax+b} D_{as}$, where $b, x \in \mathbb{R}$ and $a, s > 0$. The group multiplication on $\mathbb{R} \times \mathbb{R}^+$ is:

$$(b, a) \cdot (x, s) = (ax + b, as)$$

This generates the "$ax + b$"-group, sometimes called the "affine group."

Gabor systems in the time-frequency plane:
Recall that $g_{kl} = T_{ka} M_{lb} g$.



Wavelets:
Recall $\psi_{kl} = T_{kl} D_{ls} \psi$

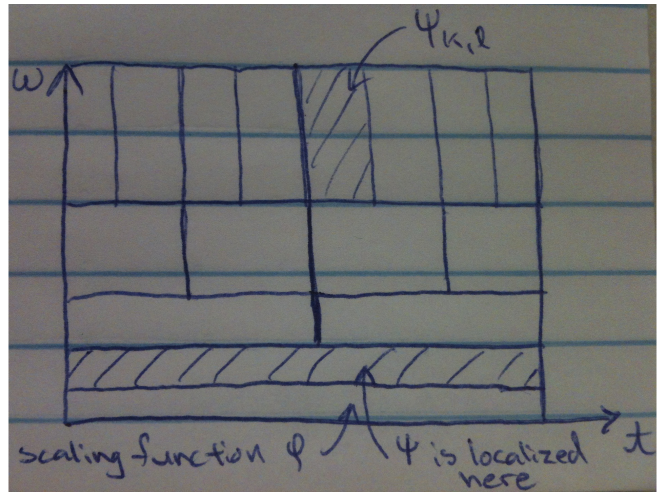


Figure 7: Logarithmic tiling of the time-frequency plane.

15.3 Linear and Nonlinear Approximation

15.3.1 Linear Approximation

Assume we have a Hilbert space $\mathcal{H}$ and an ONB $\{g_k\}_{k\in\mathbb{N}}$ for $\mathcal{H}$. Let $f \in \mathcal{H}$, $f = \sum_{k=1}^{\infty} \langle f, g_k \rangle g_k$. Approximate f by using the first n vectors: $g_1, g_2, \ldots, g_n$. In other words, project onto the subspace $U_n = \text{span}\{g_1, \ldots, g_n\}$.

$$f_n = \sum_{k=1}^{n} \langle f, g_k \rangle g_k$$

Error:

$$f - f_n = \sum_{k=n+1}^{\infty} \langle f, g_k \rangle g_k$$

$$\|f - f_n\|_2^2 = \sum_{k=n+1}^{\infty} |\langle f, g_k \rangle|^2$$

$\Rightarrow$ The approximation error depends on the decay of $\langle f, g_k \rangle$ as n increases.

Example 15.1.

Consider the Fourier ONB $\{e^{2\pi int}\}_{n\in\mathbb{Z}}$ for $L^2[0,1]$.

$$f_n(t) = \sum_{k=-n/2}^{n/2} \hat{f}(k) e^{2\pi ikt}$$

If f is s times differentiable then $\|f - f_n\|_2 = O(n^{-s})$ (Sobolev spaces). This is linear approximation.

15.3.2 Nonlinear Approximation

Define $\Sigma_n = \{p \in \sum_{k \in I} c_k g_k,\ |I| = n\}$. The best n-term approximation error $\sigma_n(f)$ is

$$\sigma_n(f) := \inf_{p \in \Sigma_n} \|f - p\|_2$$

We can find this optimal p easily if $\{g_k\}$ is an ONB. (Otherwise it is very hard!)

Compute $c_k = \langle f, g_k \rangle$ for all k and sort c_k by magnitude in decreasing order: $\to c_{\pi(k)}$, where π is the corresponding permutation of the index set. Then take the first n coefficients of $\{c_{\pi(k)}\}_{k=1}^{n}$ and form p_{opt}:

$$p_{\text{opt}} = \sum_{k=1}^{n} c_{\pi(k)} g_{\pi(k)}.$$

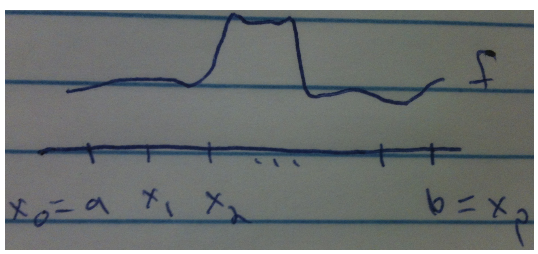
(Compute all c_k, take the n largest by magnitude, ignore the others.)

$\to$ Nonlinear approximation: Σ_n is not a linear subspace, i.e. $\Sigma_n + \Sigma_n \neq \Sigma_n,\ \Sigma_n + \Sigma_n = \Sigma_{2n}$.

We can characterize many function spaces in terms of the best n-term approximation error. If $|\langle f, g_{\pi(k)} \rangle| = O(k^{-\alpha})$, then $\sigma_n(f) = O(n^{-\alpha})$. For L^2-Sobolev W_x^m

- Fourier: $O(n^{-m})$
- Smooth wavelet ONB: $O(n^{-m})$

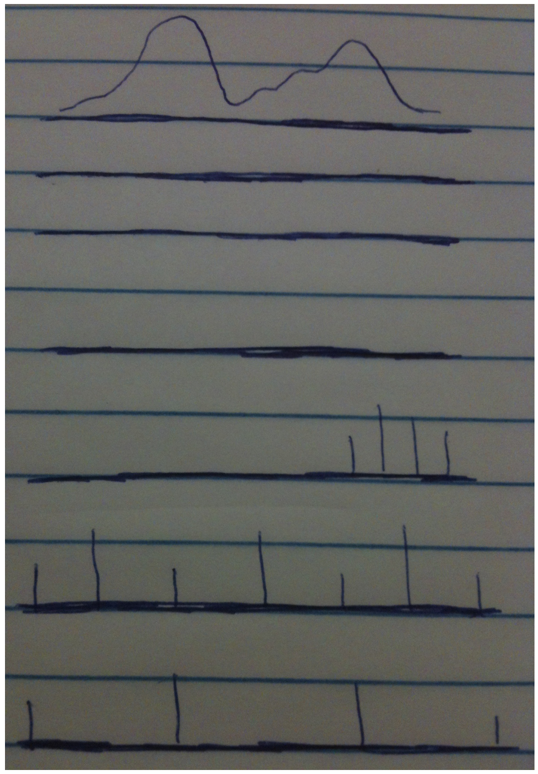
Space of bounded variations: The total variation of f on $[a,b]$ is $V_{[a,b]}(f) = \sup_{p \in \mathcal{P}} \sum_{k=1}^{p} |f(x_{k+1}) - f(x_k)|$, where $\mathcal{P} = \{x_0, \ldots, x_p\}$ is a partition of $[a,b]$.



$f \in BV[a,b] \ \Leftrightarrow\ V_{[a,b]}(f) < \infty$
(for differentiable f: $V_{[a,b]}(f) = \int_{[a,b]} |f'(x)|\, dx$)
BV is useful to model "natural images."

Wavelets (Haar or smooth): $\sigma_n = O(n^{-1})$ for $f \in BV$

Fourier: $O(n^{-1/2})$



16 3-8-12

16.1 Image Compression

Steps:

1. Acquire a high-resolution image, x, by measuring image values at many pixels. Say, $4096 \times 4096 \Rightarrow 2^{24} = m$ pixels.

2. Compress this image by applying an orthogonal transform, such as the wavelet transform: $y = Wx = \{\langle x, \psi_{m,n}\rangle\}_{m,n}$, y is an $m \times 1$ vector.

3. Keep only the N largest wavelet coefficients, y_k, and their corresponding indices. Throw away the rest.

4. Quantization: quantize the y_k's (i.e., represent them by binary numbers).

Reconstruct from the compressed image:
Apply the inverse wavelet transform to $\tilde{y}$, where

$$\tilde{y} = \begin{cases} y_k & k \text{ is a "surviving" index} \\ 0 & \text{otherwise} \end{cases}$$
$$\tilde{x} = W^{-1}\tilde{y}$$

$\tilde{x}$ is the best N-term approximation to x.

Typically, $N \ll m$. So we take m measurements and then throw most of them away. But these measurements can be very expensive: battery life, acquisition time (e.g. MRI), time and money (e.g. biological measurement).

Question: Can we take measurements in compressed form?
We don't know a priori which are the N largest coefficients!
$\Rightarrow$ Compressive sensing shows that we can do this with minimal overhead.

16.2 Compressive Sensing

$$Wx = \{\langle x, \psi_{m,n}\rangle\}_{m,n}$$

Instead of taking the full wavelet transform, we do the following: L: Let A be of size $n \times m$, $n \ll m$, and let it be a Gaussian random matrix ($A_{kl} \sim \mathcal{N}(0,1)$) and form a sensing matrix.

$$\underbrace{B}_{n\times m} := Aw$$

Compute $\underbrace{y}_{n\times 1} = \underbrace{B}_{n\times m}\underbrace{x}_{m\times 1} = AWx$.
Question: Can we recover x from y?

Problem: This system is (highly) underdetermined. There are infinitely many solutions, so how do we get the right one?

Note that Wx is "approximately sparse," meaning that most of the coefficients $\langle x, \psi_{k,l}\rangle$ are small or zero. Set $z := Wx$. Consider

$$Bx = y \quad \Leftrightarrow \quad A\underbrace{Wx}_{z} = y \quad \Leftrightarrow \quad Az = y,$$

where z is sparse. We try to solve $Az = y$, and get x from $x = W^{-1}z$. But A is $n \times m$, so $Az = y$ is also underdetermined. But we know something about the solution: z is sparse. So we look for the sparsest solution.

Note: we are shifting the burden from the measurement part (hardware) to the reconstruction part (software).

Key: How can we find the sparsest solution to $Az = y$?

16.3 Redundancy of Frames Revisited

Let $\{f_k\}_{k=1}^m$ be a frame for $\mathbb{C}^n$, $n \leq m$. Any $f \in \mathbb{C}^n$ can be written as $f = \sum c_k f_k$ for appropriate coefficients c_k. The coefficients are not unique.

Canonical choice: $c_k = \langle f, g_k \rangle$, where $\{g_k\}$ is the canonical dual frame.

Write $F = \begin{bmatrix} f_1 & f_2 & \cdots & f_m \end{bmatrix}$, F is an $n \times m$ matrix. Consider $f = \sum c_k f_k = Fc$, $c = \{c_k\}_{k=1}^m$. The system $Fc = f$ is underdetermined. We can try to enforce uniqueness by regularization: introduce constraints on c.

Standard constraint: minimize the "energy" of $c \Leftrightarrow$ minimize $\|c\|_2$. That is,

$$\min_c \|c\|_2 \quad \text{such that } Fc = f$$

$$\text{equivalent form:} \qquad \min \|Fc - f\|_2 + \lambda \|c\|_2$$

This second form is called *Tykhonov regularization*. There is a λ such that both solutions coincide.

The canonical dual frame coefficients have minimal ℓ^2-norm.

Instead of $\min \|c\|_2$ such that $Fc = y$, we can consider $\min \|c\|_p$ such that $Fc = y$.

16.3.1 Compressive Sensing Setup

Consider $Ax = y$, where A is $n \times m$ and $n \leq m$ (and possibly $n \ll m$). Given A and y, we want to find x. Assumption: x has few nonzero entries. Problem: we don't know the locations of these nonzero entries.

Definition 16.1. ***"Zero-Norm" (not really a norm)***

$$\|x\|_0 := \text{nnz}(x)$$

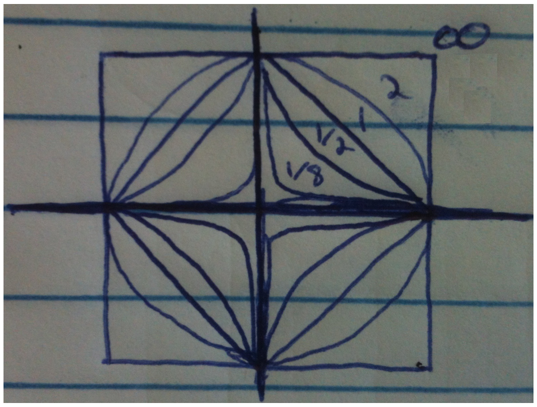
Definition 16.2. ***ℓ^p-norm***

$$\|x\|_p = \left(\sum |x_k|^p\right)^{1/p}$$

This is a norm for $1 \leq p \leq \infty$.

$$B_{p,\alpha} = \{x \in \mathbb{C}^n \mid \|x\|_p \leq \alpha\}$$

If $\alpha = 1$, we write B_p.



Consider $x = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{C}^n$. Then $\|x\|_2 = 1,\ \|x\|_1 = 1$.
Consider $y = \begin{bmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \end{bmatrix} \in \mathbb{C}^n$. Then $\|y\|_2 = 1,\ \|y\|_1 = \sqrt{n}$.
$\Rightarrow$ The ℓ^1-norm "favors" sparsity.

Example 16.3.

$$A = \begin{bmatrix} 2 & 3 \end{bmatrix}, \quad x = \begin{bmatrix} 0 & \frac{4}{3} \end{bmatrix}, \quad y = Ax = 4$$

There are infinitely many solutions $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. All solutions satisfy $A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1 + 3x_2 = 4$.

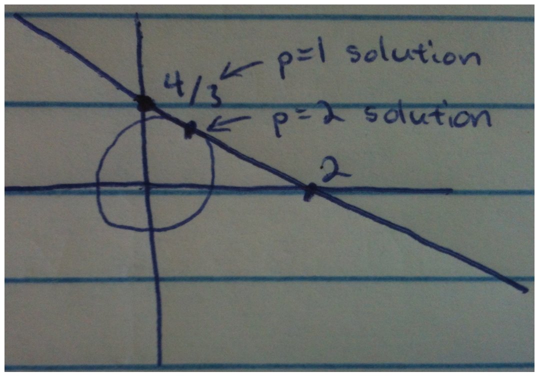


Figure 8: All solutions to $Ax = y$ lie on this subspace.

We can solve $\min \|x\|_p$ such that $Ax = y$ for $p \geq 1$. If we solve $\min \|x\|_1$ such that $Ax = y$, then we find the correct sparse solution. The "pointiness" of the B_1 ball "promotes" sparsity. Furthermore, the B_1 ball becomes even more pointy in higher dimensions. Why not use the $p = 1/2$ ball? Because it yields a non-convex problem.

Given $Ax = y$, where A is $n \times m$, $n \leq m$, and x is sparse. We could try to solve

$$\min_x \|x\|_0 \quad \text{such that } Ax = y \quad \Leftrightarrow \quad \text{find the sparsest among all possible solutions}$$

Problem: solving this problem is NP-hard.

Assume $\|x\|_0 = k (= \text{nnz}(x))$.

$$\left[\ \hat{A}\ \middle|\ X\ \right] \begin{bmatrix} \hat{x} \\ \hline 0 \\ \vdots \\ 0 \end{bmatrix} = y$$

$$\hat{A}\hat{x} = y$$

There are $\binom{m}{k}$ possibilities to place the k coefficients into a vector of length m.

$$\binom{m}{k} \approx e^{m \log(m/k)}$$

17 3-13-12

17.1 Compressive Sensing (Continued)

We are trying to find

$$\min \|x\|_1 \qquad \text{such that} \quad Ax = y.$$

We are considering the L_1 norm as a convex relaxation of the L_0 "norm."

Example 17.1.

Let $m = 100,\ n = 40,\ \|x\|_0 = 15$. Solving

$$\min \|x\|_0 \qquad \text{such that} \quad Ax = y$$

on a 10 GHz computer takes 1 year. Solving

$$\min \|x\|_1 \qquad \text{such that} \quad Ax = y$$

takes 0.1 seconds.

Questions:

1. When does the solution of the L_1 problem coincide with the solution of the L_0 problem?
   - There are problems when the L_0 and L_1 solutions do not coincide. This is due to properties of A.
2. How can we solve the L_1 problem fast?

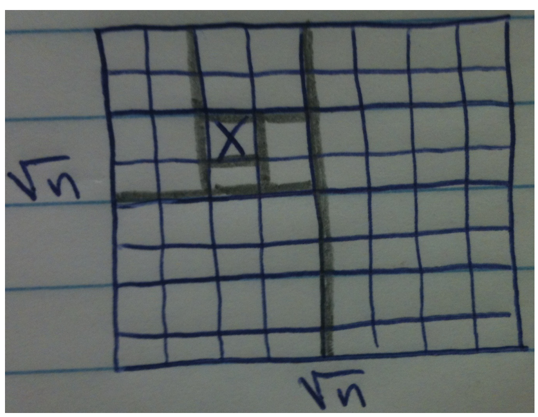


Figure 9: Find the star in the night sky.

Taking measurements pixel by pixel, it will take on average $\frac{n}{2}$ measurements to find the pixel containing the star. Taking "adaptive measurements" takes $\sim \log_2 n$ measurements. However, in many applications we cannot take adaptive measurements. (For example, if we have two stars, one $+$ and one $-$. We could miss them completely due to cancellation.)

Consider $x \in \mathbb{C}^m$ which has only one nonzero entry at k_*:

$$x(k_*) = \alpha$$
$$x(k) = 0, \qquad k \neq k_*.$$

But we don't know k_*. Can we find k_* with less than $O(m)$ nonadaptive measurements? Choose "sensing matrix" $A \in \mathbb{C}^{n\times m}$ with $n < m$ (maybe $n \ll m$) and measure $y = Ax$, $y \in \mathbb{C}^n$. Can we recover x from y?

Let

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix}, \qquad a_k \in \mathbb{C}^n, \quad \|a_k\|_2 = 1 \ \forall\ k.$$

Remember, this is a highly underdetermined system so we cannot invert A. Compute

$$z = A^*y = A^*Ax = A^*a_{k_*}\alpha = \{\langle a_k, a_{k_*}\alpha\rangle\}_{k=1}^m.$$

Look at

$$\max_k |z_k| = \max_{k=1,\dots,m} |\langle a_k, \alpha a_{k_*}\rangle| = \alpha \max_{k=1,\dots,m} |\langle a_k, a_{k_*}\rangle|$$

If

$$|\langle a_k, a_{k_l}\rangle| \begin{cases} = 1 & k = l \\ < 1 & k \neq l \end{cases},$$

then $\max |A^*y|$ will be at k_*.

But...

- what if we have noise?
- what if there is more than 1 nonzero entry?

We look at the *coherence* of A:

$$\mu(A) = \max_{k\neq l} \frac{|\langle a_k, a_l\rangle|}{\|a_k\|_2\|a_l\|_2}.$$

Theorem 17.2. ***Donahue (2001)***

Suppose $Ax = y$, where $A \in \mathbb{C}^{n\times m}$, $n < m$, and rank $A = n$. If a solution x exists with $\|x\|_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(A)}\right)$, then x is the unique solution for the L_0 problem and the L_1 problem.

Example 17.3.

Let $A = \begin{bmatrix} I_n & F_n \end{bmatrix}$, where F_n is the $n \times n$ DFT matrix. Then

$$A^*A = \begin{bmatrix} I_n \\ F_n^* \end{bmatrix} \begin{bmatrix} I_n & F_n \end{bmatrix} = \begin{bmatrix} I_n & F_n \\ F_n^* & I_n \end{bmatrix}$$

$$\mu(A) = \frac{1}{\sqrt{n}}$$

This is the smallest covariance we can have when we form a matrix from 2 orthonormal bases.

From the theorem, we can recover any x with sparsity $\|x\|_0 < \frac{1}{2}(1 + \sqrt{n})$. So if $n = 256$, then $\mu(A) = \frac{1}{16}$ and the allowed sparsity is ≤ 8.

Examples of matrices with small μ: equiangular tight frames, MUB (mutually unbiased bases).

Simulations show that this L_0-L_1 equivalence holds way beyond the $\sqrt{n}$ threshold for $\|x\|_0$. CVX-software: easy to implement convex optimization problems in Matlab (Boyd).

$$\min \|x\|_1 \qquad \text{such that} \quad Ax = y,\ x \geq 0$$

Definition 17.4. ***Restricted Isometry Property (RIP)***

An $n \times m$ matrix A has the *restricted isometry property* (with constants δ, l) if every submatrix A_I, $I \subset \{1, \ldots, m\}$, formed by at most l columns of A satisfies

$$\|z\|_2^2(1-\delta) \leq \|A_I z\|_2^2 \leq (1+\delta)\|z\|_2^2 \qquad \forall\ z = \{z_k\}_{k \in I}$$

for some $\delta > 0$. (An isometry would give us $\|A_I z\|_2 = \|z\|_2$.)

For $l = 2$, this is equivalent to the coherence version.

Theorem 17.5. ***Condes, Tao (2004)***

($\delta \geq 0.41$) If a matrix A satisfies the RIP$(0.41, 2k)$, then the L_0 and L_1 problems have identical and unique solutions for all k-sparse x.

But how do we check the RIP property? Checking it directly: $\binom{m}{2k}$ $\Rightarrow$ NP-hard.

Examples of matrices that satisfy the RIP with high probability:

- A is a Gaussian random matrix: $A_{kl} \sim \mathcal{N}(0,1)$ (`randn(n,m)`). Then A satisfies the RIP (with high probability) if $k < c\left(1 + \log \frac{m}{n}\right) \cdot n$, where c is a small constant. Or, if $n \geq ck \log \frac{m}{n}$, then A satisfies the RIP.
- Choose F = DFT matrix, randomly choose n rows of F and put them into A. In this case, we need $n \geq ck \log^4 m$.

Given $x \in \mathbb{C}^m$, let x_s be the best s-term approximation to x. That is, x_s coincides with x at the s largest entries of x and is zero elsewhere. We call x *compressible* if $\|x - x_s\| \approx s^{-p}$ for $p > 1$.

Consider $y = Ax + w$, where w is noise.

Theorem 17.6. ***Condes (2008)***

Assume $A \in \mathbb{C}^{n \times m}$ satisfies the RIP with $\delta_{2k} < \sqrt{2} - 1$. Then the solution x_* to

$$\min \|x\|_1 \qquad \text{such that} \quad \|Ax - y\| \leq \epsilon \geq \|w\|_2 \qquad \text{(Noisy } L_1 \text{ problem)}$$

obeys

$$\|x_* - x\|_2 \leq \frac{c_0}{\sqrt{k}}\|x - x_k\|_1 + c_1\epsilon,$$

where c_0, c_1 are numerical constants and x_k is the best k-term approximation to x.

18 3-15-12

18.1 Greedy Algorithms

We are looking at problems of the form min $\|x\|_0$ such that $Ax = b \Rightarrow$ min $\|x\|_1$ such that $Ax = b$.

Greedy algorithms are recursive algorithms where in each iteration step we compute $A^* b_k$, and we look for the largest r entries of $|A^* b_k|$. We "subtract" the corresponding nonzero entries in x from the RHS b_k and compute the new RHS b_{k+1}. They are not as good as true L_1 minimization.

LASSO: min $\|x\|_1 + \lambda \|Ax - b\|_2^2$
L_1-min: min $\|x\|_1$ such that $\|Ax - b\|_2 \leq \epsilon$

There exists λ such that LASSO and L_1 minimization are equivalent.

In Matlab, the standard solvers for L_1 minimization are based on

- Interior Point methods
- gradient descent methods
- augmented Lagrange multipliers
- primal-dual methods

Definition 18.1. ***Soft-Thresholding, Hard-Thresholding***

For $z \in \mathbb{C}^n$, the *soft-thresholding* operator, S_τ, is defined by:

$$S_\tau(z) = \begin{cases} \frac{z_k}{|z_k|}(|z_k| - \tau) & |z_k| > \tau \\ 0 & \text{otherwise} \end{cases}$$

The *hard-thresholding* operator is defined by

$$H_\tau(z) = \begin{cases} z_k & |z_k| > \tau \\ 0 & \text{otherwise} \end{cases}$$

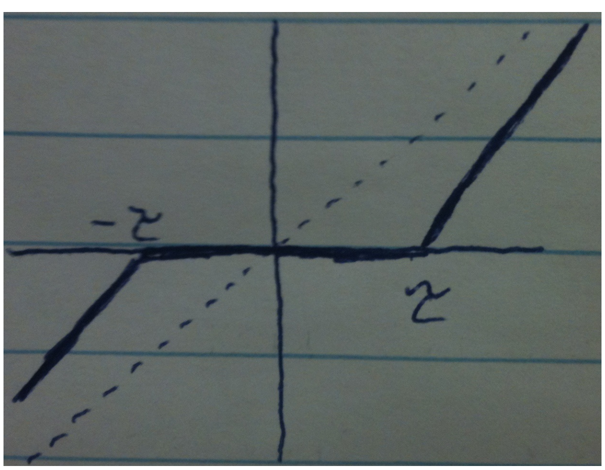


Figure 10: Soft-thresholding.

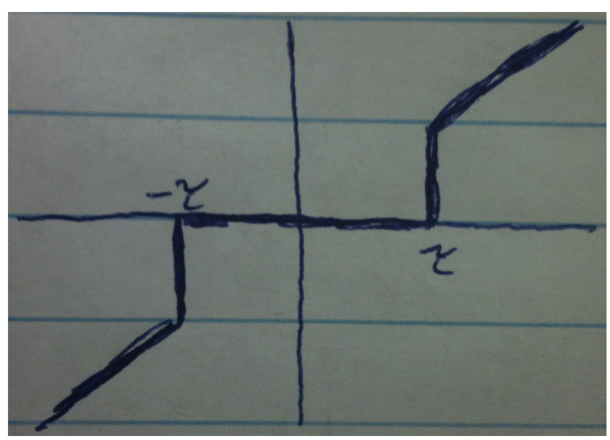


Figure 11: Hard-thresholding.

Remark 18.2. ***Chamboille Algorithm***

- Choose x_0, ξ_0 as initial guesses. Let $\theta \in [0,1]$, $\tau, \sigma > 0$ with $\tau\sigma\|A\|_{\text{op}} < 1$.
- Compute iteratively:

$$\begin{aligned}
\xi_{n+1} &= \xi_n + \sigma(A\overline{x}_n - y)\\
x_{n+1} &= S_\tau(x_n - \tau A^* \xi_{n+1})\\
\overline{x}_{n+1} &= x_{n+1} + \theta(x_{n+1} - x_n)
\end{aligned}$$

This solves $\min \|x\|_1$ such that $Ax = b$.

18.2 Sparse Representations and Image Compression

Assumption: "natural" images consist of edges, contours, and texture.

- Edges/contours: modelled very well by wavelets or curvelets
- Texture: modelled very well by Gabor systems, block-DCT (JPEG)

Let A_1 be an orthonormal basis consisting of wavelets.
Let A_2 be an orthonormal basis consisting of an orthonormal Gabor system.

Classical image compression: pick one ONB, say A_1, and compute $y = A_1^* x$.

Instead: consider $A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$. This is a tight frame, since it is the union of 2 ONB's. We could compute $y = A^* x$ and threshold. We *should* solve:

$$\min \|y\|_1 \qquad \text{such that } Ay = x. \tag{18.1}$$

Let $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Then we can write (18.1) as

$$\min \left\| \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\| \qquad \text{such that } \begin{bmatrix} A_1 & A_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x.$$

This will ideally force the edges and contours into y_1 and the textures into y_2.

If our image has more qualities, e.g. "pointy," then we can add a 3rd ONB in order to capture it.

18.3 Matrix Completion

Given an $n \times n$ matrix X with $\text{rank}(X) \ll n$ (X has low rank). Assume we know only a few entries of X: X_{kj} for some $(k, j) \subset \{(1,1),\ (1,2),\ \ldots,\ (n,n)\}$. Question: can we recover x from these x_{kj}?

Example 18.3.

$$X = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

X has rank 1. In this case, we cannot recover X from X_{kj} for some $(k, j) \subset \{(1,1),\ (1,2),\ \ldots,\ (n,n)\}$.

Example 18.4.

$$X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

X has rank 1. In this case, we can recover X from X_{kj} for some $(k, j) \subset \{(1,1),\ (1,2),\ \ldots,\ (n,n)\}$.

Assume we are given $y = \mathcal{A}(X)$, where $\mathcal{A}$ is a linear mapping from $\mathbb{C}^{n\times n} \to \mathbb{C}^m,\ m \ll n^2$. We have m linear measurements of the $n \times n$ matrix X, and X has low rank.

Goal: recover X.

We could do this by solving the following the optimization problem:

$$\min \operatorname{rank}(X) \qquad \text{such that } \mathcal{A}(X) = y.$$

(Compare to $\min \|x\|_0$ such that $Ax = y \Leftrightarrow \min \|x\|_1$ such that $Ax = y$, from earlier.)
However, in practice this is impossible to solve (NP-hard). All known algorithms have double-exponential complexity.

For example, Netflix movie rankings. Rows correspond to users, columns correspond to movies.

- Some users rate a lot of movies $\Rightarrow$ dense rows.
- Some movies are rated a lot $\Rightarrow$ dense columns.
- Most of the matrix is not filled in.

Definition 18.5. ***Schatten-p-Norm, Nuclear Norm***

Define the *Schatten-p-norm* of X as follows:
Let $\sigma_k,\ k = 1, \dots, n$, be the singular values of X.

$$\|X\|_p = \left(\sum_{k=1}^{n} |\sigma_k|^p \right)^{1/p}$$

For $p = 1$, this is sometimes called the "*nuclear norm*" and is denoted $\|X\|_*$ $(= \|X\|_1)$. So

$$\|X\|_* = \sum_{k=1}^{n} |\sigma_k|$$

$$\|X\|_0 = \operatorname{rank}(X)$$

Consider the convex relaxation of the rank minimization problem:

$$\min \|X\|_* \qquad \text{such that } \mathcal{A}(X) = y.$$

This is a semidefinite program. Many results from compressive sensing carry over to matrix completion.

Definition 18.6. ***Matrix-RIP***

Let $\mathcal{A} : \mathbb{R}^{n \times n} \to \mathbb{R}^m$. For every integer r with $1 \le r \le m$, we define the restricted isometry property as follows:

$$(1 - \delta_r)\|X\|_{\mathrm{F}} \le \|\mathcal{A}(X)\|_{\mathrm{F}} \le (1 + \delta_r)\|X\|_{\mathrm{F}},$$

where $\operatorname{rank}(X) \le r$.

Index