6.342 Lecture 9 — March 2, 2011

Today:
• Time and frequency spread
• Uncertainty principles
• Finite sequence recovery based on time-frequency sparsity

Next time:
• Compressed sensing

Reading:
• Chapter 6

Time and Frequency

(a) Equal time divisions at all frequencies – like STFT

(b) Finer time divisions at lower frequencies – possible w/wavelet packets

(c) Finer time divisions at higher frequencies – like DWT
\textbf{Definition 3.8 (Fourier transform)} The Fourier transform of a function \( x(t) \) is
\[
X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} \, dt, \quad \omega \in \mathbb{R}.
\] (3.43a)

It exists when (3.43a) converges for all \( \omega \in \mathbb{R} \); we then call it the \textit{spectrum} of \( x \).

The inverse Fourier transform of \( X(\omega) \) is
\[
x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)e^{j\omega t} \, d\omega, \quad t \in \mathbb{R}.
\] (3.43b)

When the Fourier transform exists, we denote the Fourier-transform pair as
\[
x(t) \xrightarrow{\text{FT}} X(\omega).
\]

\textbf{Definition 6.1 (Time center and spread for functions)} Given a function \( x(t) \in L^2(\mathbb{R}) \) with norm \( E_x = \|x\|^2 \), its time center \( \mu_t \) is
\[
\mu_t = \frac{1}{E_x} \int_{t \in \mathbb{R}} t|x(t)|^2 \, dt, \quad (6.1a)
\]

and its time spread \( \Delta_t \),
\[
\Delta_t^2 = \frac{1}{E_x} \int_{t \in \mathbb{R}} (t - \mu_t)^2|x(t)|^2 \, dt. \quad (6.1b)
\]
DEFINITION 6.2 (Frequency Center and Spread for Functions) Given a function \( f \in L^2(\mathbb{R}) \) with the Fourier transform \( X(\omega) \) of norm \( E_{\omega} = \| X \|^2 \), its frequency center \( \mu_\omega \) is

\[
\mu_\omega = \frac{1}{2\pi E_{\omega}} \int_{\omega \in \mathbb{R}} \omega |X(\omega)|^2 d\omega,
\]

and its frequency spread \( \Delta_\omega \),

\[
\Delta^2_\omega = \frac{1}{2\pi E_{\omega}} \int_{\omega \in \mathbb{R}} (\omega - \mu_\omega)^2 |X(\omega)|^2 d\omega.
\]

Heisenberg boxes

A function \( f \in L^2(\mathbb{R}) \) has a four-tuple \( (\mu_t, \Delta_t, \mu_\omega, \Delta_\omega) \). Its Heisenberg box is centered at \( (\mu_t, \mu_\omega) \) with dimensions \( \Delta_t \)-by-\( \Delta_\omega \).
<table>
<thead>
<tr>
<th>Function</th>
<th>Time Center</th>
<th>Time Spread</th>
<th>Freq. Center</th>
<th>Freq. Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x(t)$</td>
<td>$\mu_t$</td>
<td>$\Delta_t$</td>
<td>$\mu_\omega$</td>
<td>$\Delta_\omega$</td>
</tr>
<tr>
<td>$x(t-T)$</td>
<td>$\mu_t + T$</td>
<td>$\Delta_t$</td>
<td>$\mu_\omega$</td>
<td>$\Delta_\omega$</td>
</tr>
<tr>
<td>$e^{j\Omega t}x(t)$</td>
<td>$\mu_t$</td>
<td>$\Delta_t$</td>
<td>$\mu_\omega + \Omega$</td>
<td>$\Delta_\omega$</td>
</tr>
<tr>
<td>$x(t/\alpha)/\sqrt{\alpha}$</td>
<td>$\mu_t/\alpha$</td>
<td>$\Delta_t/\alpha$</td>
<td>$\mu_\omega/\alpha$</td>
<td>$\Delta_\omega/\alpha$</td>
</tr>
</tbody>
</table>

Table 6.1: Effect of a shift, modulation and scaling on a Heisenberg box ($\mu_t, \Delta_t, \mu_\omega, \Delta_\omega$)

![Figure 6.2](image)

Figure 6.2. Elementary operations on a basis function $f$ and their effects on a time-frequency tile. (a) Shift in time by $\tau$ producing $f'$ and modulation by $\omega_0$ producing $f''$. (b) Scaling $f'(t) = f(at)$ ($a = 1/3$ is shown).

**Theorem 6.3 (Uncertainty Principle)** Given a function $x \in L^2(\mathbb{R})$, the product of its squared time and frequency spreads is lower bounded as

$$\Delta_t^2 \Delta_\omega^2 \geq \frac{1}{4}. \quad (6.8)$$

The lower bound is attained by Gaussian functions from

$$x(t) = \gamma e^{-\alpha t^2}, \quad \alpha > 0. \quad (6.9)$$
6.342 Wavelets, Approximation, and Compression

March 2, 2011

**Lecture 9**

**Definition 6.4 (Time Center and Spread for Sequences)**

Given a sequence $x_n \in \ell^2(\mathbb{Z})$ with norm $E_x = \|x\|^2$, its time center $\mu_n$ is

$$\mu_n = \frac{1}{E_x} \sum_{n \in \mathbb{Z}} n |x_n|^2$$

and its time spread $\Delta_n$.

**Definition 6.5 (Frequency Center and Spread for Sequences)**

Given a sequence $x_n \in \ell^2(\mathbb{Z})$ with the discrete-time Fourier transform $X(e^{i\omega})$ of norm $E_\omega = \|X\|^2$, its frequency center $\rho_\omega$ is

$$\rho_\omega = \frac{1}{2\pi E_\omega} \int_{-\pi}^{\pi} \omega |X(e^{i\omega})|^2 \, d\omega,$$

and its frequency spread $\Delta_\omega$.

$$\Delta_\omega^2 = \frac{1}{2\pi E_\omega} \int_{-\pi}^{\pi} (\omega - \rho_\omega)^2 |X(e^{i\omega})|^2 \, d\omega.$$  

**Theorem 6.6 (Discrete-time Uncertainty Principle)**

Given a sequence $x \in \ell^2(\mathbb{Z})$ with $X(e^{i\omega}) = 0$, the product of its squared time and frequency spreads is lower bounded as

$$\Delta_n^2 \Delta_\omega^2 \geq \frac{1}{4}.$$  

---

**Uncertainty principle via (approximate) support**

We say $f$ is $\epsilon$-concentrated on $T$ when there is a function $g$ supported on $T$ such that $\|f - g\| \leq \epsilon$.

Let $\|f\|_{L^2(\mathbb{R})} = 1$.

Suppose $f$ is $\epsilon_T$-concentrated on $T$ and $f'$ is $\epsilon_W$-concentrated on $W$. Then

$$|W| \cdot |T| \geq (1 - (\epsilon_T + \epsilon_W))^2.$$

($T$ and $W$ could be unions of intervals, not just intervals.)
Finite-Dimensional Uncertainty Principle

Let \( \{x_n\}_{n=0}^{N-1} \) have DFT \( \{X_k\}_{k=0}^{N-1} \).
Let \( N_t \) be the number of nonzero \( x_n \)s.
Let \( N_\omega \) be the number of nonzero \( X_k \)s.

Theorem [Donoho & Stark (1989)]:

For any nonzero signal, \( N_t N_\omega \geq N \)

Corollary: \( N_t + N_\omega \geq 2\sqrt{N} \)

Proof of theorem

Use the following lemma (proved afterward).

Lemma: If \( \{x_n\}_{n=0}^{N-1} \) has \( N_t \) nonzero elements,
then \( \{X_k\}_{k=0}^{N-1} \) cannot have \( N_t \) consecutive zeros.
(Interpret “consecutive” modulo \( N \).)

Suppose \( \{x_n\}_{n=0}^{N-1} \) has \( N_t \) nonzero elements.
Using the lemma, each interval of length \( N/N_t \)
must contain at least one nonzero \( X_k \).
Thus \( N_\omega \geq N/N_t \).
Proof of lemma

Let \( \{x_n\}_{n=\tau_1, \ldots, \tau_{N_t}} \) be nonzeros and let \( b_j = x_{\tau_j} \). We'll show \( X_k \) can't be zero for \( k = m+1, m+2, \ldots, m+N_t \). Define

\[
g_k = X_{k+m} = \sum_{j=1}^{N_t} b_j W_{N}^{\tau_j(m+k)}, \quad k = 1, 2, \ldots, N_t.
\]

We have a matrix equation \( g = Zb \) where \( Z_{kj} = W_{N}^{\tau_j(m+k)} \).

Since \( Z \) is nonsingular and \( b \neq 0 \), \( g \) must have a nonzero.

Proof of corollary

Note the arithmetic-geometric mean inequality:

\[
\frac{1}{2} (a + b) \geq (a \cdot b)^{1/2}.
\]

So \( \frac{1}{2}(N_{l} + N_{\omega}) \geq (N_{l}N_{\omega})^{1/2} \geq \sqrt{N} \). Thm.
Ramification [Donoho & Stark (1989)]

Easy: Suppose true signal \( \{ s_n \}_{n=0}^{N-1} \) has DFT-domain support of \( N_\omega \). Observations of \( s_n \) are made in the time domain, with only \( M < N \) observations available. Unique exact recovery is always possible when 
\[ 2(N - M)N_\omega < N. \]

Harder: If noisy observations \( s_n + \eta_n \) are made, with \( K \) missing and \( \| \eta \| \leq \epsilon \), recovery of estimate \( \hat{s} \) with 
\[ \| s - \hat{s} \| \leq 2 \left( 1 - \frac{2KN_\omega}{N} \right)^{-1/2} \epsilon \]
is possible.

Finite-Dimensional Uncertainty Principles

Let \( \{ x_n \}_{n=0}^{N-1} \) have DFT \( \{ X_k \}_{k=0}^{N-1} \).
Let \( N_t \) be the number of nonzero \( x_n \)s.
Let \( N_\omega \) be the number of nonzero \( X_k \)s.

Theorem [Donoho & Stark (1989)]:

For any nonzero signal, \( N_t N_\omega \geq N \).

Corollary: \( N_t + N_\omega \geq 2\sqrt{N} \)

- Sparsity simultaneously in two different bases may be limited
- Is there something special about (standard,DFT) basis pair?
Theorem [Elad & Bruckstein (2002)]:
Suppose nonzero \( x \in \mathbb{R}^N \) has representations
\( x = \Phi \alpha \) and \( x = \Psi \beta \) with
orthonormal bases (matrices) \( \Phi \) and \( \Psi \).
Then \( \|\alpha\|_0 \|\beta\|_0 \geq (\mu(\Phi, \Psi))^{-2} \), where
\( \mu(\Phi, \Psi) = \max_{i,j} |\langle \varphi_i, \psi_j \rangle| \)
\( \mu(\text{std, DFT}) = 1/\sqrt{N} \), so this generalizes previous result
\( \mu(\Phi, \Psi) \geq 1/\sqrt{N} \), so (std,DFT) is a “best” pair

- Sparsity simultaneously in two different bases may be limited
- Is there something special about (standard,DFT) basis pair? Yes

Ramification [Donoho & Stark (1989)]

Suppose true signal \( \{s_n\}_{n=0}^{N-1} \) has DFT-domain support of \( N_w \). Observations of \( s_n \) are made in the time domain, with only \( M < N \) observations available.
Unique exact recovery is always possible when
\( 2(N - M)N_w < N \).
(Also a stability result as seen last time.)

- Sparsity in one basis counteracts missing data when observing in the other basis
- Sparser signal implies fewer observations are needed
Simple $\ell^1$-based recovery [Donoho & Stark (1989)]

Suppose true signal $\{s_n\}_{n=0}^{N-1}$ has DFT-domain support of $N_\omega$. Observations of $s_n$ are made in the time domain, with only $M < N$ observations available: $\tilde{s} = Ps$, where $P$ sets unobserved components to zero. Under the same condition $2(N - M)N_\omega < N$,

$$\hat{s} = \arg\min_{s'} \|s'\|_1 \text{ subject to } Ps' = \tilde{s}$$

recovers $s$ exactly.

- Sometimes, sparsity can be exploited in a computationally-tractable manner

Further result [Donoho & Huo (2001)]

Suppose a signal $\{s_n\}_{n=0}^{N-1}$ has a “sparse” representation when mixing time and frequency domains. Specifically,

$$s_n = \sum_{i=0}^{N-1} a_i \delta_n - i + \sum_{i=0}^{N-1} b_i W_n^i$$

where $T$ of the $a_i$s are nonzero and $W$ of the $b_i$s are nonzero. (It is “sparse” when $T + W < N$.)

If $T + W < \sqrt{N}$, the decomposition is unique.

Furthermore, if $T' + W < \frac{1}{2}\sqrt{N}$, the unique sparsest decomposition can be found with a linear program.