Computational Harmonic Analysis (Wavelet Tutorial)
Part I

Understanding Many Particle Systems with Machine Learning

Tutorials

Matthew Hirn
Michigan State University
Department of Computational Mathematics, Science & Engineering
Department of Mathematics

Sparsity

The Role of Sparsity

- Function $f \in \mathbf{L}^2(\mathbb{R})$
- Dictionary $\mathcal{D} = \{\phi_\gamma\}_\gamma$
- Analysis of f in Φ:

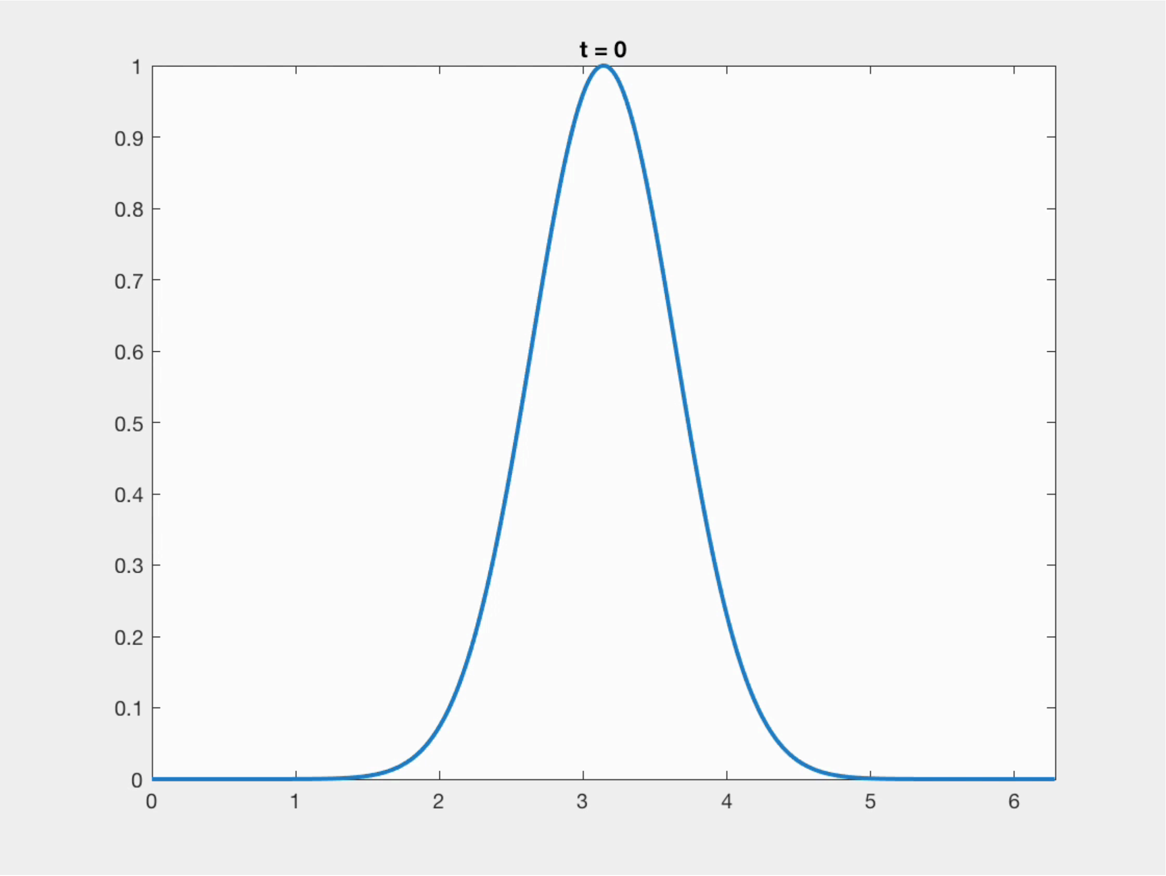
$$\Phi f(\gamma) = \langle f, \phi \rangle$$

- If $\Phi f(\gamma)$ is sparse, it means that the dictionary $\mathcal{D}$ is well adapted to the (regularity) properties of f.
- On the other hand, if $\mathcal{D}$ is not well adapted to f, then it will be distributed over large number of the $\{\phi_\gamma\}_\gamma$ and $\Phi f(\gamma)$ will be non-sparse.

Fourier Analysis

Joseph Fourier and the Heat Equation

$$\partial_t u(x,t) = \Delta_x u(x,t)$$
$$u(x,0) = f(x)$$



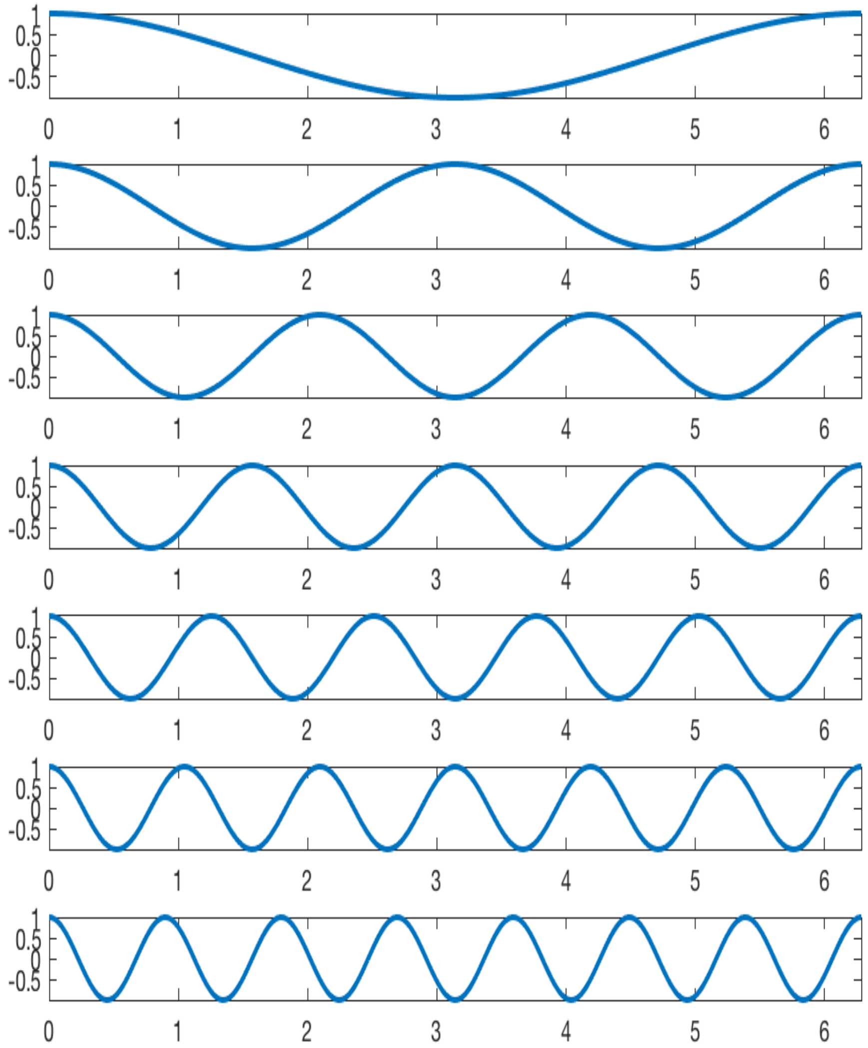
For $x \in [0, 2\pi]$ and $t > 0$:

$$u(x,t) = \sum_n a_n e^{-n^2 t} e^{inx}$$
$$\lim_{t \to 0} u(x,t) = \sum_n a_n e^{inx} = f(x)$$
$$a_n = \text{Fourier coefficients of } f$$

Fourier Transform

- $f \in \mathbf{L}^p(\mathbb{R})$:

$$\|f\|_p = \left(\int_{-\infty}^{\infty} |f(t)|^p \, dt\right)^{\frac{1}{p}} < \infty$$



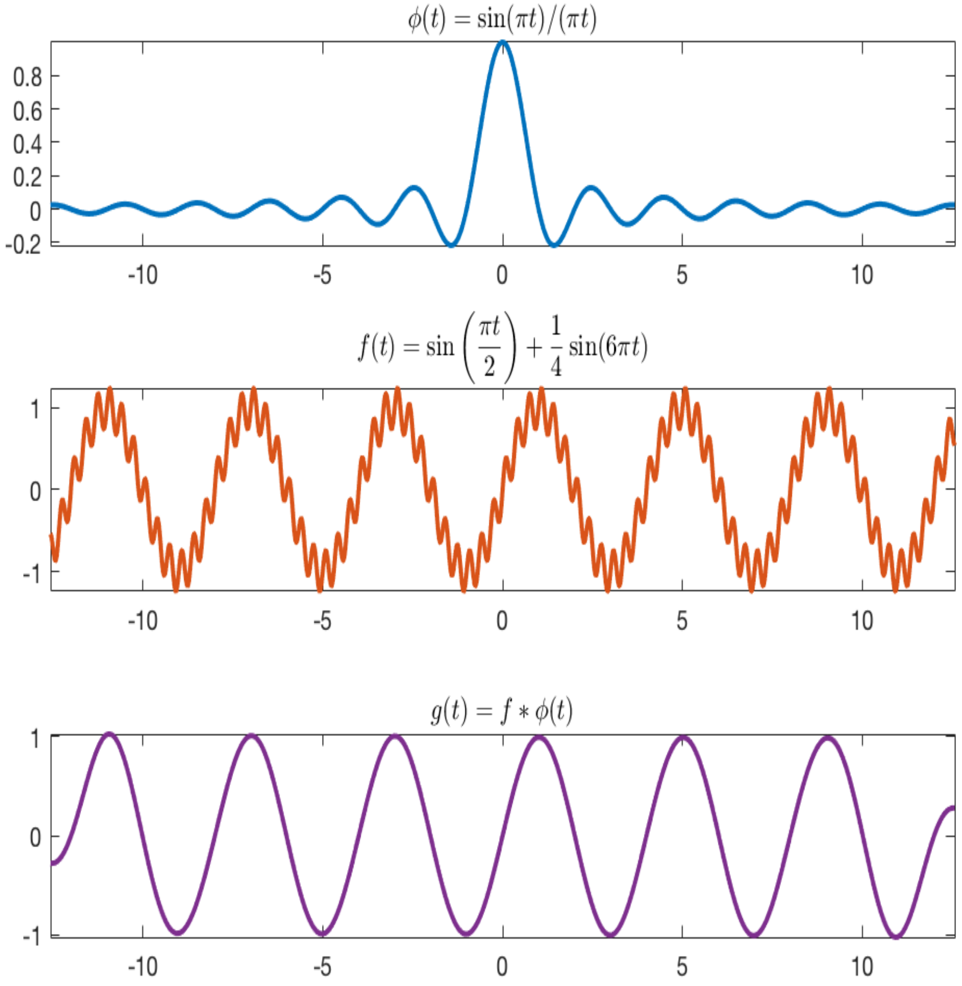
- Fourier transform for $f \in \mathbf{L}^1(\mathbb{R})$:

$$\widehat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} \, dt$$

- Inverse Fourier transform: If $f, \widehat{f} \in \mathbf{L}^1(\mathbb{R})$,

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(\omega) e^{i\omega t} \, dt$$

Ideal Low Pass Filter



- The Fourier transform is extended to $f \in \mathbf{L}^2(\mathbb{R})$ by a "density" argument.

- Ideal low pass filter example:

$$\phi(t) = \frac{\sin(\xi t)}{\pi t} \in \mathbf{L}^2(\mathbb{R}) \setminus \mathbf{L}^1(\mathbb{R})$$

$$\widehat{\phi}(\omega) = \mathbf{1}_{[-\xi,\xi]}(\omega)$$

- Convolution theorem: $f, h \in \mathbf{L}^1(\mathbb{R}) \cup \mathbf{L}^2(\mathbb{R})$,

$$\widehat{f * h}(\omega) = \widehat{f}(\omega)\widehat{h}(\omega)$$

- Filtering f with ϕ retains the low frequency content of f:

$$\widehat{f * \phi}(\omega) = \widehat{f}(\omega)\widehat{\phi}(\omega) = \widehat{f}(\omega)\mathbf{1}_{[-\xi,\xi]}(\omega)$$

Time Invariant Linear Operators

- Time delay of a function of f:

$$f_\tau(t) = f(t - \tau)$$

- Time invariant linear operator L:

$$g(t) = Lf(t) \Rightarrow g_\tau(t) = Lf_\tau(t)$$

- Time invariant linear operators are convolution operators:

$$Lf(t) = f * h(t), \quad h(t) = L\delta(t)$$

- They are thus diagonalized by Fourier:

$$Le^{i\omega t} = \widehat{h}(\omega)e^{i\omega t}$$

- Example: $L = \Delta$ with:

$$\Delta e^{i\omega t} = -\omega^2 e^{i\omega t}$$

Parseval and Plancherel Formulas

- Let $f, h \in \mathbf{L}^2(\mathbb{R})$
- Inner product:

$$\langle f, h \rangle = \int_{-\infty}^{\infty} f(t)\overline{h(t)}\,dt$$

- Parseval formula:

$$\langle f, h \rangle = \frac{1}{2\pi}\langle \hat{f}, \hat{h} \rangle$$

- Plancherel formula (just set $f = h$):

$$\|f\|_2 = \frac{1}{\sqrt{2\pi}}\|\hat{f}\|_2$$

  Fourier transform preserves the energy of f up to a constant factor.

Fourier Transform Properties

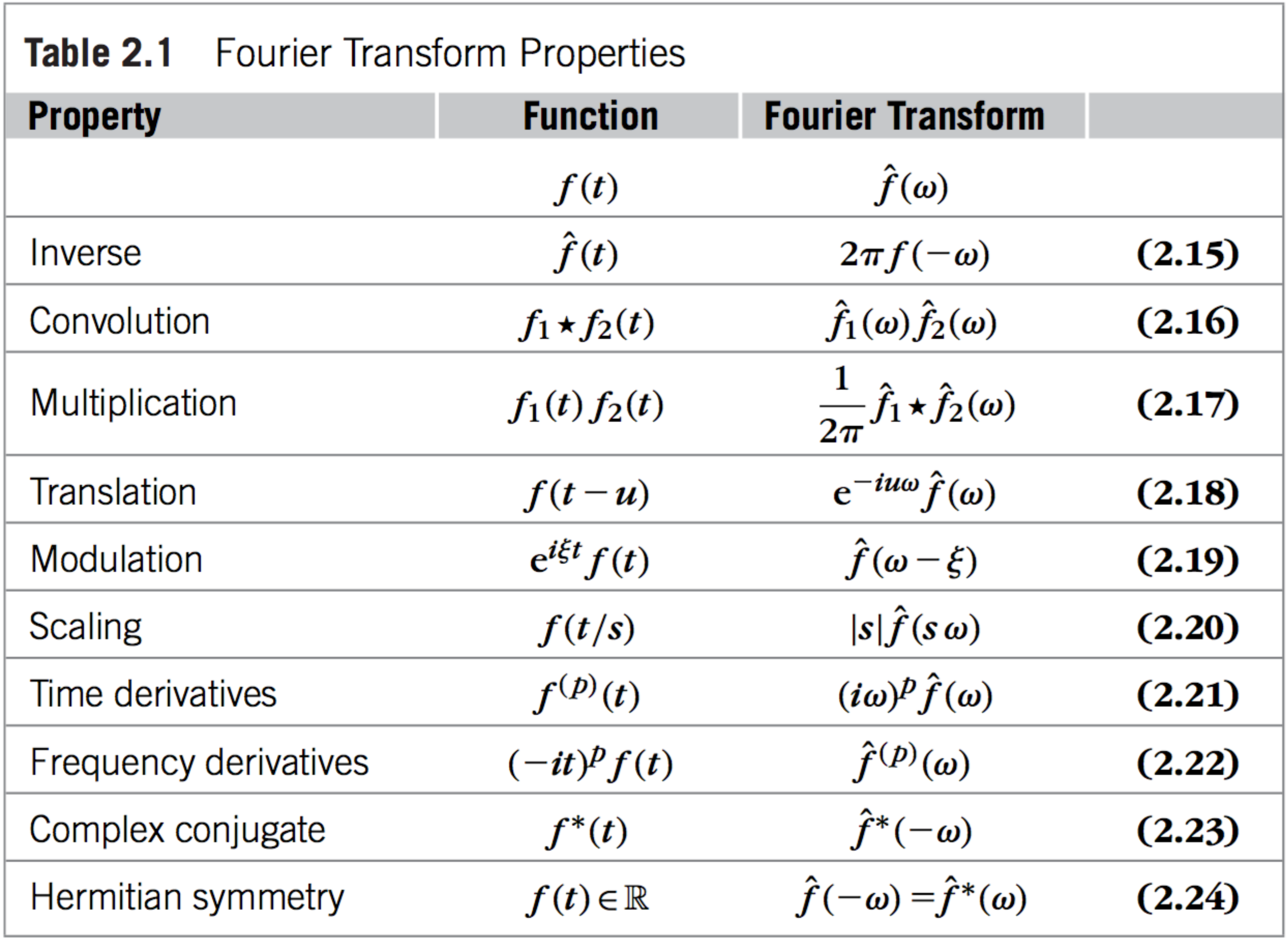


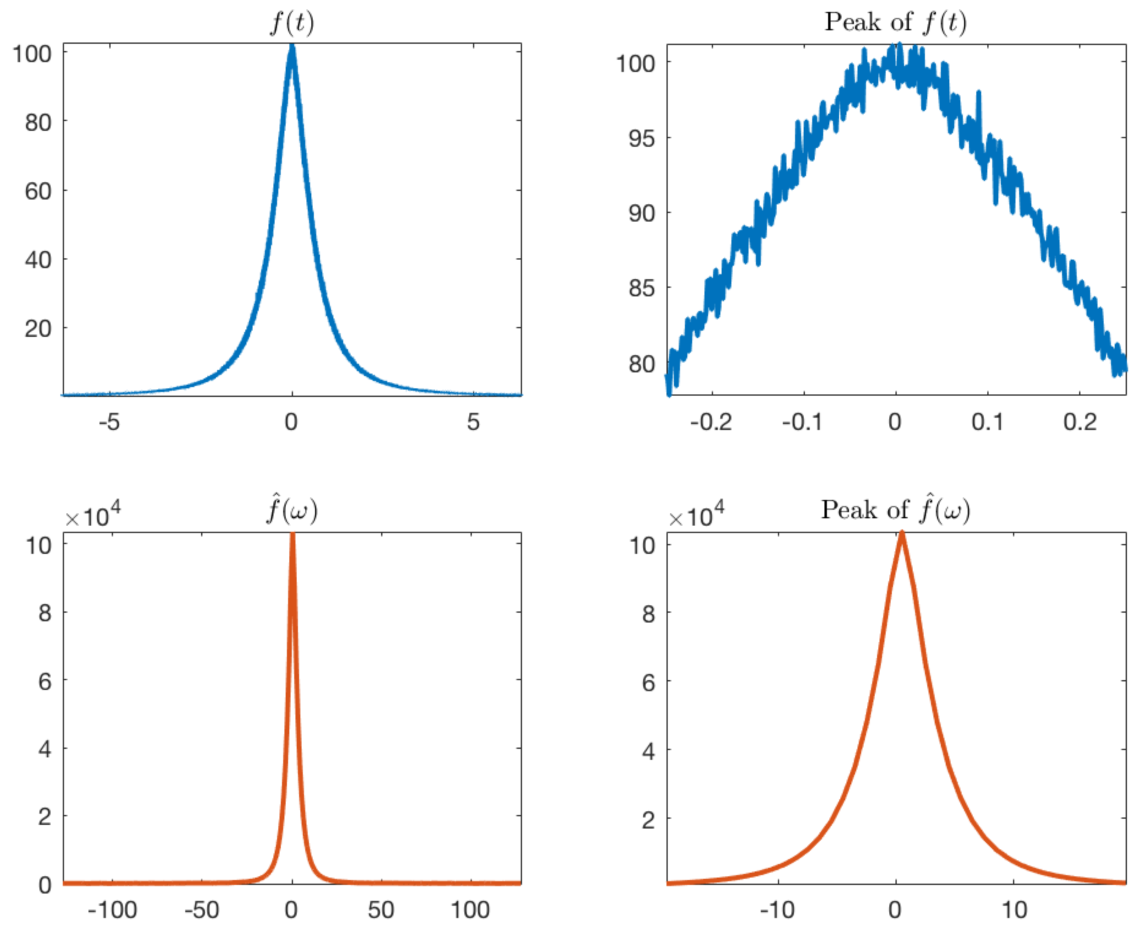
Table 2.1 Fourier Transform Properties

Property	**Function**	**Fourier Transform**	
	$f(t)$	$\hat{f}(\omega)$	
Inverse	$\hat{f}(t)$	$2\pi f(-\omega)$	**(2.15)**
Convolution	$f_1 \star f_2(t)$	$\hat{f}_1(\omega)\hat{f}_2(\omega)$	**(2.16)**
Multiplication	$f_1(t)\, f_2(t)$	$\frac{1}{2\pi}\hat{f}_1 \star \hat{f}_2(\omega)$	**(2.17)**
Translation	$f(t-u)$	$e^{-iu\omega}\hat{f}(\omega)$	**(2.18)**
Modulation	$e^{i\xi t} f(t)$	$\hat{f}(\omega-\xi)$	**(2.19)**
Scaling	$f(t/s)$	$\|s\|\hat{f}(s\,\omega)$	**(2.20)**
Time derivatives	$f^{(p)}(t)$	$(i\omega)^p\hat{f}(\omega)$	**(2.21)**
Frequency derivatives	$(-it)^p f(t)$	$\hat{f}^{(p)}(\omega)$	**(2.22)**
Complex conjugate	$f^*(t)$	$\hat{f}^*(-\omega)$	**(2.23)**
Hermitian symmetry	$f(t) \in \mathbb{R}$	$\hat{f}(-\omega) = \hat{f}^*(\omega)$	**(2.24)**

Continuity and Riemann-Lebesgue

- $f \in \mathbf{L}^1(\mathbb{R}) \Rightarrow \widehat{f}$ is continous
- Riemann-Lebesgue:

$$f \in \mathbf{L}^1(\mathbb{R}) \Rightarrow \lim_{\omega \to \infty} \widehat{f}(\omega) = 0$$

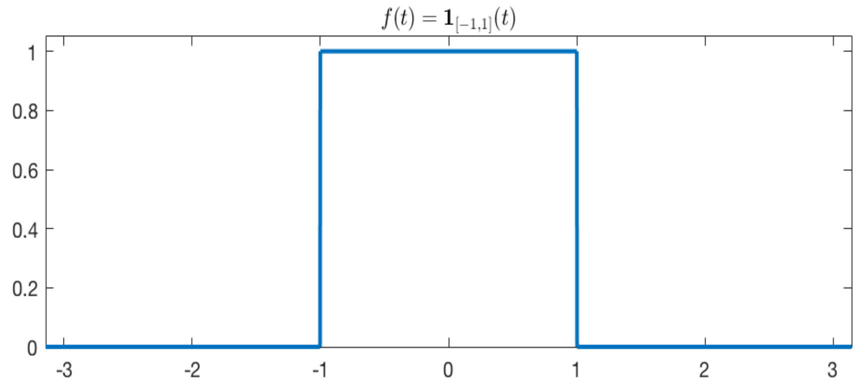


Regularity and Decay

- $\mathbf{C}^p(\mathbb{R})$ = Bounded continuous functions with p bounded continuous derivatives.

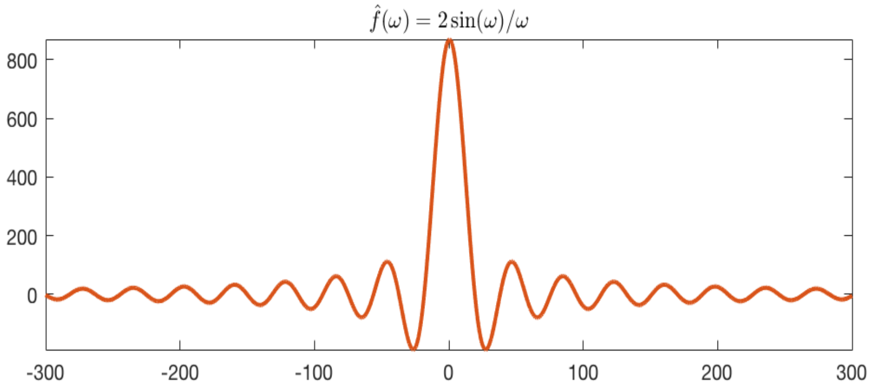
- Regularity of f implies decay of $\hat{f}$:

$$f^{(k)} \in \mathbf{L}^1(\mathbb{R}) \;\; \forall\, k \leq p \Rightarrow \hat{f}(\omega) = O(|\omega|^{-p})$$


$f(t) = \mathbf{1}_{[-1,1]}(t)$

- Faster decay implies more regularity:

$$|\hat{f}(\omega)| \leq \frac{K}{1 + |\omega|^{p+1+\epsilon}} \Rightarrow f \in \mathbf{C}^p(\mathbb{R})$$


$\hat{f}(\omega) = 2\sin(\omega)/\omega$

- The decay of $|\hat{f}(\omega)|$ depends on the *worst* singular behavior of f, e.g.,

$$f(t) = \mathbf{1}_{[-T,T]}(t) \Rightarrow \hat{f}(\omega) = \frac{2\sin(T\omega)}{\omega}$$

Time-Frequency Analysis

Time-Frequency Localization

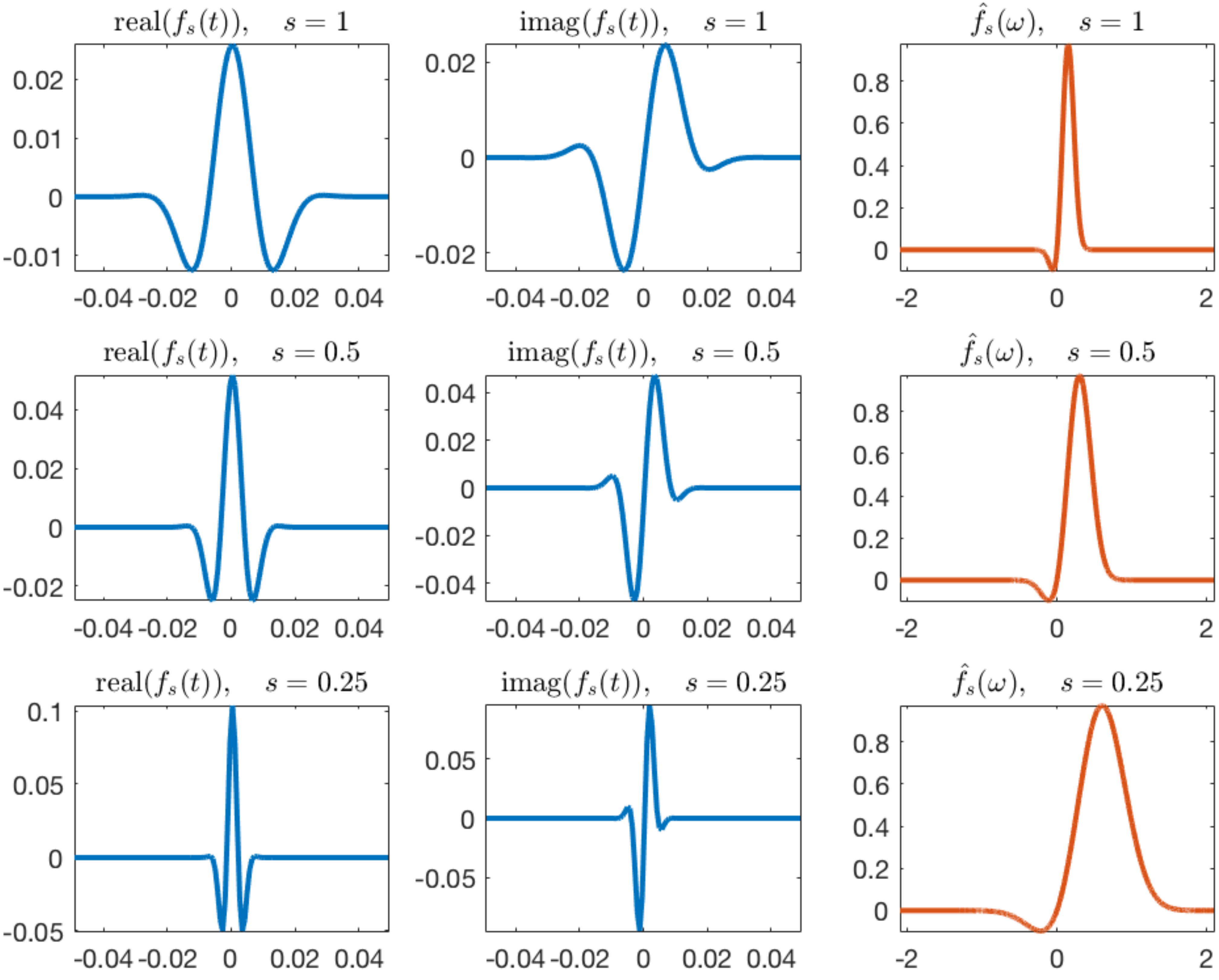
- Can we find a function f highly localized in time and with Fourier transform $\widehat{f}$ concentrated on a small frequency interval?
- $f(t) = \delta(t-u)$ is concentrated at $t = u$ in time, but $\widehat{f}(\omega) = e^{-iu\omega}$.
- Dilation of f:

$$f_s(t) = \frac{1}{\sqrt{s}} f(t/s) \quad (\|f_s\|_2 = \|f\|_2)$$

$$\widehat{f_s}(\omega) = \sqrt{s}\,\widehat{f}(s\omega)$$

So $s < 1$ reduces spread of f but increases spread of $\widehat{f}$.

Time-Frequency Localization



Heisenberg's Uncertainty Principle

- Interpret f with $\|f\|_2 = 1$ as a wavefunction describing the state of a 1D particle.
- $|f(t)|^2$: Probability density to find particle at position t.
- $\frac{1}{2\pi}|\hat{f}(\omega)|^2$: Probability density to find particle with momentum ω.
- Average location u and variance σ_t^2 are:

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

- Average momentum ξ and variance σ_ω^2 are:

$$\xi = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \omega|\hat{f}(\omega)|^2\,d\omega, \quad \sigma_\omega^2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty} (\omega-\xi)^2|\hat{f}(\omega)|^2\,d\omega$$

- For any $f \in \mathbf{L}^2(\mathbb{R})$: $\sigma_t\sigma_\omega \geq 1/2$

Time-Frequency Dictionary

- Dictionary $\mathcal{D} = \{\phi_\gamma\}_\gamma$ of time-frequency atoms, i.e., waveforms concentrated in time and frequency with $\|\phi_\gamma\|_2 = 1$.
- Corresponding time-frequency transform of $f \in \mathbf{L}^2(\mathbb{R})$ (2nd equality is Parseval):

$$\Phi f(\gamma) = \langle f, \phi_\gamma \rangle = \frac{1}{2\pi} \langle \hat{f}, \hat{\phi}_\gamma \rangle$$

- As in previous slide, let u_γ be center of $|\phi_\gamma(t)|^2$ with variance $\sigma_t^2(\gamma)$.
- Central frequency of $\hat{\phi}_\gamma$ is ξ_γ with variance $\sigma_\omega^2(\gamma)$.

Time-Frequency Dictionary

- Time-frequency resolution of ϕ_γ in the time-frequency plane (t, ω) is measured by a Heisenberg box centered at (u_γ, ξ_γ) having time width $\sigma_t(\gamma)$ and frequency width $\sigma_\omega(\gamma)$.
- Limited by the uncertainty principle:

$$\sigma_t(\gamma)\sigma_\omega(\gamma) \geq \frac{1}{2}$$

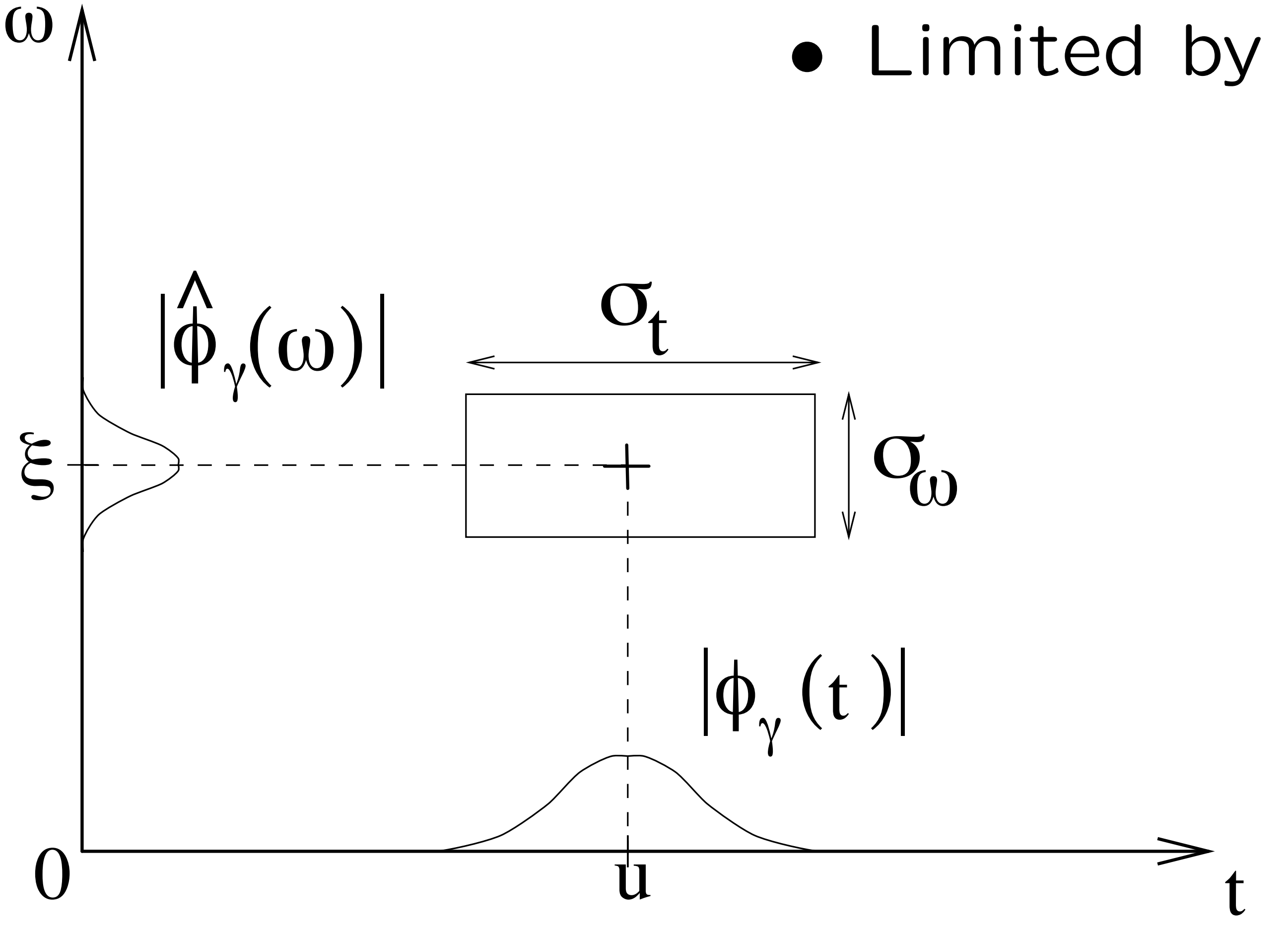


Fig. 1.3. A Wavelet Tour of Signal Processing, 3^{rd} ed. Heisenberg box representing an atom ϕ_γ.

Translation Invariant Dictionaries

- Observe:

$$\langle f_u, \phi_\gamma \rangle = \int_{-\infty}^{+\infty} f(t-u)\overline{\phi_\gamma(t)}\,dt = \int_{-\infty}^{\infty} f(t)\overline{\phi_\gamma(t+u)}\,dt$$

- Thus $\mathcal{D}$ is translation invariant invariant if it contains $\phi_\gamma(t+u)$.

- Write as:

$$\mathcal{D} = \{\phi_{u,\gamma}\}_{\gamma\in\Gamma, u\in\mathbb{R}}, \quad \phi_{u,\gamma}(t) = \phi_\gamma(t-u)$$

- Note:

$$\Phi f(u,\gamma) = \langle f, \phi_{u,\gamma} \rangle = f * \tilde{\phi}_\gamma(u), \quad \tilde{\phi}_\gamma(t) = \overline{\phi_\gamma(-t)}$$

Windowed Fourier Transform

Windowed Fourier Transform

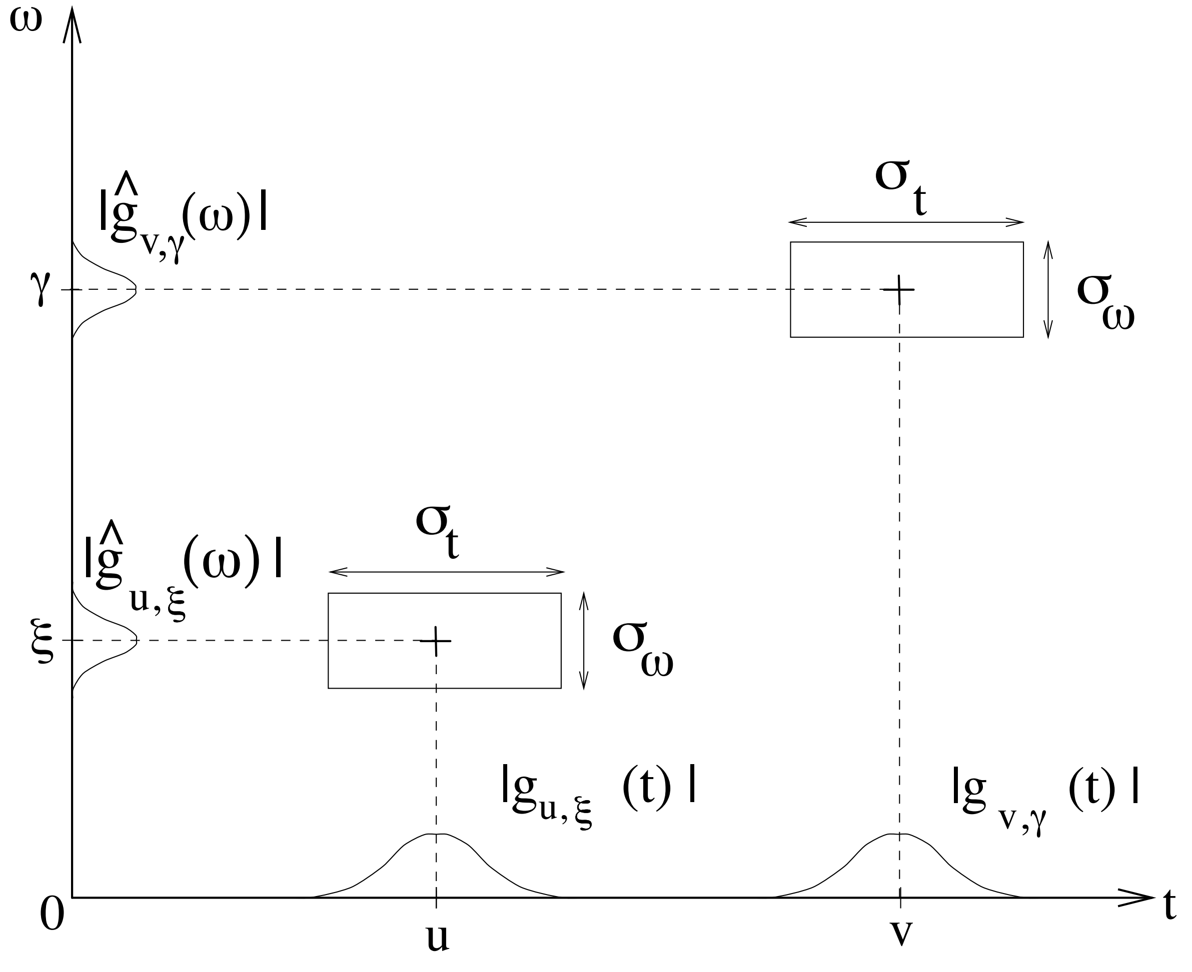
- Introduced by Gabor in 1946.
- Real and symmetric window $g(t) = g(-t)$, translated by u and modulated by a frequency ξ:

$$g_{u,\xi}(t) = e^{i\xi t} g(t-u)$$

- Translation invariant dictionary $\mathcal{D} = \{g_{u,\xi}\}_{u,\xi \in \mathbb{R}}$
- Windowed Fourier transform:

$$Sf(u,\xi) = \langle f, g_{u,\xi} \rangle = \int_{-\infty}^{+\infty} f(t) g(t-u) e^{-i\xi t} \, dt$$

Time-Frequency Spread of Windowed Fourier Transform



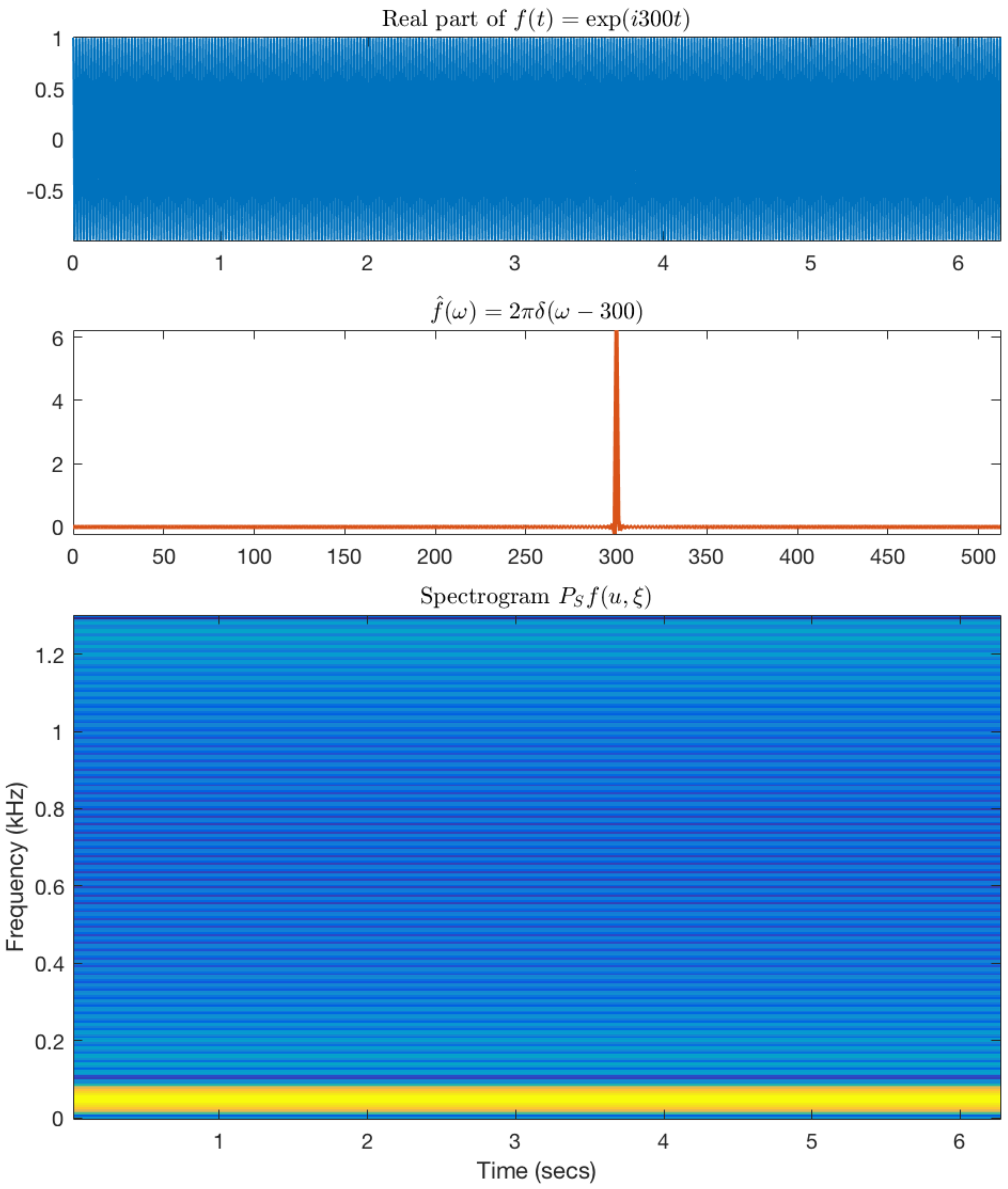
Wavelet Tour of Signal Processing, 3rd ed. Time-frequency boxes ("Heisenberg rectangles") representing the energy spread of two windowed Fourier atoms.

- $g_{u,\xi}$ is centered in time at u and in frequency at ξ.
- Time width σ_t and frequency width σ_ω of $g_{u,\xi}$ does not depend on (u,ξ).
- The spectrogram

$$P_S f(u,\xi) = |Sf(u,\xi)|^2$$

measures the energy of f in the Heisenberg box of $g_{u,\xi}$.

Sinusoid Example

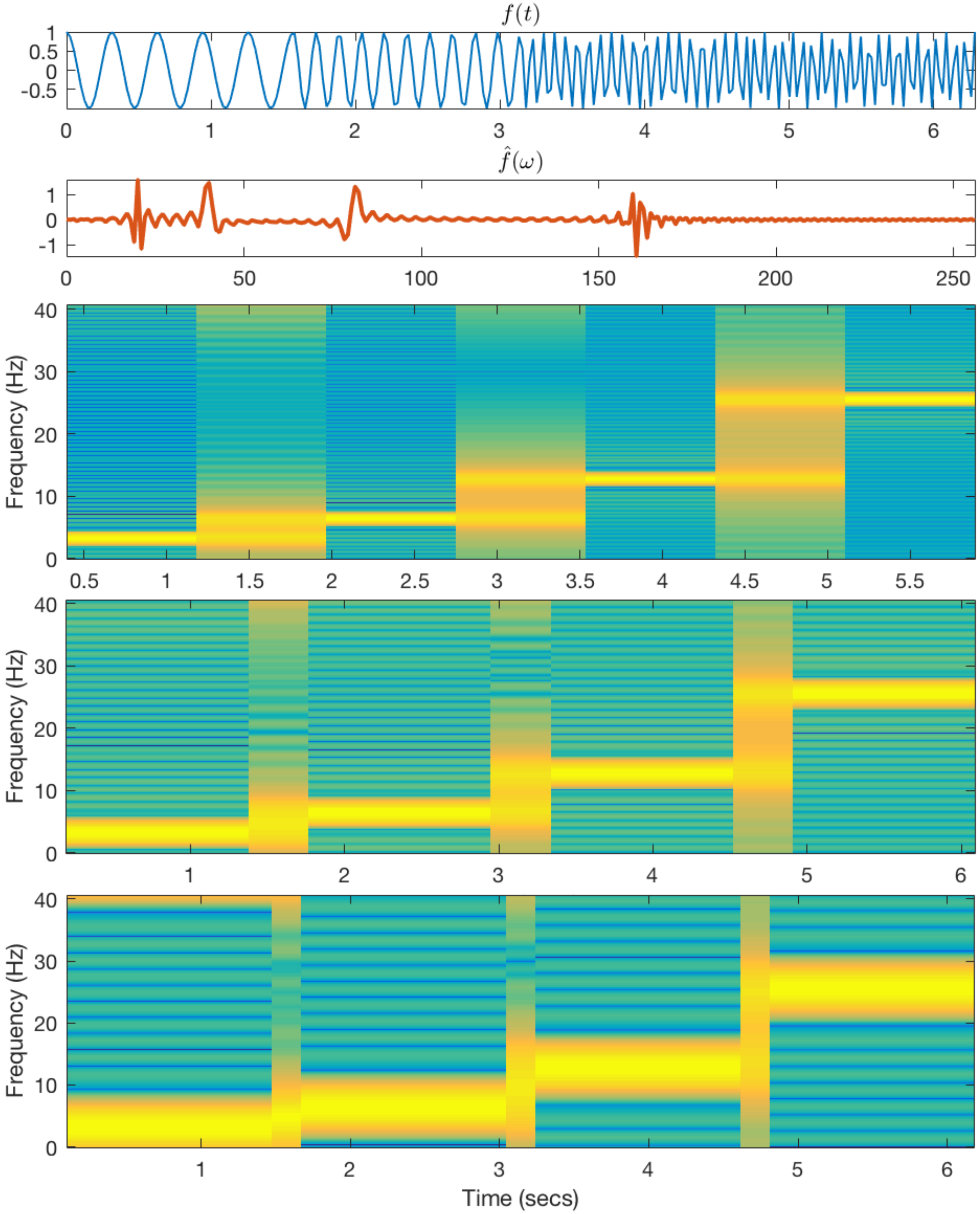


- $f(t) = e^{i\xi_0 t}$
- $\widehat{f}(\omega) = 2\pi\delta(\omega - \xi_0)$
- Windowed Fourier Transform

  $$Sf(u, \xi) = \widehat{g}(\xi - \xi_0)\,\mathrm{exp}(-iu(\xi - \xi_0))$$

  has energy spread over frequency interval $[\xi_0 - \sigma_\omega/2, \xi_0 + \sigma_\omega/2]$

Piecewise Sinusoid



- Function:

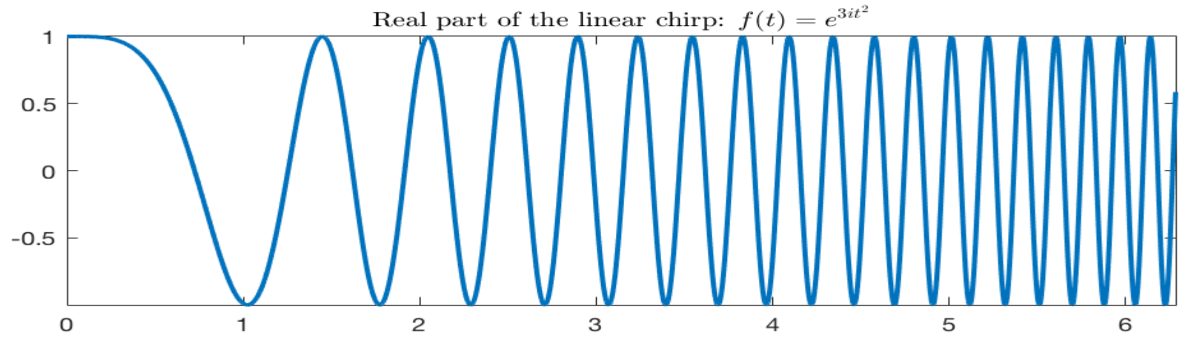
$$f(t) = \begin{cases} \cos(20t) & 0 \leq t \leq \frac{\pi}{2} \\ \cos(40t) & \frac{\pi}{2} < t \leq \pi \\ \cos(80t) & \pi < t \leq \frac{3\pi}{2} \\ \cos(160t) & \frac{3\pi}{2} < t \leq 2\pi \end{cases}$$

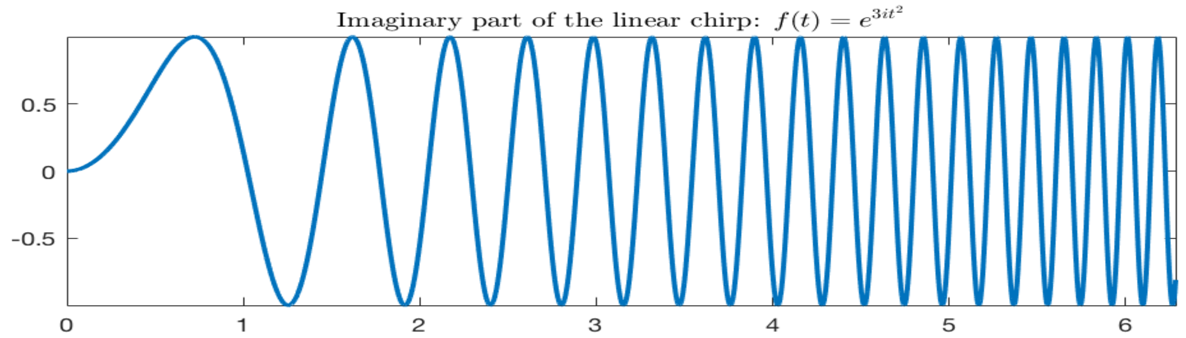
- $\hat{f}$ somewhat captures frequencies but gives no time information.

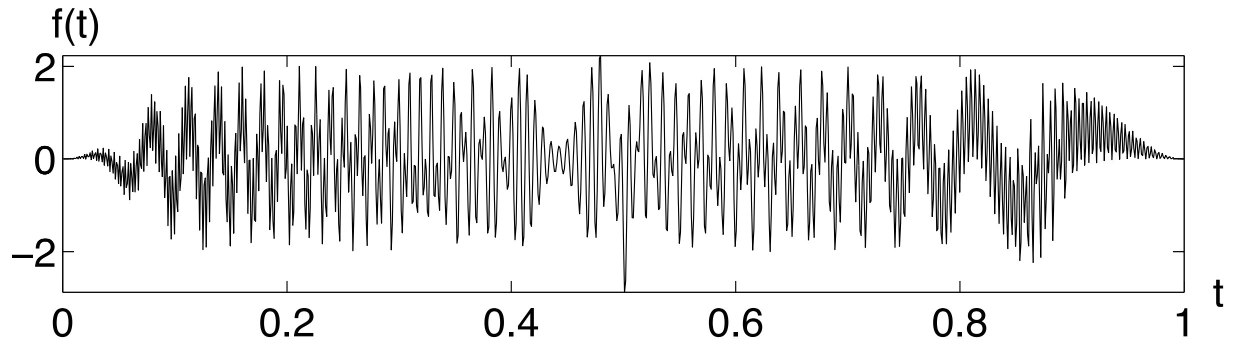
- Windowed Fourier transform $Sf(u, \xi)$ gives time and frequency information.

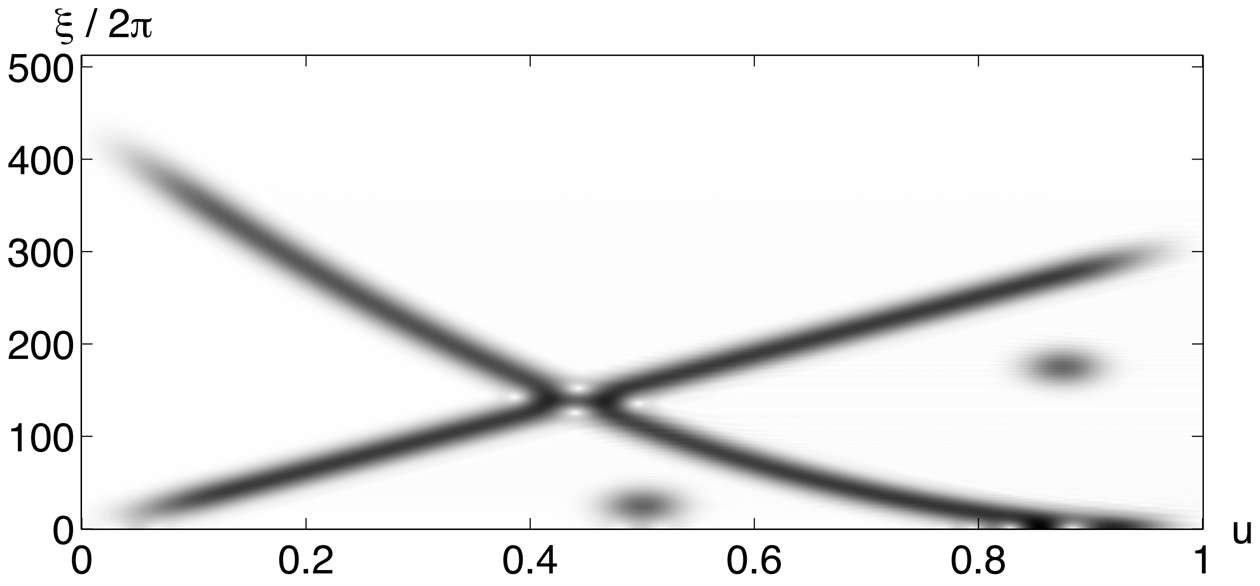
- The size of the window g balances time vs. frequency resolution.

Linear and Quadratic Chirps


Real part of the linear chirp: $f(t) = e^{3it^2}$


Imaginary part of the linear chirp: $f(t) = e^{3it^2}$





- Chirp: $c(t) = e^{i\phi(t)}$
- Linear if $\phi \in \mathcal{P}_2$
- Quadratic if $\phi \in \mathcal{P}_3$
- Instantaneous frequency is defined as $\phi'(t)$. Measures the frequency of $c(t)$ at time t.
- Function:

$$f(t) = e^{iat^2} + e^{ib(t-1)^3} + \theta(t-t_1)e^{i\xi_1 t} + \theta(t-t_2)e^{i\xi_2 t}$$

  where $\theta(t)$ is a Gaussian.
- Spectrogram $P_S f(u,\xi)$ yields large amplitude coefficients along the trajectories of the instantaneous frequencies of the linear and quadratic chirps, as well as low/high frequency blobs at the occurrence in time of the modulated Gaussians.

Completeness and Stability

- Intuitively the Heisenberg time-frequency boxes of the dictionary $\mathcal{D} = \{g_{u,\xi}\}_{u,\xi \in \mathbb{R}}$ cover the time-frequency plane (t, ω).
- Theorem: If $f \in \mathbf{L}^2(\mathbb{R})$, then:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} Sf(u,\xi) g(t-u) e^{i\xi t} \, d\xi \, du$$

$$\|f\|_2^2 = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |Sf(u,\xi)|^2 \, d\xi \, du$$

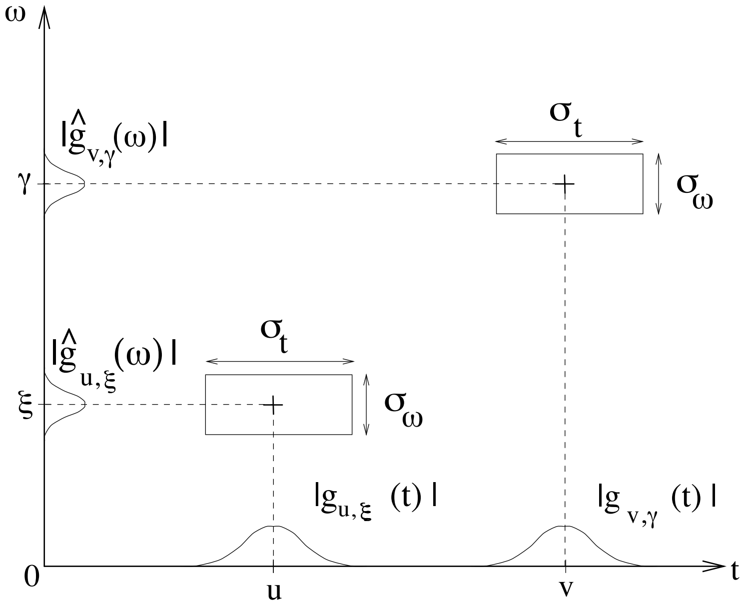
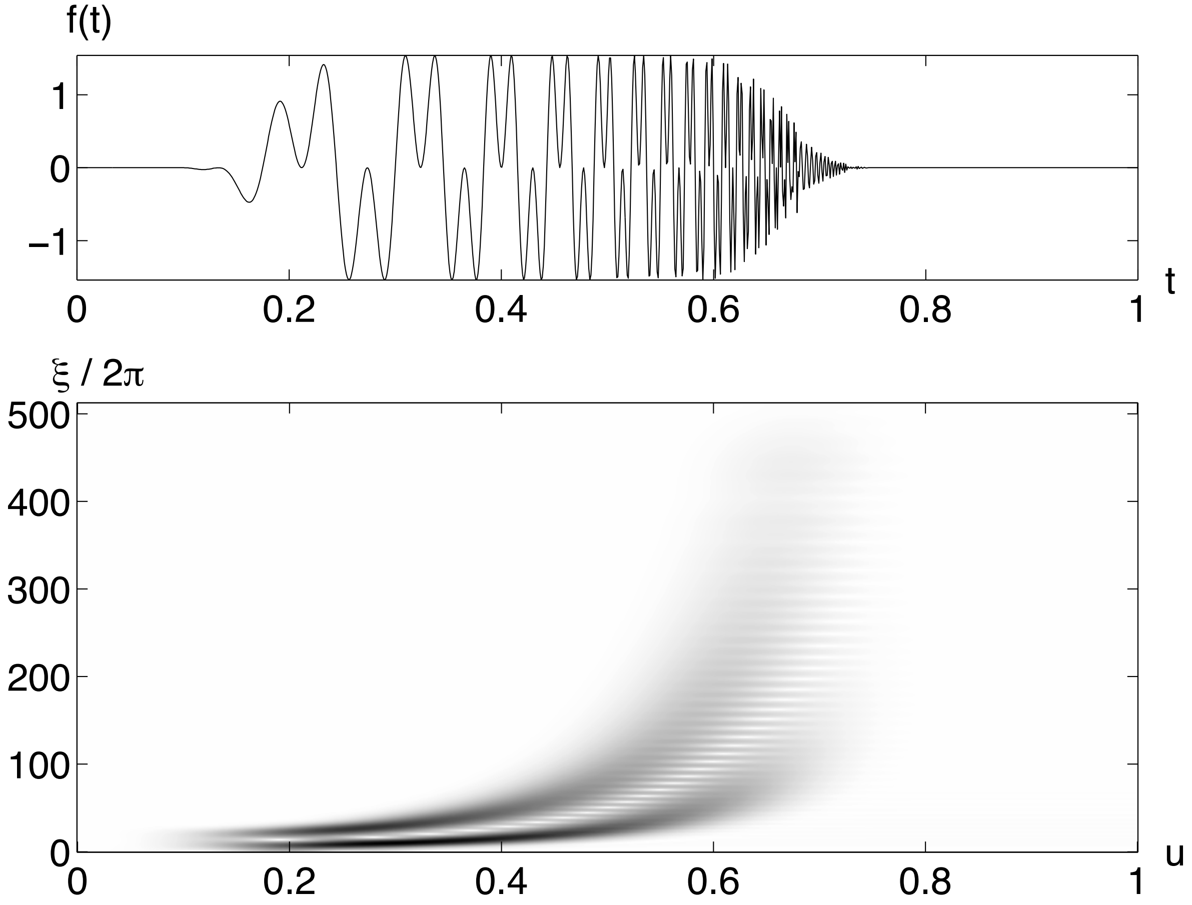


Fig. 1.4. A Wavelet Tour of Signal Processing, 3rd ed. Time-frequency boxes (“Heisenberg rectangles”) representing the energy spread of two windowed Fourier atoms.

Hyperbolic Chirps



- Hyperbolic chirp:
$$c(t) = \cos\left(\frac{\alpha}{\beta - t}\right)$$
- Instantaneous frequency is
$$\phi'(t) = \frac{\alpha}{(\beta - t)^2}$$
which varies quickly as $t \to \beta$.
- Function:
$$f(t) = a_1 \cos\left(\frac{\alpha_1}{\beta_1 - t}\right) + a_2 \cos\left(\frac{\alpha_2}{\beta_2 - t}\right)$$
- For $|\beta_1 - \beta_2|$ small, spectrogram $P_S f(u, \xi)$ with large time window can distinguish the similar instantaneous frequencies at small times.
- However, as t gets larger the instantaneous frequencies vary too quickly and the fixed frequency window size cannot keep up, leading to interference in the spectrogram.

Multiscale Signal Phenomena

- The windowed Fourier transform can analyze local time-frequency signal phenomena.
- However, the hyperbolic chirp example illustrates that the windowed Fourier transform fails when a signal contains structures of very different scales.
- For such signals we need time-frequency dictionaries with different time supports.