II

Fourier Kingdom

The story begins in 1807 when Fourier presents a memoir to the Institut de France, where he claims that any periodic function can be represented as a series of harmonically related sinusoids. This idea had a profound impact in mathematical analysis, physics and engineering, but it took one and a half centuries to understand the convergence of Fourier series and complete the theory of Fourier integrals.

Fourier was motivated by the study of heat diffusion, which is governed by a linear differential equation. However, the Fourier transform diagonalizes all linear time-invariant operators, which are the building blocks of signal processing. It is therefore not only the starting point of our exploration but the basis of all further developments.

2.1 Linear Time-Invariant Filtering

Classical signal processing operations such as signal transmission, stationary noise removal or predictive coding are implemented with linear time-invariant operators. The time invariance of an operator L means that if the input $f(t)$ is delayed by τ, $f_\tau(t) = f(t-\tau)$, then the output is also delayed by τ:

$$g(t) = Lf(t) \quad \Rightarrow \quad g(t-\tau) = Lf_\tau(t). \tag{2.1}$$

For numerical stability, the operator L must have a weak form of continuity, which means that Lf is modified by a small amount if f is slightly modified. This weak continuity is formalized by the theory of distributions [60, 63], which guarantees that we are on a safe ground without further worrying about it.

2.1.1 Impulse Response

Linear time-invariant systems are characterized by their response to a Dirac impulse, defined in Appendix A.7. If f is continuous, its value at t is obtained by an "integration" against a Dirac located at t. Let $\delta_u(t) = \delta(t-u)$:

$$f(t) = \int_{-\infty}^{+\infty} f(u)\, \delta_u(t)\, du.$$

The continuity and linearity of L imply that

$$Lf(t) = \int_{-\infty}^{+\infty} f(u)\, L\delta_u(t)\, du.$$

Let h be the impulse response of L:

$$h(t) = L\delta(t).$$

The time-invariance proves that $L\delta_u(t) = h(t-u)$ and hence

$$Lf(t) = \int_{-\infty}^{+\infty} f(u)\, h(t-u)\, du = \int_{-\infty}^{+\infty} h(u) f(t-u)\, du = h \star f\ (t). \tag{2.2}$$

A time-invariant linear filter is thus equivalent to a convolution with the impulse response h. The continuity of f is not necessary. This formula remains valid for any signal f for which the convolution integral converges.

Let us recall a few useful properties of convolution products:

- Commutativity
$$f \star h(t) = h \star f(t). \tag{2.3}$$
- Differentiation
$$\frac{d}{dt}(f \star h)(t) = \frac{df}{dt} \star h(t) = f \star \frac{dh}{dt}(t). \tag{2.4}$$
- Dirac convolution
$$f \star \delta_\tau(t) = f(t-\tau). \tag{2.5}$$

Stability and Causality A filter is said to be *causal* if $Lf(t)$ does not depend on the values $f(u)$ for $u > t$. Since

$$Lf(t) = \int_{-\infty}^{+\infty} h(u)\, f(t-u)\, du,$$

this means that $h(u) = 0$ for $u < 0$. Such impulse responses are said to be *causal*.

The *stability* property guarantees that $Lf(t)$ is bounded if $f(t)$ is bounded. Since

$$|Lf(t)| \leqslant \int_{-\infty}^{+\infty} |h(u)|\, |f(t-u)|\, du \leqslant \sup_{u\in\mathbb{R}} |f(u)| \int_{-\infty}^{+\infty} |h(u)|\, du,$$

it is sufficient that $\int_{-\infty}^{+\infty} |h(u)|\, du < +\infty$. One can verify that this condition is also necessary if h is a function. We thus say that h is *stable* if it is integrable.

Example 2.1. *An amplification and delay system is defined by*

$$Lf(t) = \lambda\, f(t-\tau).$$

The impulse response of this filter is $h(t) = \lambda\, \delta(t-\tau)$.

Example 2.2. *A uniform averaging of* f *over intervals of size* T *is calculated by*

$$Lf(t) = \frac{1}{T} \int_{t-T/2}^{t+T/2} f(u)\, du.$$

This integral can be rewritten as a convolution of f *with the impulse response* $h = 1/T\, \mathbf{1}_{[-T/2,T/2]}$.

2.1.2 Transfer Functions

Complex exponentials $\mathrm{e}^{i\omega t}$ are eigenvectors of convolution operators. Indeed

$$L\mathrm{e}^{i\omega t} = \int_{-\infty}^{+\infty} h(u)\, \mathrm{e}^{i\omega(t-u)}\, du,$$

which yields

$$L\mathrm{e}^{i\omega t} = \mathrm{e}^{it\omega} \int_{-\infty}^{+\infty} h(u)\, \mathrm{e}^{-i\omega u}\, du = \hat{h}(\omega)\, \mathrm{e}^{i\omega t}.$$

The eigenvalue

$$\hat{h}(\omega) = \int_{-\infty}^{+\infty} h(u)\, \mathrm{e}^{-i\omega u}\, du$$

is the Fourier transform of h at the frequency ω. Since complex sinusoidal waves $\mathrm{e}^{i\omega t}$ are the eigenvectors of time-invariant linear systems, it is tempting to try to decompose any function f as a sum of these eigenvectors. We are then able to express Lf directly from the eigenvalues $\hat{h}(\omega)$. The Fourier analysis proves that under weak conditions on f, it is indeed possible to write it as a Fourier integral.

2.2 Fourier Integrals

To avoid convergence issues, the Fourier integral is first defined over the space $\mathbf{L^1}(\mathbb{R})$ of integrable functions [53]. It is then extended to the space $\mathbf{L^2}(\mathbb{R})$ of finite energy functions [22].

2.2.1 Fourier Transform in $\mathbf{L^1}(\mathbb{R})$

The Fourier integral

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(t)\, \mathrm{e}^{-i\omega t}\, dt \tag{2.6}$$

measures "how much" oscillations at the frequency ω there is in f. If $f \in \mathbf{L^1}(\mathbb{R})$ this integral does converge and

$$|\hat{f}(\omega)| \leqslant \int_{-\infty}^{+\infty} |f(t)|\, dt < +\infty. \tag{2.7}$$

The Fourier transform is thus bounded, and one can verify that it is a continuous function of ω (Exercise 2.1). If $\hat{f}$ is also integrable, the following theorem gives the inverse Fourier transform.

Theorem 2.1 (Inverse Fourier Transform)**.** *If $f \in \mathbf{L^1}(\mathbb{R})$ and $\hat{f} \in \mathbf{L^1}(\mathbb{R})$ then*

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega)\, \mathrm{e}^{i\omega t}\, d\omega. \tag{2.8}$$

Proof. Replacing $\hat{f}(\omega)$ by its integral expression yields

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega)\, \exp(i\omega t)\, d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(u)\, \exp[i\omega(t-u)]\, du \right) d\omega.$$

We cannot apply the Fubini Theorem A.2 directly because $f(u)\exp[i\omega(t-u)]$ is not integrable in $\mathbb{R}^2$. To avoid this technical problem, we multiply by $\exp(-\varepsilon^2\omega^2/4)$ which converges to 1 when ε goes to 0. Let us define

$$I_\varepsilon(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(u)\, \exp\left(\frac{-\varepsilon^2\omega^2}{4}\right) \exp[i\omega(t-u)]\, du \right) d\omega. \tag{2.9}$$

We compute I_ε in two different ways using the Fubini theorem. The integration with respect to u gives

$$I_\varepsilon(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega)\, \exp\left(\frac{-\varepsilon^2\omega^2}{4}\right) \exp(i\omega t)\, d\omega.$$

Since

$$\left| \hat{f}(\omega)\, \exp\left(\frac{-\varepsilon^2\omega^2}{4}\right) \exp[i\omega(t-u)] \right| \leqslant |\hat{f}(\omega)|$$

and since $\hat{f}$ is integrable, we can apply the dominated convergence Theorem A.1, which proves that

$$\lim_{\varepsilon \to 0} I_\varepsilon(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega)\, \exp(i\omega t)\, d\omega. \tag{2.10}$$

Let us now compute the integral (2.9) differently by applying the Fubini theorem and integrating with respect to ω:

$$I_\varepsilon(t) = \int_{-\infty}^{+\infty} g_\varepsilon(t-u)\, f(u)\, du, \tag{2.11}$$

with

$$g_\varepsilon(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(ix\omega)\, \exp\left(\frac{-\varepsilon^2\omega^2}{4}\right) d\omega.$$

A change of variable $\omega' = \varepsilon\omega$ shows that $g_\varepsilon(x) = \varepsilon^{-1} g_1(\varepsilon^{-1}x)$, and it is proved in (2.32) that $g_1(x) = \pi^{-1/2}\, \mathrm{e}^{-x^2}$. The Gaussian g_1 has an integral equal to 1 and a fast decay. The squeezed Gaussians g_ε have an integral that remains equal to 1, and thus they converge to a Dirac δ when ε goes to 0. By inserting (2.11) one can thus verify that

$$\lim_{\varepsilon \to 0} \int_{-\infty}^{+\infty} |I_\varepsilon(t) - f(t)|\, dt = \lim_{\varepsilon \to 0} \iint g_\varepsilon(t-u)\, |f(u) - f(t)|\, du\, dt = 0.$$

Inserting (2.10) proves (2.8). ■ ■

The inversion formula (2.8) decomposes f as a sum of sinusoidal waves $\mathrm{e}^{i\omega t}$ of amplitude $\hat{f}(\omega)$. By using this formula, we can show (Exercise 2.1) that the hypothesis $\hat{f} \in \mathbf{L^1}(\mathbb{R})$ implies that f must be continuous. The reconstruction (2.8) is therefore not proved for discontinuous functions. The extension of the Fourier transform to the space $\mathbf{L^2}(\mathbb{R})$ will address this issue.

The most important property of the Fourier transform for signal processing applications is the convolution theorem. It is another way to express the fact that sinusoidal waves $\mathrm{e}^{it\omega}$ are eigenvalues of convolution operators.

Theorem 2.2 (Convolution)**.** *Let $f \in \mathbf{L^1}(\mathbb{R})$ and $h \in \mathbf{L^1}(\mathbb{R})$. The function $g = h \star f$ is in $\mathbf{L^1}(\mathbb{R})$ and*

$$\hat{g}(\omega) = \hat{h}(\omega)\, \hat{f}(\omega). \tag{2.12}$$

Proof.

$$\hat{g}(\omega) = \int_{-\infty}^{+\infty} \exp(-it\omega) \left(\int_{-\infty}^{+\infty} f(t-u)\, h(u)\, du \right) dt.$$

Since $|f(t-u)||h(u)|$ is integrable in $\mathbb{R}^2$, we can apply the Fubini Theorem A.2, and the change of variable $(t,u) \to (v = t-u, u)$ yields

$$\begin{aligned} \hat{g}(\omega) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp[-i(u+v)\omega]\, f(v)\, h(u)\, du\, dv \\ &= \left(\int_{-\infty}^{+\infty} \exp(-iv\omega)\, f(v)\, dv \right) \left(\int_{-\infty}^{+\infty} \exp(-iu\omega)\, h(u)\, du \right), \end{aligned}$$

which verifies (2.12). ■ ■

The response $Lf = g = f \star h$ of a linear time-invariant system can be calculated from its Fourier transform $\hat{g}(\omega) = \hat{f}(\omega)\, \hat{h}(\omega)$ with the inverse Fourier formula

$$g(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{g}(\omega)\, \mathrm{e}^{i\omega t}\, d\omega, \tag{2.13}$$

which yields

$$Lf(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{h}(\omega)\, \hat{f}(\omega)\, \mathrm{e}^{i\omega t}\, d\omega. \tag{2.14}$$

Each frequency component $\mathrm{e}^{it\omega}$ of amplitude $\hat{f}(\omega)$ is amplified or attenuated by $\hat{h}(\omega)$. Such a convolution is thus called a *frequency filtering*, and $\hat{h}$ is the *transfer function* of the filter.

The following table summarizes important properties of the Fourier transform, often used in calculations. Most of these formulas are proved with a change of variable in the Fourier integral.

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	$f(t-u)$	$\mathrm{e}^{-iu\omega}\, \hat{f}(\omega)$	(2.18)
Modulation	$\mathrm{e}^{i\xi t}\, f(t)$	$\hat{f}(\omega - \xi)$	(2.19)
Scaling	$f(t/s)$	$\lvert s \rvert\, \hat{f}(s\,\omega)$	(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)

2.2.2 Fourier Transform in $\mathbf{L^2}(\mathbb{R})$

The Fourier transform of the indicator function $f = \mathbf{1}_{[-1,1]}$ is

$$\hat{f}(\omega) = \int_{-1}^{1} \mathrm{e}^{-i\omega t}\, dt = \frac{2\sin\omega}{\omega}.$$

This function is not integrable because f is not continuous, but its square is integrable. The inverse Fourier transform Theorem 2.1 thus does not apply. This motivates the extension of the Fourier transform to the space $\mathbf{L^2}(\mathbb{R})$ of functions f with a finite energy $\int_{-\infty}^{+\infty} |f(t)|^2\, dt < +\infty$. By working in the Hilbert space $\mathbf{L^2}(\mathbb{R})$, we also have access to all the facilities provided by the existence of an inner product. The inner product of $f \in \mathbf{L^2}(\mathbb{R})$ and $g \in \mathbf{L^2}(\mathbb{R})$ is

$$\langle f, g \rangle = \int_{-\infty}^{+\infty} f(t)\, g^*(t)\, dt,$$

and the resulting norm in $\mathbf{L^2}(\mathbb{R})$ is

$$\|f\|^2 = \langle f, f \rangle = \int_{-\infty}^{+\infty} |f(t)|^2\, dt.$$

The following theorem proves that inner products and norms in $\mathbf{L^2}(\mathbb{R})$ are conserved by the Fourier transform up to a factor of 2π. Equations (2.25) and (2.26) are called respectively the *Parseval* and *Plancherel* formulas.

Theorem 2.3. *If f and h are in $\mathbf{L^1}(\mathbb{R}) \cap \mathbf{L^2}(\mathbb{R})$ then*

$$\int_{-\infty}^{+\infty} f(t)\, h^*(t)\, dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega)\, \hat{h}^*(\omega)\, d\omega. \tag{2.25}$$

For $h = f$ it follows that

$$\int_{-\infty}^{+\infty} |f(t)|^2\, dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |\hat{f}(\omega)|^2\, d\omega. \tag{2.26}$$

Proof. Let $g = f \star \bar{h}$ with $\bar{h}(t) = h^*(-t)$. The convolution Theorem 2.2 and property (2.23) show that $\hat{g}(\omega) = \hat{f}(\omega)\, \hat{h}^*(\omega)$. The reconstruction formula (2.8) applied to $g(0)$ yields

$$\int_{-\infty}^{+\infty} f(t)\, h^*(t)\, dt = g(0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{g}(\omega)\, d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega)\, \hat{h}^*(\omega)\, d\omega.$$

■ ■

Density Extension in $\mathbf{L^2}(\mathbb{R})$ If $f \in \mathbf{L^2}(\mathbb{R})$ but $f \notin \mathbf{L^1}(\mathbb{R})$, its Fourier transform cannot be calculated with the Fourier integral (2.6) because $f(t)\, \mathrm{e}^{i\omega t}$ is not integrable. It is defined as a limit using the Fourier transforms of functions in $\mathbf{L^1}(\mathbb{R}) \cap \mathbf{L^2}(\mathbb{R})$.

Since $\mathbf{L^1}(\mathbb{R}) \cap \mathbf{L^2}(\mathbb{R})$ is dense in $\mathbf{L^2}(\mathbb{R})$, one can find a family $\{f_n\}_{n\in\mathbb{Z}}$ of functions in $\mathbf{L^1}(\mathbb{R}) \cap \mathbf{L^2}(\mathbb{R})$ that converges to f:

$$\lim_{n\to+\infty} \|f - f_n\| = 0.$$

Since $\{f_n\}_{n\in\mathbb{Z}}$ converges, it is a Cauchy sequence, which means that $\|f_n - f_p\|$ is arbitrarily small if n and p are large enough. Moreover, $f_n \in \mathbf{L^1}(\mathbb{R})$, so its Fourier transform $\hat{f}_n$ is well defined. The Plancherel formula (2.26) proves that $\{\hat{f}_n\}_{n\in\mathbb{Z}}$ is also a Cauchy sequence because

$$\|\hat{f}_n - \hat{f}_p\| = \sqrt{2\pi}\, \|f_n - f_p\|$$

is arbitrarily small for n and p large enough. A Hilbert space (Appendix A.2) is complete, which means that all Cauchy sequences converge to an element of the space. Hence, there exists $\hat{f} \in \mathbf{L^2}(\mathbb{R})$ such that

$$\lim_{n\to+\infty} \|\hat{f} - \hat{f}_n\| = 0.$$

By definition, $\hat{f}$ is the Fourier transform of f. This extension of the Fourier transform to $\mathbf{L^2}(\mathbb{R})$ satisfies the convolution theorem, the Parseval and Plancherel formulas, as well as all properties (2.15-2.24).

Diracs Diracs are often used in calculations; their properties are summarized in Appendix A.7. A Dirac δ associates to a function its value at $t = 0$. Since $e^{i\omega t} = 1$ at $t = 0$ it seems reasonable to define its Fourier transform by

$$\hat{\delta}(\omega) = \int_{-\infty}^{+\infty} \delta(t)\, e^{-i\omega t}\, dt = 1. \tag{2.27}$$

This formula is justified mathematically by the extension of the Fourier transform to tempered distributions [60, 63].

2.2.3 Examples

The following examples often appear in Fourier calculations. They also illustrate important Fourier transform properties.

- The *indicator function* $f = \mathbf{1}_{[-T,T]}$ is discontinuous at $t = \pm T$. Its Fourier transform is therefore not integrable:
$$\hat{f}(\omega) = \int_{-T}^{T} e^{-i\omega t}\, dt = \frac{2\sin(T\omega)}{\omega}. \tag{2.28}$$

- An *ideal low-pass filter* has a transfer function $\hat{\phi} = \mathbf{1}_{[-\xi,\xi]}$ that selects low frequencies over $[-\xi, \xi]$. The impulse response is calculated with the inverse Fourier integral (2.8):
$$\phi(t) = \frac{1}{2\pi} \int_{-\xi}^{\xi} e^{i\omega t}\, d\omega = \frac{\sin(\xi t)}{\pi t}. \tag{2.29}$$

- A *passive electronic circuit* implements analog filters with resistances, capacities and inductors. The input voltage $f(t)$ is related to the output voltage $g(t)$ by a differential equation with constant coefficients:
$$\sum_{k=0}^{K} a_k\, f^{(k)}(t) = \sum_{k=0}^{M} b_k\, g^{(k)}(t). \tag{2.30}$$
Suppose that the circuit is not charged for $t < 0$, which means that $f(t) = g(t) = 0$. The output g is a linear time-invariant function of f and can thus be written $g = f \star \phi$. Computing the Fourier transform of (2.30) and applying (2.22) proves that
$$\hat{\phi}(\omega) = \frac{\hat{g}(\omega)}{\hat{f}(\omega)} = \frac{\sum_{k=0}^{K} a_k\, (i\omega)^k}{\sum_{k=0}^{M} b_k\, (i\omega)^k}. \tag{2.31}$$
It is therefore a rational function of $i\omega$. An ideal low-pass transfer function $\mathbf{1}_{[-\xi,\xi]}$ thus cannot be implemented by an analog circuit. It must be approximated by a rational function. Chebyshev or Butterworth filters are often used for this purpose [13].

- A *Gaussian* $f(t) = \exp(-t^2)$ is a $\mathbf{C}^\infty$ function with a fast asymptotic decay. Its Fourier transform is also a Gaussian:
$$\hat{f}(\omega) = \sqrt{\pi}\, \exp(-\omega^2/4). \tag{2.32}$$
This Fourier transform is computed by showing with an integration by parts that $\hat{f}(\omega) = \int_{-\infty}^{+\infty} \exp(-t^2)\, e^{-i\omega t} dt$ is differentiable and satisfies the differential equation
$$2\, \hat{f}'(\omega) + \omega\, \hat{f}(\omega) = 0. \tag{2.33}$$
The solution of this equation is a Gaussian $\hat{f}(\omega) = K \exp(-\omega^2/4)$, and since $\hat{f}(0) = \int_{-\infty}^{+\infty} \exp(-t^2)\, dt = \sqrt{\pi}$, we obtain (2.32).

- A Gaussian *chirp* $f(t) = \exp[-(a - ib)t^2]$ has a Fourier transform calculated with a similar differential equation:
$$\hat{f}(\omega) = \sqrt{\frac{\pi}{a - ib}}\, \exp\left(\frac{-(a + ib)\omega^2}{4(a^2 + b^2)}\right). \tag{2.34}$$

- A translated *Dirac* $\delta_\tau(t) = \delta(t-\tau)$ has a Fourier transform calculated by evaluating $\mathrm{e}^{-i\omega t}$ at $t = \tau$:
$$\hat{\delta}_\tau(\omega) = \int_{-\infty}^{+\infty} \delta(t-\tau)\,\mathrm{e}^{-i\omega t}\,dt = \mathrm{e}^{-i\omega\tau}. \tag{2.35}$$

- The *Dirac comb* is a sum of translated Diracs
$$c(t) = \sum_{n=-\infty}^{+\infty} \delta(t-nT)$$
that is used to uniformly sample analog signals. Its Fourier transform is derived from (2.35):
$$\hat{c}(\omega) = \sum_{n=-\infty}^{+\infty} \mathrm{e}^{-inT\omega}. \tag{2.36}$$
The Poisson formula proves that it is also equal to a Dirac comb with a spacing equal to $2\pi/T$.

Theorem 2.4 (Poisson Formula)**.** *In the sense of distribution equalities (A.29),*
$$\sum_{n=-\infty}^{+\infty} \mathrm{e}^{-inT\omega} = \frac{2\pi}{T} \sum_{k=-\infty}^{+\infty} \delta\left(\omega - \frac{2\pi k}{T}\right). \tag{2.37}$$

Proof. The Fourier transform $\hat{c}$ in (2.36) is periodic with period $2\pi/T$. To verify the Poisson formula, it is therefore sufficient to prove that the restriction of $\hat{c}$ to $[-\pi/T, \pi/T]$ is equal to $2\pi/T\,\delta$. The formula (2.37) is proved in the sense of a distribution equality (A.29) by showing that for any test function $\hat{\theta}(\omega)$ with a support included in $[-\pi/T, \pi/T]$,
$$\langle \hat{c}, \hat{\theta} \rangle = \lim_{N\to+\infty} \int_{-\infty}^{+\infty} \sum_{n=-N}^{N} \exp(-inT\omega)\,\hat{\theta}(\omega)\,d\omega = \frac{2\pi}{T}\,\hat{\theta}(0).$$

The sum of the geometric series is
$$\sum_{n=-N}^{N} \exp(-inT\omega) = \frac{\sin[(N+1/2)T\omega]}{\sin[T\omega/2]}. \tag{2.38}$$

Hence
$$\langle \hat{c}, \hat{\theta} \rangle = \lim_{N\to+\infty} \frac{2\pi}{T} \int_{-\pi/T}^{\pi/T} \frac{\sin[(N+1/2)T\omega]}{\pi\omega}\,\frac{T\omega/2}{\sin[T\omega/2]}\,\hat{\theta}(\omega)\,d\omega. \tag{2.39}$$

Let
$$\hat{\psi}(\omega) = \begin{cases} \hat{\theta}(\omega)\,\frac{T\omega/2}{\sin[T\omega/2]} & \text{if } |\omega| \leqslant \pi/T \\ 0 & \text{if } |\omega| > \pi/T \end{cases}$$
and $\psi(t)$ be the inverse Fourier transform of $\hat{\psi}(\omega)$. Since $2\omega^{-1}\sin(a\omega)$ is the Fourier transform of $\mathbf{1}_{[-a,a]}(t)$, the Parseval formula (2.25) implies
$$\begin{aligned} \langle \hat{c}, \hat{\theta} \rangle &= \lim_{N\to+\infty} \frac{2\pi}{T} \int_{-\infty}^{+\infty} \frac{\sin[(N+1/2)T\omega]}{\pi\omega}\,\hat{\psi}(\omega)\,d\omega \\ &= \lim_{N\to+\infty} \frac{2\pi}{T} \int_{-(N+1/2)T}^{(N+1/2)T} \psi(t)\,dt. \end{aligned} \tag{2.40}$$

When N goes to $+\infty$ the integral converges to $\hat{\psi}(0) = \hat{\theta}(0)$. ■ ■

2.3 Properties

2.3.1 Regularity and Decay

The global regularity of a signal f depends on the decay of $|\hat{f}(\omega)|$ when the frequency ω increases. The differentiability of f is studied. If $\hat{f} \in \mathbf{L^1}(\mathbb{R})$, then the Fourier inversion formula (2.8) implies

that f is continuous and bounded:

$$|f(t)| \leqslant \frac{1}{2\pi} \int_{-\infty}^{+\infty} |e^{i\omega t} \hat{f}(\omega)| \, d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |\hat{f}(\omega)| \, d\omega < +\infty \,. \tag{2.41}$$

The next theorem applies this property to obtain a sufficient condition that guarantees the differentiability of f at any order p.

Theorem 2.5. *A function f is bounded and p times continuously differentiable with bounded derivatives if*

$$\int_{-\infty}^{+\infty} |\hat{f}(\omega)| \, (1 + |\omega|^p) \, d\omega < +\infty \,. \tag{2.42}$$

Proof. The Fourier transform of the k^{th} order derivative $f^{(k)}(t)$ is $(i\omega)^k \hat{f}(\omega)$. Applying (2.41) to this derivative proves that

$$|f^{(k)}(t)| \leqslant \int_{-\infty}^{+\infty} |\hat{f}(\omega)| \, |\omega|^k \, d\omega.$$

Condition (2.42) implies that $\int_{-\infty}^{+\infty} |\hat{f}(\omega)||\omega|^k \, d\omega < +\infty$ for any $k \leqslant p$, so $f^{(k)}(t)$ is continuous and bounded. ■ ■

This result proves that if there exist a constant K and $\varepsilon > 0$ such that

$$|\hat{f}(\omega)| \leqslant \frac{K}{1 + |\omega|^{p+1+\varepsilon}}, \quad \text{then} \quad f \in \mathbf{C}^p.$$

If $\hat{f}$ has a compact support then (2.42) implies that $f \in \mathbf{C}^\infty$.

The decay of $|\hat{f}(\omega)|$ depends on the worst singular behavior of f. For example, $f = \mathbf{1}_{[-T,T]}$ is discontinuous at $t = \pm T$, so $|\hat{f}(\omega)|$ decays like $|\omega|^{-1}$. In this case, it could also be important to know that $f(t)$ is regular for $t \neq \pm T$. This information cannot be derived from the decay of $|\hat{f}(\omega)|$. To characterize local regularity of a signal f it is necessary to decompose it over waveforms that are well localized in time, as opposed to sinusoidal waves $e^{i\omega t}$. Section 6.1.3 explains that wavelets are particularly well adapted to this purpose.

2.3.2 Uncertainty Principle

Can we construct a function f whose energy is well localized in time and whose Fourier transform $\hat{f}$ has an energy concentrated in a small frequency neighborhood? The Dirac $\delta(t-u)$ has a support restricted to $t = u$ but its Fourier transform $e^{-iu\omega}$ has an energy uniformly spread over all frequencies. We know that $|\hat{f}(\omega)|$ decays quickly at high frequencies only if f has regular variations in time. The energy of f must therefore be spread over a relatively large domain.

To reduce the time spread of f, we can scale it by $s < 1$ while maintaining constant its total energy. If

$$f_s(t) = \frac{1}{\sqrt{s}} \, f\left(\frac{t}{s}\right) \quad \text{then} \quad \|f_s\|^2 = \|f\|^2.$$

The Fourier transform $\hat{f}_s(\omega) = \sqrt{s} \, \hat{f}(s\omega)$ is dilated by $1/s$ so we lose in frequency localization what we gained in time. Underlying is a trade-off between time and frequency localization.

Time and frequency energy concentrations are restricted by the Heisenberg uncertainty principle. This principle has a particularly important interpretation in quantum mechanics as an uncertainty as to the position and momentum of a free particle. The state of a one-dimensional particle is described by a wave function $f \in \mathbf{L^2}(\mathbb{R})$. The probability density that this particle is located at t is $\frac{1}{\|f\|^2}|f(t)|^2$. The probability density that its momentum is equal to ω is $\frac{1}{2\pi\|f\|^2}|\hat{f}(\omega)|^2$. The average location of this particle is

$$u = \frac{1}{\|f\|^2} \int_{-\infty}^{+\infty} t \, |f(t)|^2 \, dt, \tag{2.43}$$

and the average momentum is

$$\xi = \frac{1}{2\pi\|f\|^2}\int_{-\infty}^{+\infty} \omega\,|\hat{f}(\omega)|^2\,d\omega. \tag{2.44}$$

The variances around these average values are respectively

$$\sigma_t^2 = \frac{1}{\|f\|^2}\int_{-\infty}^{+\infty} (t-u)^2\,|f(t)|^2\,dt \tag{2.45}$$

and

$$\sigma_\omega^2 = \frac{1}{2\pi\|f\|^2}\int_{-\infty}^{+\infty} (\omega-\xi)^2\,|\hat{f}(\omega)|^2\,d\omega. \tag{2.46}$$

The larger σ_t, the more uncertainty there is concerning the position of the free particle; the larger σ_ω, the more uncertainty there is concerning its momentum.

Theorem 2.6 (Heisenberg Uncertainty)**.** *The temporal variance and the frequency variance of $f \in \mathbf{L^2}(\mathbb{R})$ satisfy*

$$\sigma_t^2\,\sigma_\omega^2 \geqslant \frac{1}{4}\,. \tag{2.47}$$

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

$$f(t) = a\,\exp[i\xi t - b(t-u)^2]. \tag{2.48}$$

Proof. The following proof due to Weyl [68] supposes that $\lim_{|t|\to+\infty}\sqrt{t}f(t) = 0$, but the theorem is valid for any $f \in \mathbf{L^2}(\mathbb{R})$. If the average time and frequency localization of f is u and ξ, then the average time and frequency location of $\exp(-i\xi t)\,f(t+u)$ is zero. It is thus sufficient to prove the theorem for $u = \xi = 0$. Observe that

$$\sigma_t^2\,\sigma_\omega^2 = \frac{1}{2\pi\|f\|^4}\int_{-\infty}^{+\infty} |t\,f(t)|^2\,dt\int_{-\infty}^{+\infty} |\omega\,\hat{f}(\omega)|^2\,d\omega. \tag{2.49}$$

Since $i\omega\hat{f}(\omega)$ is the Fourier transform of $f'(t)$, the Plancherel identity (2.26) applied to $i\omega\hat{f}(\omega)$ yields

$$\sigma_t^2\sigma_\omega^2 = \frac{1}{\|f\|^4}\int_{-\infty}^{+\infty} |t\,f(t)|^2\,dt\int_{-\infty}^{+\infty} |f'(t)|^2\,dt. \tag{2.50}$$

Schwarz's inequality implies

$$\begin{aligned}
\sigma_t^2\sigma_\omega^2 &\geqslant \frac{1}{\|f\|^4}\left[\int_{-\infty}^{+\infty} |t\,f'(t)\,f^*(t)|\,dt\right]^2\\
&\geqslant \frac{1}{\|f\|^4}\left[\int_{-\infty}^{+\infty} \frac{t}{2}\,[f'(t)\,f^*(t) + f'^*(t)\,f(t)]\,dt\right]^2\\
&\geqslant \frac{1}{4\|f\|^4}\left[\int_{-\infty}^{+\infty} t\,(|f(t)|^2)'\,dt\right]^2.
\end{aligned}$$

Since $\lim_{|t|\to+\infty}\sqrt{t}\,f(t) = 0$, an integration by parts gives

$$\sigma_t^2\sigma_\omega^2 \geqslant \frac{1}{4\|f\|^4}\left[\int_{-\infty}^{+\infty} |f(t)|^2\,dt\right]^2 = \frac{1}{4}. \tag{2.51}$$

To obtain an equality, Schwarz's inequality applied to (2.50) must be an equality. This implies that there exists $b \in \mathbb{C}$ such that

$$f'(t) = -2\,b\,t\,f(t). \tag{2.52}$$

Hence, there exists $a \in \mathbb{C}$ such that $f(t) = a\,\exp(-bt^2)$. The other steps of the proof are then equalities so that the lower bound is indeed reached. When $u \neq 0$ and $\xi \neq 0$ the corresponding time and frequency translations yield (2.48). ■ ■

In quantum mechanics, this theorem shows that we cannot reduce arbitrarily the uncertainty as to the position and the momentum of a free particle. In signal processing, the modulated Gaussians (2.48) that have a minimum joint time-frequency localization are called Gabor chirps. As expected, they are smooth functions with a fast time asymptotic decay.

Compact Support Despite the Heisenberg uncertainty bound, we might still be able to construct a function of compact support whose Fourier transform has a compact support. Such a function would be very useful in constructing a finite impulse response filter with a band-limited transfer function. Unfortunately, the following theorem proves that it does not exist.

Theorem 2.7. *If $f \neq 0$ has a compact support then $\hat{f}(\omega)$ cannot be zero on a whole interval. Similarly, if $\hat{f} \neq 0$ has a compact support then $f(t)$ cannot be zero on a whole interval.*

Proof. We prove only the first statement, since the second is derived from the first by applying the Fourier transform. If $\hat{f}$ has a compact support included in $[-b, b]$ then

$$f(t) = \frac{1}{2\pi} \int_{-b}^{b} \hat{f}(\omega) \, \exp(i\omega t) \, d\omega. \tag{2.53}$$

If $f(t) = 0$ for $t \in [c, d]$, by differentiating n times under the integral at $t_0 = (c + d)/2$, we obtain

$$f^{(n)}(t_0) = \frac{1}{2\pi} \int_{-b}^{b} \hat{f}(\omega) \, (i\omega)^n \, \exp(i\omega t_0) \, d\omega = 0. \tag{2.54}$$

Since

$$f(t) = \frac{1}{2\pi} \int_{-b}^{b} \hat{f}(\omega) \, \exp[i\omega(t - t_0)] \, \exp(i\omega t_0) \, d\omega, \tag{2.55}$$

developing $\exp[i\omega(t - t_0)]$ as an infinite series yields for all $t \in \mathbb{R}$

$$f(t) = \frac{1}{2\pi} \sum_{n=0}^{+\infty} \frac{[i(t - t_0)]^n}{n!} \int_{-b}^{b} \hat{f}(\omega) \, \omega^n \, \exp(i\omega t_0) \, d\omega = 0. \tag{2.56}$$

This contradicts our assumption that $f \neq 0$. ■ ■

2.3.3 Total Variation

The total variation measures the total amplitude of signal oscillations. It plays an important role in image processing, where its value depends on the length of the image level sets. We show that a low-pass filter can considerably amplify the total variation by creating Gibbs oscillations.

Variations and Oscillations If f is differentiable, its total variation is defined by

$$\|f\|_V = \int_{-\infty}^{+\infty} |f'(t)| \, dt \, . \tag{2.57}$$

If $\{x_p\}_p$ are the abscissa of the local extrema of f where $f'(x_p) = 0$, then

$$\|f\|_V = \sum_p |f(x_{p+1}) - f(x_p)| \, .$$

It thus measures the total amplitude of the oscillations of f. For example, if $f(t) = \exp(-t^2)$, then $\|f\|_V = 2$. If $f(t) = \sin(\pi t)/(\pi t)$, then f has a local extrema at $x_p \in [p, p + 1]$ for any $p \in \mathbb{Z}$. Since $|f(x_{p+1}) - f(x_p)| \sim |p|^{-1}$, we derive that $\|f\|_V = +\infty$.

The total variation of non-differentiable functions can be calculated by considering the derivative in the general sense of distributions [60, 73]. This is equivalent to approximating the derivative by a finite difference on an interval h that goes to zero:

$$\|f\|_V = \lim_{h \to 0} \int_{-\infty}^{+\infty} \frac{|f(t) - f(t - h)|}{|h|} \, dt \, . \tag{2.58}$$

The total variation of discontinuous functions is thus well defined. For example, if $f = \mathbf{1}_{[a,b]}$ then (2.58) gives $\|f\|_V = 2$. We say that f has a *bounded variation* if $\|f\|_V < +\infty$.

Whether f' is the standard derivative of f or its generalized derivative in the sense of distributions, its Fourier transform is $\widehat{f'}(\omega) = i\omega \hat{f}(\omega)$. Hence

$$|\omega| \, |\hat{f}(\omega)| \leqslant \int_{-\infty}^{+\infty} |f'(t)| dt = \|f\|_V \, ,$$

which implies that

$$|\hat{f}(\omega)| \leqslant \frac{\|f\|_V}{|\omega|} \ . \tag{2.59}$$

However, $|\hat{f}(\omega)| = O(|\omega|^{-1})$ is not a sufficient condition to guarantee that f has bounded variation. For example, if $f(t) = \sin(\pi t)/(\pi t)$, then $\hat{f} = \mathbf{1}_{[-\pi,\pi]}$ satisfies $|\hat{f}(\omega)| \leqslant \pi|\omega|^{-1}$ although $\|f\|_V = +\infty$. In general, the total variation of f cannot be evaluated from $|\hat{f}(\omega)|$.

Discrete Signals Let $f_N[n] = f \star \phi_N(n/N)$ be a discrete signal obtained with an averaging filter $\phi_N(t) = \mathbf{1}_{[0,N^{-1}]}(t)$, and a uniform sampling at intervals N^{-1}. The discrete total variation is calculated by approximating the signal derivative by a finite difference over the sampling distance $h = N^{-1}$, and replacing the integral (2.58) by a Riemann sum, which gives:

$$\|f_N\|_V = \sum_n |f_N[n] - f_N[n-1]| \ . \tag{2.60}$$

If n_p are the abscissa of the local extrema of f_N, then

$$\|f_N\|_V = \sum_p |f_N[n_{p+1}] - f_N[n_p]| \ .$$

The total variation thus measures the total amplitude of the oscillations of f. In accordance with (2.58), we say that the discrete signal has a *bounded variation* if $\|f_N\|_V$ is bounded by a constant independent of the resolution N.

Gibbs Oscillations Filtering a signal with a low-pass filter can create oscillations that have an infinite total variation. Let $f_\xi = f \star \phi_\xi$ be the filtered signal obtained with an ideal low-pass filter whose transfer function is $\hat{\phi}_\xi = \mathbf{1}_{[-\xi,\xi]}$. If $f \in \mathbf{L^2}(\mathbb{R})$, then f_ξ converges to f in $\mathbf{L^2}(\mathbb{R})$ norm: $\lim_{\xi\to+\infty} \|f - f_\xi\| = 0$. Indeed, $\hat{f}_\xi = \hat{f}\,\mathbf{1}_{[-\xi,\xi]}$ and the Plancherel formula (2.26) implies that

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

which goes to zero as ξ increases. However, if f is discontinuous in t_0, then we show that f_ξ has Gibbs oscillations in the neighborhood of t_0, which prevents $\sup_{t\in\mathbb{R}} |f(t) - f_\xi(t)|$ from converging to zero as ξ increases.

Let f be a bounded variation function $\|f\|_V < +\infty$ that has an isolated discontinuity at t_0, with a left limit $f(t_0^-)$ and right limit $f(t_0^+)$. It is decomposed as a sum of f_c, which is continuous in the neighborhood of t_0, plus a Heaviside step of amplitude $f(t_0^+) - f(t_0^-)$:

$$f(t) = f_c(t) + [f(t_0^+) - f(t_0^-)]\, u(t - t_0),$$

with

$$u(t) = \begin{cases} 1 & \text{if } t \geqslant 0 \\ 0 & \text{otherwise} \end{cases} \ . \tag{2.61}$$

Hence

$$f_\xi(t) = f_c \star \phi_\xi(t) + [f(t_0^+) - f(t_0^-)]\, u \star \phi_{\,\xi}(t - t_0). \tag{2.62}$$

Since f_c has bounded variation and is uniformly continuous in the neighborhood of t_0, one can prove (Exercise 2.15) that $f_c \star \phi_\xi(t)$ converges uniformly to $f_c(t)$ in a neighborhood of t_0. The following theorem shows that this is not true for $u \star \phi_\xi$, which creates Gibbs oscillations.

Theorem 2.8 (Gibbs)**.** *For any $\xi > 0$,*

$$u \star \phi_\xi(t) = \int_{-\infty}^{\xi t} \frac{\sin x}{\pi x}\, dx. \tag{2.63}$$

Proof. The impulse response of an ideal low-pass filter, calculated in (2.29), is $\phi_\xi(t) = \sin(\xi t)/(\pi t)$. Hence

$$u \star \phi_\xi(t) = \int_{-\infty}^{+\infty} u(\tau) \frac{\sin \xi(t-\tau)}{\pi(t-\tau)}\, d\tau = \int_0^{+\infty} \frac{\sin \xi(t-\tau)}{\pi(t-\tau)}\, d\tau.$$

The change of variable $x = \xi(t-\tau)$ gives (2.63). ■ ■

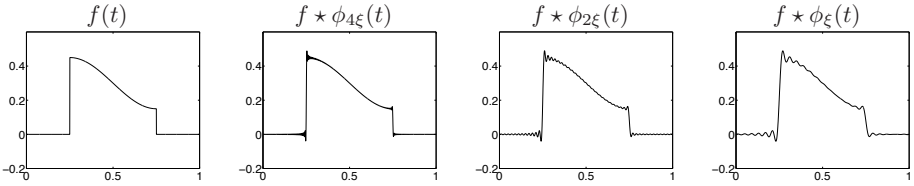


Figure 2.1: Gibbs oscillations created by low-pass filters with cut-off frequencies that decrease from left to right.

The function

$$s(\xi t) = \int_{-\infty}^{\xi t} \frac{\sin x}{\pi x}\, dx$$

is a sigmoid that increases from 0 at $t = -\infty$ to 1 at $t = +\infty$, with $s(0) = 1/2$. It has oscillations of period π/ξ, which are attenuated when the distance to 0 increases, but their total variation is infinite: $\|s\|_V = +\infty$. The maximum amplitude of the Gibbs oscillations occurs at $t = \pm\pi/\xi$, with an amplitude independent of ξ:

$$A = s(\pi) - 1 = \int_{-\infty}^{\pi} \frac{\sin x}{\pi x} dx - 1 \approx 0.045\ .$$

Inserting (2.63) in (2.62) shows that

$$f(t) - f_\xi(t) = [f(t_0^+) - f(t_0^-)]\, s(\xi(t-t_0)) + \varepsilon(\xi, t)\ , \tag{2.64}$$

where $\lim_{\xi \to +\infty} \sup_{|t-t_0|<\alpha} |\varepsilon(\xi, t)| = 0$ in some neighborhood of size $\alpha > 0$ around t_0. The sigmoid $s(\xi(t-t_0))$ centered at t_0 creates a maximum error of fixed amplitude for all ξ. This is seen in Figure 2.1, where the Gibbs oscillations have an amplitude proportional to the jump $f(t_0^+) - f(t_0^-)$ at all frequencies ξ.

Image Total Variation The total variation of an image $f(x_1, x_2)$ depends on the amplitude of its variations as well as the length of the contours along which they occur. Suppose that $f(x_1, x_2)$ is differentiable. The total variation is defined by

$$\|f\|_V = \int\int |\vec{\nabla} f(x_1, x_2)|\, dx_1\, dx_2\ , \tag{2.65}$$

where the modulus of the gradient vector is

$$|\vec{\nabla} f(x_1, x_2)| = \left(\left| \frac{\partial f(x_1, x_2)}{\partial x_1} \right|^2 + \left| \frac{\partial f(x_1, x_2)}{\partial x_2} \right|^2 \right)^{1/2}.$$

As in one dimension, the total variation is extended to discontinuous functions by taking the derivatives in the general sense of distributions. An equivalent norm is obtained by approximating the partial derivatives by finite differences:

$$|\Delta_h f(x_1, x_2)| = \left(\left| \frac{f(x_1, x_2) - f(x_1 - h, x_2)}{h} \right|^2 + \left| \frac{f(x_1, x_2) - f(x_1, x_2 - h)}{h} \right|^2 \right)^{1/2}.$$

One can verify that

$$\|f\|_V \leqslant \lim_{h\to 0} \int\int |\Delta_h f(x_1,x_2)|\, dx_1\, dx_2 \leqslant \sqrt{2}\, \|f\|_V. \tag{2.66}$$

The finite difference integral gives a larger value when $f(x_1,x_2)$ is discontinuous along a diagonal line in the (x_1,x_2) plane.

The total variation of f is related to the length of it level sets. Let us define

$$\Omega_y = \{(x_1,x_2) \in \mathbb{R}^2 \quad : \quad f(x_1,x_2) > y\}\ .$$

If f is continuous then the boundary $\partial\Omega_y$ of Ω_y is the level set of all (x_1,x_2) such that $f(x_1,x_2) = y$. Let $H^1(\partial\Omega_y)$ be the length of $\partial\Omega_y$. Formally, this length is calculated in the sense of the mono-dimensional Hausdorff measure. The following theorem relates the total variation of f to the length of its level sets.

Theorem 2.9 (Co-area Formula)**.** *If* $\|f\|_V < +\infty$ *then*

$$\|f\|_V = \int_{-\infty}^{+\infty} H^1(\partial\Omega_y)\, dy. \tag{2.67}$$

Proof. The proof is a highly technical result that is given in [73]. We give an intuitive explanation when f is continuously differentiable. In this case $\partial\Omega_y$ is a differentiable curve $x(y,s) \in \mathbb{R}^2$, which is parameterized by the arc-length s. Let $\vec{\tau}(x)$ be the vector tangent to this curve in the plane. The gradient $\vec{\nabla} f(x)$ is orthogonal to $\vec{\tau}(x)$. The Frenet coordinate system along $\partial\Omega_y$ is composed of $\vec{\tau}(x)$ and of the unit vector $\vec{n}(x)$ parallel to $\vec{\nabla} f(x)$. Let ds and dn be the Lebesgue measures in the direction of $\vec{\tau}$ and $\vec{n}$. We have

$$|\vec{\nabla} f(x)| = \vec{\nabla} f(x)\,.\,\vec{n} = \frac{dy}{dn}\ , \tag{2.68}$$

where dy is the differential of amplitudes across level sets. The idea of the proof is to decompose the total variation integral over the plane as an integral along the level sets and across level sets, which we write:

$$\|f\|_V = \int\int |\vec{\nabla} f(x_1,x_2)|\, dx_1\, dx_2 = \int \int_{\partial\Omega_y} |\vec{\nabla} f(x(y,s))|\, ds\, dn. \tag{2.69}$$

By using (2.68) we can get

$$\|f\|_V = \int \int_{\partial\Omega_y} ds\, dy\ .$$

But $\int_{\partial\Omega_y} ds = H^1(\partial\Omega_y)$ is the length of the level set, which justifies (2.67). ■ ■

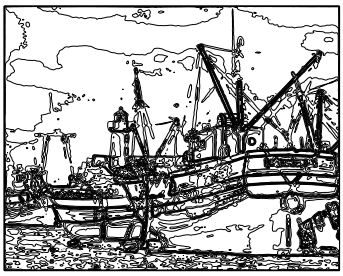
The co-area formula gives an important geometrical interpretation of the total image variation. Images are uniformly bounded so the integral (2.67) is calculated over a finite interval and is proportional to the average length of level sets. It is finite as long as the level sets are not fractal curves. Let $f = \alpha\, \mathbf{1}_\Omega$ be proportional to the indicator function of a set $\Omega \subset \mathbb{R}^2$ which has a boundary $\partial\Omega$ of length L. The co-area formula (2.9) implies that $\|f\|_V = \alpha\, L$. In general, bounded variation images must have step edges of finite length.

Discrete Images A camera measures light intensity with photoreceptors that perform an averaging and a uniform sampling over a grid that is supposed to be uniform. For a resolution N, the sampling interval is N^{-1}. The resulting image can be written $f_N[n_1,n_2] = f \star \phi_N(n_1/N, n_2/N)$, where $\phi_N = \mathbf{1}_{[0,N^{-1}]^2}$ and f is the averaged analog image. Its total variation is defined by approximating derivatives by finite differences and the integral (2.66) by a Riemann sum:

$$\|f_N\|_V = \frac{1}{N} \sum_{n_1} \sum_{n_2} \Big(\Big| f_N[n_1,n_2] - f_N[n_1-1,n_2] \Big|^2 + \Big| f_N[n_1,n_2] - f_N[n_1,n_2-1] \Big|^2 \Big)^{1/2}\ . \tag{2.70}$$

In accordance with (2.66) we say that the image has bounded variation if $\|f_N\|_V$ is bounded by a constant independent of the resolution N. The co-area formula proves that it depends on the

(a)



(b)

Figure 2.2: (a): The total variation of this image remains nearly constant when the resolution N increases. (b): Level sets $\partial\Omega_y$ obtained by sampling uniformly the amplitude variable y.

length of the level sets as the image resolution increases. The $\sqrt{2}$ upper bound factor in (2.66) comes from the fact that the length of a diagonal line can be increased by $\sqrt{2}$ if it is approximated by a zig-zag line that remains on the horizontal and vertical segments of the image sampling grid. Figure 2.2(a) shows a bounded variation image and Figure 2.2(b) displays the level sets obtained by discretizing uniformly the amplitude variable y. The total variation of this image remains nearly constant as the resolution varies.

2.4 Two-Dimensional Fourier Transform

The Fourier transform in $\mathbb{R}^n$ is a straightforward extension of the one-dimensional Fourier transform. The two-dimensional case is briefly reviewed for image processing applications. The Fourier transform of a two-dimensional integrable function $f \in \mathbf{L^1}(\mathbb{R}^2)$ is

$$\hat{f}(\omega_1, \omega_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x_1, x_2)\, \exp[-i(\omega_1 x_1 + \omega_2 x_2)]\, dx_1\, dx_2. \tag{2.71}$$

In polar coordinates $\exp[i(\omega_1 x + \omega_2 y)]$ can be rewritten

$$\exp[i(\omega_1 x_1 + \omega_2 x_2)] = \exp[i\xi(x_1 \cos\theta + x_2 \sin\theta)]$$

with $\xi = \sqrt{\omega_1^2 + \omega_2^2}$. It is a plane wave that propagates in the direction of θ and oscillates at the frequency ξ. The properties of a two-dimensional Fourier transform are essentially the same as in one dimension. We summarize a few important results. We write $\omega = (\omega_1, \omega_2)$, $x = (x_1, x_2)$, $\omega \cdot x = \omega_1 x_1 + \omega_2 x_2$ and $\iint f(x_1, x_2) dx_1\, dx_2 = \iint f(x)\, dx$.

- If $f \in \mathbf{L^1}(\mathbb{R}^2)$ and $\hat{f} \in \mathbf{L^1}(\mathbb{R}^2)$ then

$$f(x) = \frac{1}{4\pi^2} \iint \hat{f}(\omega)\, \exp[i(\omega \cdot x)]\, d\omega. \tag{2.72}$$

- If $f \in \mathbf{L^1}(\mathbb{R}^2)$ and $h \in \mathbf{L^1}(\mathbb{R}^2)$ then the convolution

$$g(x) = f \star h(x) = \iint f(u)\, h(x-u)\, du$$

has a Fourier transform

$$\hat{g}(\omega) = \hat{f}(\omega)\, \hat{h}(\omega). \tag{2.73}$$

- The Parseval formula proves that

$$\iint f(x)\, g^*(x)\, dx = \frac{1}{4\pi^2} \iint \hat{f}(\omega)\, \hat{g}^*(\omega)\, d\omega\ . \tag{2.74}$$

If $f = g$, we obtain the Plancherel equality

$$\iint |f(x)|^2 \, dx = \frac{1}{4\pi^2} \iint |\hat{f}(\omega)|^2 \, d\omega \, . \tag{2.75}$$

The Fourier transform of a finite energy function thus has finite energy. With the same density based argument as in one dimension, energy equivalence makes it possible to extend the Fourier transform to any function $f \in \mathbf{L^2}(\mathbb{R}^2)$.

- If $f \in \mathbf{L^2}(\mathbb{R}^2)$ is separable, which means that

$$f(x) = f(x_1, x_2) = g(x_1)\, h(x_2),$$

then its Fourier transform is

$$\hat{f}(\omega) = \hat{f}(\omega_1, \omega_2) = \hat{g}(\omega_1)\, \hat{h}(\omega_2),$$

where $\hat{h}$ and $\hat{g}$ are the one-dimensional Fourier transforms of g and h. For example, the indicator function

$$f(x_1, x_2) = \begin{cases} 1 & \text{if } |x_1| \leqslant T \ , \ |x_2| \leqslant T \\ 0 & \text{otherwise} \end{cases} = \mathbf{1}_{[-T,T]}(x_1) \times \mathbf{1}_{[-T,T]}(x_2)$$

is a separable function whose Fourier transform is derived from (2.28):

$$\hat{f}(\omega_1, \omega_2) = \frac{4 \sin(T\omega_1)\, \sin(T\omega_2)}{\omega_1\, \omega_2}.$$

- If $f(x_1, x_2)$ is rotated by θ:

$$f_\theta(x_1, x_2) = f(x_1 \cos\theta - x_2 \sin\theta, x_1 \sin\theta + x_2 \cos\theta),$$

then its Fourier transform is rotated by θ:

$$\hat{f}_\theta(\omega_1, \omega_2) = \hat{f}(\omega_1 \cos\theta - \omega_2 \sin\theta, \omega_1 \sin\theta + \omega_2 \cos\theta). \tag{2.76}$$

Radon Transform A Radon transform computes integrals of $f \in \mathbf{L^2}(\mathbb{R}^2)$ along rays. It is provides a good model for some tomographic systems such as X-ray measurements in medical imaging. Inverting the Radon transform is then needed to reconstruct the 2D or 3D body from these integrals.

Let us write $\tau_\theta = (\cos\theta, \sin\theta)$. A ray $\Delta_{t,\theta}$ is a line defined by its equation

$$x \cdot \tau_\theta = x_1 \cos\theta + x_2 \sin\theta = t \ .$$

The projection p_θ of f along a parallel line of orientation θ is defined by

$$\forall\, \theta \in [0, \pi), \forall\, t \in \mathbb{R}, \quad p_\theta(t) = \int_{\Delta_{t,\theta}} f(x)\, ds = \iint f(x)\, \delta(x \cdot \tau_\theta - t)\, dx, \tag{2.77}$$

where δ is the Dirac distribution. The Radon transform maps $f(x)$ to $p_\theta(t)$ for $\theta \in [0, \pi)$. In medical imaging applications, a scanner is rotated around an object to compute the projection p_θ for many angles $\theta \in [0, \pi)$, as illustrated in Figure 2.3. The Fourier slice theorem relates the Fourier transform of p_θ to slices of the Fourier transform of f.

Theorem 2.10 (Fourier slice)**.** *The Fourier transform of projections satisfies*

$$\forall\, \theta \in [0, \pi) \ , \ \forall\, \xi \in \mathbb{R} \quad \hat{p}_\theta(\xi) = \hat{f}(\xi \cos\theta, \xi \sin\theta).$$

Proof. The Fourier transform of the projection is

$$\begin{aligned} \hat{p}_\theta(\xi) &= \int_{-\infty}^{+\infty} \left(\iint f(x)\, \delta(x \cdot \tau_\theta - t)\, dx \right) e^{-it\xi} dt \\ &= \int \int f(x) \exp\left(-i(x \cdot \tau_\theta)\xi\right) dx = \hat{f}(\xi \tau_\theta). \end{aligned}$$

■

An image f can be recovered from its projections p_θ thanks to the projection slice theorem. Indeed, the Fourier transform $\hat{f}$ is known along each ray of direction θ and f is thus obtained with the 2D inverse Fourier transform (2.71). The following back projection theorem gives an inversion formula.

Theorem 2.11 (Back projection)**.** *The image f is recovered using a one-dimensional filter $h(t)$:*

$$f(x) = \frac{1}{2\pi} \int_0^\pi p_\theta * h(x \cdot \tau_\theta)\, d\theta \quad \textit{with} \quad \hat{h}(\xi) = |\xi| \ .$$

Proof. The inverse Fourier transform (2.72) in polar coordinates $(\omega_1, \omega_2) = (\xi \cos\theta, \xi \sin\theta)$, with $d\omega_1\, d\omega_2 = \xi\, d\theta\, d\xi$, can be written

$$f(x) = \frac{1}{4\pi^2} \int_0^{+\infty} \int_0^{2\pi} \hat{f}(\xi \cos\theta, \xi \sin\theta) \exp\left(i(x \cdot \tau_\theta)\xi\right)\, \xi\, d\theta\, d\xi \ .$$

Using the Fourier slice Theorem 2.10 with $p_{\theta+\pi}(t) = p_\theta(-t)$, this is rewritten as

$$f(x) = \frac{1}{2\pi} \int_0^\pi \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} |\xi|\, \hat{p}_\theta(\xi) \exp\left(i(x \cdot \tau_\theta)\xi\right) d\xi \right) d\theta \ .$$

The inner integral is the inverse Fourier transform of $\hat{p}_\theta(\xi)\,|\xi|$ evaluated at $x \cdot \tau_\theta \in \mathbb{R}$. The convolution formula (2.73) shows that it is equal to $p_\theta * h(x \cdot \tau_\theta)$. ■

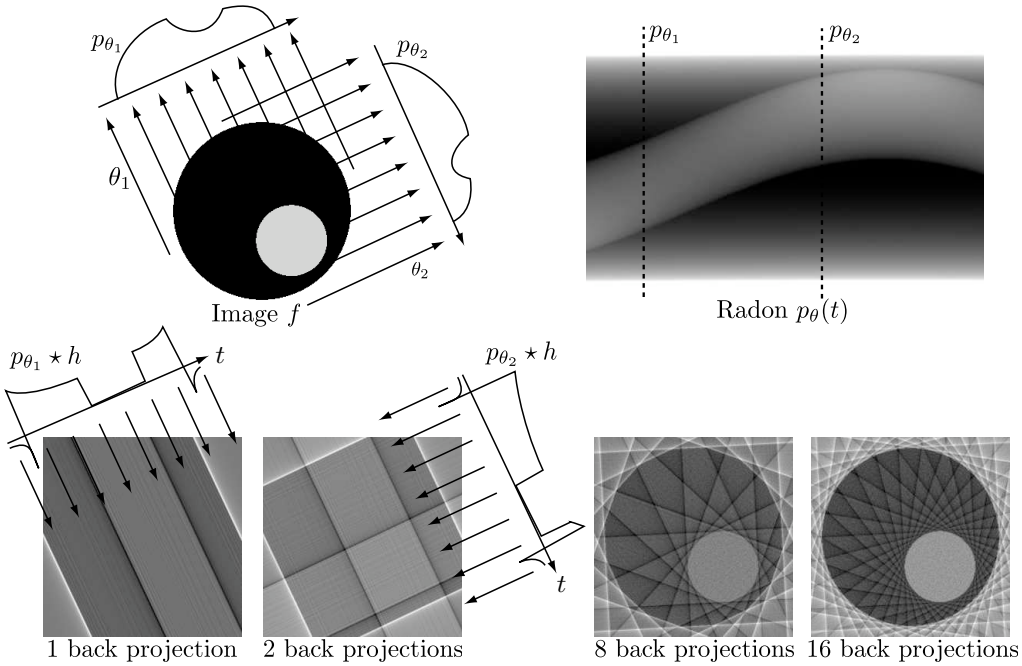


Figure 2.3: The Radon transform and its reconstruction with an increasing number of back projections.

In medical imaging applications, only a limited number of projections is available, and the Fourier transform $\hat{f}$ is thus partially known. In this case, an approximation of f can still be recovered by summing the corresponding filtered back projections $p_\theta * h(x \cdot \tau_\theta)$. Figure 2.3 describes this process, and shows the reconstruction of an image with a geometric object, using an increasing number of evenly spaced projections. Section 13.3 describes a non-linear super-resolution reconstruction algorithm, that recovers a more precise image by using a sparse representation.