Math 994-003: Computational Harmonic Analysis and Data Science

Matthew Hirn

January 23, 2018

Abstract

This course will cover aspects of modern computational harmonic analysis at the interface of signal processing and data science. A central theme of the course is to find "good" representations of functional data (e.g., time series, images, etc), where the quality of the representation is measured through notions of sparsity, characterization of certain functional classes, and eventually empirical data driven measures.

The prologue of the course will cover the rudiments of Fourier analysis and discrete signal processing. The shortcomings of the Fourier transform will motivate us to study localized time-frequency representations of functions, which will introduce the windowed Fourier transform as well as the continuous and dyadic wavelet transforms. Unlike the Fourier transform, which characterizes only the global regularity of a function, wavelet transforms characterize the local regularity of functions, and we will prove fundamental results along these lines. Windowed Fourier and wavelet transforms will be placed in a more general mathematical context via the study of redundant dictionaries and frame theory. Motivated in part by the sparsity of wavelet transforms, we will then aim to understand how to find sparse representations in general dictionaries, and at the conclusion of this part of the course look at recent methods that learn, in a data driven fashion, dictionaries of functions that yield sparse representations. Dictionary learning in turn leads to the study of more complicated learning models; convolutional neural networks are a natural place to turn. In the final part of the course we will study these networks, as well as mathematically tractable models for them (i.e. ones in which we can prove theorems) based upon nonlinear cascades of semi-discrete frame operators.

The primary textbook for the course will be *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd edition, by Stephane Mallat [1]. The course may also draw a bit of material from *Wavelets and Operators*, by Yves Meyer

[2]. The final part of the course on convolutional neural networks will be based on current papers in the field. All parts of the course (minus the prologue) will highlight current research and papers.

The course will assume knowledge of real analysis (Lebesgue integration, $\mathbf{L}^p$ spaces, Banach and Hilbert spaces).

1 Sparse Representations

Chapter 1 of A Wavelet Tour of Signal Processing.

Exercise 1. Read Chapter 1 (*Sparse Representations*) of *A Wavelet Tour of Signal Processing.* It gives a nice overview of the book and will give you a good perspective on computational harmonic analysis heading into the course.

Exercise 2. Read the appendices in *A Wavelet Tour of Signal Processing*, as we will not cover these in class. We will immediately need some of the material contained in them.

Remark 1. The integral we use in this course will be the Lebesgue integral, which is usually taught in a first year graduate course in real analysis. However, if these are unfamiliar to you, you may replace some of the results with Riemann integrals from Calculus and assume that the generic functions f, g, h, etc. are Schwartz class functions. For more details on the Schwartz class and Fourier integrals, see [3].

2 The Fourier Kingdom

Chapter 2 of A Wavelet Tour of Signal Processing.

2.1 Linear time-invariant filtering

Section 2.1 of A Wavelet Tour of Signal Processing.

Fourier analysis originates with the work of Joseph Fourier, who was studying the heat equation:

$$\partial_t u = \Delta u$$
$$u(x, 0) = f(x)$$

This is a linear partial differential equation, and the Laplacian Δ is a linear time invariant operator. Let $f_\tau(t) = f(t-\tau)$ be the translation of f by τ; if t is time, then this is a time delay by τ. An operator L is *time (shift) invariant* if it commutes with the time delay of any function,

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

As we shall see all linear, continuous shift invariant operators L are diagonalized by the complex exponentials $e_\omega(t) = e^{i\omega t}$. To see this, recall the convolution of two functions f, g:

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

Now let $\delta(t)$ be a Dirac (centered at zero), and $\delta_u(t) = \delta(t-u)$ be a Dirac centered at u. We have:

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

Since L is continuous and linear,

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

Let h be the impulse response of L, defined as

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

Since L is shift invariant, we have

$$L\delta_u(t) = h(t-u)$$

and therefore

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

Thus *every* continuous, linear shift invariant operator is equivalent to a convolution with an impulse response h.

We can now use this fact to show our original goal, which was that the complex exponential functions $e_\omega(t) = e^{i\omega t}$ diagonalize L. This will in turn motivate the study of Fourier integrals. We have:

$$Le_\omega(t) = \int_{-\infty}^{+\infty} h(u)e^{i\omega(t-u)}\,du = e^{it\omega}\underbrace{\int_{-\infty}^{+\infty} h(u)e^{-i\omega u}\,du}_{\widehat{h}(\omega)} = \widehat{h}(\omega)e_\omega(t).$$

Thus $e_\omega(t)$ is an eigenfunction of L with eigenvalue $\widehat{h}(\omega)$, if $\widehat{h}(\omega)$ exists. The value $\widehat{h}(\omega)$ is the *Fourier transform* of h at the frequency ω. Since the functions $e_\omega(t) = e^{i\omega t}$ are the eigenfunctions of shift invariant operators, we would like to decompose any function f as a sum or integral of these functions. This will then allow us to write Lf directly in terms of the eigenvalues of L (as you do in linear algebra when you are able to diagonalize a matrix/operator on a finite dimensional vector space). We'll come back to this in a bit.

Exercise 3. Read Section 2.1 of *A Wavelet Tour of Signal Processing.*

2.2 Fourier integrals

Section 2.2 of A Wavelet Tour of Signal Processing.

The Fourier transform of a function f is defined as:

$$\mathcal{F}(f)(\omega) = \widehat{f}(\omega) = \int_{-\infty}^{+\infty} f(t)e^{-i\omega t}\,dt \tag{1}$$

As we alluded to earlier, we need to know when this integral converges. To that end, define:

$$\mathbf{L}^p(\mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{C} : \int_{-\infty}^{+\infty} |f(t)|^p\,dt < +\infty \right\}, \quad 0 < p < \infty$$

The space $\mathbf{L}^p(\mathbb{R})$ is a Banach space with norm:

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

The space $\mathbf{L}^2(\mathbb{R})$ is special, as it is in fact a Hilbert space with inner product

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

where we use $g^*(t)$ to denote the complex conjugate of $g(t)$. We also define $\mathbf{L}^\infty(\mathbb{R})$. Set:

$$\|f\|_\infty = \operatorname*{ess\,sup}_{t \in \mathbb{R}} |f(t)|$$

The value $\|f\|_\infty$ is the smallest number M, $0 \leq M \leq +\infty$, such that $|f(t)| \leq M$ for almost every $t \in \mathbb{R}$; if f is continuous, it is the smallest number M such that $|f(t)| \leq M$ for all $t \in \mathbb{R}$. It thus measures whether f is bounded or not. The space $\mathbf{L}^\infty(\mathbb{R})$ is the space of bounded functions:

$$\mathbf{L}^\infty(\mathbb{R}) = \{f : \|f\|_\infty < +\infty\}$$

We then have:

Proposition 2. *If $f \in \mathbf{L}^1(\mathbb{R})$, then $\widehat{f} \in \mathbf{L}^\infty(\mathbb{R})$.*

Proof. Suppose $f \in \mathbf{L}^1(\mathbb{R})$. We have:

$$|\widehat{f}(\omega)| = \left| \int_{-\infty}^{+\infty} f(t) e^{-i\omega t}\, dt \right| \leq \int_{-\infty}^{+\infty} |f(t) e^{-i\omega t}|\, dt = \int_{-\infty}^{+\infty} |f(t)|\, dt = \|f\|_1 < +\infty$$

□

Recall from Section 2.1 that we would like to write $f(t)$ in terms of $\widehat{f}(\omega)$. This requires a Fourier inversion formula. However, the above proposition only guarantees that $\widehat{f} \in \mathbf{L}^\infty(\mathbb{R})$, which will not help with convergence issues. We thus assume that $\widehat{f} \in \mathbf{L}^1(\mathbb{R})$ as well.

Theorem 3 (*Fourier inversion*)**.** *If $f \in \mathbf{L}^1(\mathbb{R})$ and $\widehat{f} \in \mathbf{L}^1(\mathbb{R})$ then*

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \widehat{f}(\omega) e^{i\omega t}\, d\omega, \quad \textit{for almost every } t \in \mathbb{R} \tag{2}$$

Proof. We first recall two theorems from real analysis. The first is the *Fubini theorem*, which states that if

$$\int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} |f(x_1, x_2)|\, dx_1 \right) dx_2 < +\infty$$

then

$$\begin{aligned} \iint_{\mathbb{R}^2} f(x_1, x_2)\, dx_1\, dx_2 &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x_1, x_2)\, dx_1 \right) dx_2 \\ &= \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x_1, x_2)\, dx_2 \right) dx_1 \end{aligned}$$

The other is the *dominated convergence theorem*. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of functions such that $\lim_{n\to\infty} f_n = f$. If

$$\forall\, n \in \mathbb{N}, \quad |f_n(t)| \leq g(t) \quad \text{and} \quad \int_{-\infty}^{+\infty} g(t)\, dt < +\infty$$

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

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} f_n(t)\, dt = \int_{-\infty}^{+\infty} f(t)\, dt$$

Now we turn to the proof. Plugging in the formula of $\widehat{f}(\omega)$ into the right hand side of (2) yields

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \widehat{f}(\omega) e^{i\omega t} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(u) e^{i\omega(t-u)}\, du \right) d\omega$$

However we cannot apply Fubini directly because the function $F(u,\omega) = f(u)e^{i\omega(t-u)}$ is not integrable in $\mathbb{R}^2$. Therefore we instead consider the following integral:

$$I_\varepsilon(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(u) e^{-\varepsilon^2\omega^2/4} e^{i\omega(t-u)}\, du \right) d\omega$$

The Gaussian yields a new integrand $F_\varepsilon(u,\omega) = f(u)e^{-\varepsilon^2\omega^2/4}e^{i\omega(t-u)}$ which is integrable on $\mathbb{R}^2$, and for which $\lim_{\varepsilon\to 0} F_\varepsilon = F$. We can thus apply the Fubini theorem to $I_\varepsilon(t)$; we do so in two ways. For the first, we integrate with respect to u, giving:

$$I_\varepsilon(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \widehat{f}(\omega) e^{-\varepsilon^2\omega^2/4} e^{i\omega t}\, d\omega.$$

Since

$$\left| \widehat{f}(\omega) e^{-\varepsilon^2\omega^2/4} e^{i\omega t} \right| \leq |\widehat{f}(\omega)|$$

and since $\widehat{f} \in \mathbf{L}^1(\mathbb{R})$, we can apply the dominated convergence theorem to obtain:

$$\lim_{\varepsilon\to 0} I_\varepsilon(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \widehat{f}(\omega) e^{i\omega t}\, d\omega$$

Now compute $I_\varepsilon(t)$ a second way by applying the Fubini theorem and integrating with respect to ω. We get that

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

where

$$\begin{aligned} g_\varepsilon(x) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\varepsilon^2\omega^2/4} e^{ix\omega}\, d\omega \\ &= \frac{1}{\varepsilon\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{\varepsilon}{2\sqrt{\pi}} e^{-\varepsilon^2\omega^2/4} e^{ix\omega}\, d\omega \\ &= \frac{1}{\varepsilon\sqrt{\pi}} e^{-x^2/\varepsilon^2} \end{aligned}$$

To get the last line, we used the fact that that the Fourier transform of $\theta(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-t^2/2\sigma^2}$ is equal to $\widehat{\theta}(\xi) = e^{-\sigma^2\xi^2/2}$. This is a useful identity that you should verify yourself and then remember. Another useful identity is that $\int_{-\infty}^{+\infty} \theta(t)\, dt = 1$. From this latter formula we deduce that

$$\int_{-\infty}^{+\infty} g_\varepsilon(x)\, dx = 1$$

Furthermore, we notice that

$$g_\varepsilon(x) = \varepsilon^{-1} g_1(\varepsilon^{-1}x), \quad g_1(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}$$

Thus the family $\{g_\varepsilon\}_{\varepsilon>0}$ is an *approximate identity*. For general approximate identities one can prove (see [4, Theorem 9.6]):

$$\lim_{\varepsilon\to\infty} \|I_\varepsilon - f\|_1 = 0$$

However this is not quite enough. Since our approximate identity, though, is generated by a Gaussian, it is a particularly nice one, and using [5, Theorem 1.25, p. 13] we obtain

$$\lim_{\varepsilon\to 0} I_\varepsilon(t) = f(t), \quad \text{for almost every } t \in \mathbb{R}$$

Thus the theorem is proved. □

Exercise 4. Prove that the assumptions of the previous Fourier inversion theorem imply that f must be continuous and bounded.

Remark 4. Exercise 4 shows that our Fourier inversion theorem only holds for continuous functions. However, many signals that we encounter will have discontinuities. Thus we will need to extend the theory to discontinuous functions. This will be done by extending the Fourier transform to $\mathbf{L}^2(\mathbb{R})$ (more on this later).

Recall in Section 2.1 that for a linear shift invariant operator L with impulse response h, we wanted to write Lf in terms of the eigenvalues of $\widehat{h}(\omega)$ of L by also being able to compute $\widehat{f}(\omega)$. The previous theorem gives us part of the solution; the other part is given by the *convolution theorem*, which is stated next.

Theorem 5 (*Convolution thoerem*)**.** *Let $f, g \in \mathbf{L}^1(\mathbb{R})$. Then the function $h = f * g \in \mathbf{L}^1(\mathbb{R})$ and*

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

Proof. See p. 37 of *A Wavelet Tour of Signal Processing.* □

Recall now that every linear shift invariant operator L can be written as $Lf = h * f$, where $h = L\delta$. Thus using the Fourier inversion theorem and the convolution theorem we have:

$$Lf(t) = h * f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \widehat{h * f}(\omega)e^{i\omega t}\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \widehat{h}(\omega)\widehat{f}(\omega)e^{i\omega t}\,d\omega$$

Thus at last we see that the sinusoids $e_\omega(t) = e^{i\omega t}$ diagonalize L, with eigenvalues $\widehat{h}(\omega)/2\pi$. To see this consider the case of a finite dimensional vector space V with $\dim(V) = d$ and a diagonalizable linear operator $T : V \to V$ with orthonormal eigenvectors $v_1, \ldots, v_d$ and eigenvalues $\lambda_1, \ldots, \lambda_d$. We then have for any $v \in V$,

$$Tv = \sum_{k=1}^{d} \lambda_k \langle v, v_k \rangle v_k$$

The correspondence is $\lambda_k \leftrightarrow \widehat{h}(\omega)/2\pi$, $\langle v, v_k\rangle \leftrightarrow \widehat{f}(\omega)$, and $v_k \leftrightarrow e^{i\omega t}$.

The Fourier transform has several important properties that are listed in Figure 1.

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

Figure 1: Summary of basic properties of the Fourier transform. Taken from Table 2.1 of *A Wavelet Tour of Signal Processing*.

Exercise 5. Verify all of the properties in Figure 1. No need to turn this one in, but it is important to do these verifications.

Consider now the function

$$f(t) = \mathbf{1}_{[-1,1]}(t) = \begin{cases} 1 & -1 \leq t \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

We can compute the Fourier transform of this function:

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

One can verify that this function is not integrable; we would expect this from Exercise 4 because $f(t)$ is not continuous. However, $\widehat{f}(\omega)$ is square integrable; that is $\widehat{f} \in \mathbf{L}^2(\mathbb{R})$. This motivates extending the Fourier transform to functions $f \in \mathbf{L}^2(\mathbb{R})$. Recall that $\mathbf{L}^2(\mathbb{R})$ is a Hilbert space with inner product

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

We first have the following fundamental results:

Theorem 6 (*Parseval*)**.** *Let* $f, g \in \mathbf{L}^1(\mathbb{R}) \cap \mathbf{L}^2(\mathbb{R})$*. Then:*

$$\langle f, g \rangle = \frac{1}{2\pi} \langle \widehat{f}, \widehat{g} \rangle$$

Proof. See p. 39 of *A Wavelet Tour of Signal Processing.* □

Corollary 7 (*Plancheral*)**.** *Let* $f \in \mathbf{L}^1(\mathbb{R}) \cap \mathbf{L}^2(\mathbb{R})$*. Then:*

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

Note that in the previous theorems, the inner product and norm are computable because we assume $f, g \in \mathbf{L}^2(\mathbb{R})$, but the Fourier transform is only well defined because we assume $f, g \in \mathbf{L}^1(\mathbb{R})$ as well. We would like to remedy this be extending the Fourier transform to all functions $f \in \mathbf{L}^2(\mathbb{R})$, even those for which $f \notin \mathbf{L}^1(\mathbb{R})$. We do this with a *density argument*, which will define the Fourier transform of a function $f \in \mathbf{L}^2(\mathbb{R})$ as the limit of Fourier transforms of functions in $\mathbf{L}^1(\mathbb{R}) \cap \mathbf{L}^2(\mathbb{R})$. A very useful inequality from real analysis, which we will need here, is *Hölder's inequality*:

$$\forall\, f \in \mathbf{L}^p(\mathbb{R}),\ g \in \mathbf{L}^q(\mathbb{R}),\ p, q \in [1, \infty],\ \frac{1}{p} + \frac{1}{q} = 1, \quad \|fg\|_1 \leq \|f\|_p \|g\|_q$$

Now to the density argument. The first thing to note is that $\mathbf{L}^1(\mathbb{R}) \cap \mathbf{L}^2(\mathbb{R})$ is dense in $\mathbf{L}^2(\mathbb{R})$. This means that given an $f \in \mathbf{L}^2(\mathbb{R})$, we can find a family $\{f_n\}_{n \geq 1}$ of functions in $\mathbf{L}^1(\mathbb{R}) \cap \mathbf{L}^2(\mathbb{R})$ that converges to f,

$$\lim_{n \to \infty} \|f - f_n\|_2 = 0$$

In fact it is easy to find a such a family. Define:

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

We have that $f_n \in \mathbf{L}^2(\mathbb{R})$ for all $n \geq 1$ since $|f_n(t)| \leq |f(t)|$ for all $t \in \mathbb{R}$. Furthermore, $f_n \in \mathbf{L}^1(\mathbb{R})$ since by Hölder's inequality we have:

$$\begin{aligned}
\|f_n\|_1 = \int_{-\infty}^{+\infty} |f_n(t)|\, dt &= \int_{-\infty}^{+\infty} |f(t) \mathbf{1}_{[-n,n]}(t)|\, dt \\
&\leq \left(\int_{-\infty}^{+\infty} |f(t)|^2\, dt \right)^{\frac{1}{2}} \left(\int_{-\infty}^{+\infty} |\mathbf{1}_{[-n,n]}(t)|^2 \right)^{\frac{1}{2}} \\
&= \|f\|_2 \left(\int_{-n}^{n} 1\, dt \right)^{\frac{1}{2}} \\
&= \sqrt{2n} \|f\|_2
\end{aligned}$$

We also have that

$$\|f - f_n\|_2 = \left(\int_{|t|>n} |f(t)|^2 \right)^{\frac{1}{2}} \to 0 \text{ as } n \to \infty$$

Now, since $f_n \to f$, the family $\{f_n\}_{n\geq 1}$ is also a *Cauchy sequence*, meaning that for all $\varepsilon > 0$ there exists an N such that if $n, m > N$, then $\|f_n - f_m\|_2 \leq \varepsilon$. Furthermore, since $f_n \in \mathbf{L}^1(\mathbb{R})$, its Fourier transform $\widehat{f_n}$ is well defined. The Plancheral formula (Corollary 7) then yields:

$$\|\widehat{f_n} - \widehat{f_m}\|_2 = \sqrt{2\pi}\|f_n - f_m\|_2$$

Thus since $\{f_n\}_{n\geq 1}$ is a Cauchy sequence, we see that $\{\widehat{f_n}\}_{n\geq 1}$ is a Cauchy sequence as well. Since $\mathbf{L}^2(\mathbb{R})$ is a Hilbert space, it is complete, which means that every Cauchy sequence converges to an element of $\mathbf{L}^2(\mathbb{R})$. Thus there exists an $F \in \mathbf{L}^2(\mathbb{R})$ such that

$$\lim_{n\to\infty} \|F - \widehat{f_n}\|_2 = 0$$

We define the Fourier transform of $f \in \mathbf{L}^2(\mathbb{R})$ as F, and from now on write $\widehat{f} = F$.

One can show the extension of the Fourier transform to $\mathbf{L}^2(\mathbb{R})$ satisfies the convolution theorem (Theorem 5), the Parseval formula (Theorem 6), the Plancheral formula (Corollary 7), and all properties in Figure 1. In particular, the Plancheral formula implies the following. Let $\mathcal{F}(f) = \widehat{f}$, so that $\mathcal{F}$ is the operator that maps a function f to its Fourier transform $\widehat{f}$. We see from the Plancheral formula and the extension of the Fourier transform to $\mathbf{L}^2(\mathbb{R})$ that $\mathcal{F} : \mathbf{L}^2(\mathbb{R}) \to \mathbf{L}^2(\mathbb{R})$, and furthermore that this linear operator is an isometry up to a factory $1/\sqrt{2\pi}$. The operator $\mathcal{F} : \mathbf{L}^2(\mathbb{R}) \to \mathbf{L}^2(\mathbb{R})$ is bijective, and thus is invertible; we therefore have Fourier inversion for $\mathbf{L}^2(\mathbb{R})$ functions as well.

Remark 8. To summarize the Fourier transform can be defined on $\mathbf{L}^1(\mathbb{R})$ in which case we have

$$\mathcal{F} : \mathbf{L}^1(\mathbb{R}) \to \mathbf{L}^\infty(\mathbb{R})$$

with $\|\widehat{f}\|_\infty \leq \|f\|_1$, or on $\mathbf{L}^2(\mathbb{R})$ where we have:

$$\mathcal{F} : \mathbf{L}^2(\mathbb{R}) \to \mathbf{L}^2(\mathbb{R})$$

with $\|f\|_2 = (1/\sqrt{2\pi})\|\widehat{f}\|_2$. It follows then from the Riesz-Thorin Theorem that the Fourier transform can be extended to $\mathbf{L}^p(\mathbb{R})$ for any $1 \leq p \leq 2$, where we have

$$\mathcal{F} : \mathbf{L}^p(\mathbb{R}) \to \mathbf{L}^q(\mathbb{R}), \quad \frac{1}{p} + \frac{1}{q} = 1, \ \ 1 \leq p \leq 2$$

and that

$$\|\widehat{f}\|_p \leq \left(\frac{1}{2\pi}\right)^{\frac{1}{p}} \|f\|_p \tag{3}$$

Equation (3) is called the *Hausdorff–Young Inequality*. Note that in general one only obtains equality for $p = q = 2$, and indeed $\mathcal{F}$ is not an isometry otherwise (up to the constant factor) and is not invertible. Indeed, we saw this for $\mathbf{L}^1(\mathbb{R})$, where in order to get Fourier inversion we had to assume that $\widehat{f} \in \mathbf{L}^1(\mathbb{R})$ as well.

Remark 9. The Dirac $\delta(t)$ is not a function, and hence is not in $\mathbf{L}^1(\mathbb{R})$ nor $\mathbf{L}^2(\mathbb{R})$; it is a distribution, which we will not discuss in this course. However, it will be useful to define the Fourier transform of $\delta(t)$. Recall that for any continuous function f

$$\int_{-\infty}^{+\infty} \delta(t) f(t)\, dt = f(0)$$

We use this to define $\widehat{\delta}(\omega)$ as:

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

A translated Dirac $\delta_\tau(t) = \delta(t - \tau)$ has Fourier transform calculated by evaluating $e^{-i\omega t}$ at $t = \tau$,

$$\widehat{\delta}_\tau(\omega) = \int_{-\infty}^{+\infty} \delta(t - \tau) e^{-i\omega t}\, dt = e^{-i\omega\tau}$$

The *Dirac comb* is a sum of translated Diracs:

$$c(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT) \tag{4}$$

It is used to obtain a discrete sampling of an analogue signal, as we shall see later. Its Fourier transform is:

$$\widehat{c}(\omega) = \sum_{n=-\infty}^{+\infty} e^{-inT\omega}$$

Remarkably, $\widehat{c}(\omega)$ is also a Dirac comb, as the next theorem shows.

Theorem 10 (*Poisson Formula*). *In the sense of distribution equalities,*

$$\sum_{n=-\infty}^{+\infty} e^{-inT\omega} = \frac{2\pi}{T} \sum_{k=-\infty}^{+\infty} \delta\left(\omega - \frac{2\pi k}{T}\right)$$

Proof. See p. 41–42 of *A Wavelet Tour of Signal Processing.* □

Exercise 6. Read Section 2.2 of *A Wavelet Tour of Signal Processing.*

2.3 Regularity and Decay

Section 2.3.1 of A Wavelet Tour of Signal Processing.

The global regularity of f depends on the decay of $|\widehat{f}(\omega)|$ as $\omega \to \infty$. In particular, the smoother the function, the faster the decay of $|\widehat{f}(\omega)|$. The intuition is that smooth functions vary slowly, and thus can be well represented by low frequency modes $e^{i\omega t}$, i.e., those with small values of $|\omega|$. On the other hand, if f is irregular, then it must have sharp transitions which require fast oscillations to capture. We make these intuitions precise with the following two results. First define $\mathbf{C}^n(\mathbb{R})$ as the space of functions with n continuous derivatives; $\mathbf{C}^0(\mathbb{R})$ is the space of continuous functions.

Theorem 11. *Let $f \in \mathbf{L}^1(\mathbb{R}) \cup \mathbf{L}^2(\mathbb{R})$. If there exits a constant C and $\epsilon > 0$ such that*

$$|\widehat{f}(\omega)| \leq \frac{C}{1 + |\omega|^{n+1+\epsilon}}$$

for some $n \in \mathbb{N}$, then $f \in \mathbf{C}^n(\mathbb{R}) \cap \mathbf{L}^\infty(\mathbb{R})$.

Proof. We know from Exercise 4 that if $\widehat{f} \in \mathbf{L}^1(\mathbb{R})$, then f is continuous and bounded. Notice for $n = 0$ we have:

$$\|\widehat{f}\|_1 = \int_{-\infty}^{+\infty} |\widehat{f}(\omega)|\, d\omega \leq \int_{-\infty}^{+\infty} \frac{C}{1 + |\omega|^{1+\epsilon}}\, d\omega < \infty$$

So indeed $f \in \mathbf{C}(\mathbb{R}) \cap \mathbf{L}^\infty(\mathbb{R})$. Now consider $n \in \mathbb{N}$ and $k \leq n$; define the function $F_k(\omega) = (i\omega)^k \widehat{f}(\omega)$. We see that:

$$\|F_k\|_1 \leq \int_{-\infty}^{+\infty} \frac{C|\omega|^k}{1 + |\omega|^{n+1+\epsilon}}\, d\omega < \infty$$

It thus follows that $\mathcal{F}^{-1}(F_k)$ (i.e., the inverse Fourier transform of F_k) is continuous and bounded. But from Figure 1 we know that $\mathcal{F})\|^{-1}(F) = f^{(k)}(t)$, and so the proof is completed. □

Note in particular that if $\widehat{f}$ has compact support, then $f \in \mathbf{C}^\infty(\mathbb{R})$. In the other direction we have:

Theorem 12. *Let $f \in \mathbf{C}^n(\mathbb{R})$ with $f^{(n)} \in \mathbf{L}^1(\mathbb{R})$. Then:*

$$|\widehat{f}(\omega)| \leq \frac{C}{|\omega|^n}$$

for some constant C.

Exercise 7. Prove Theorem 12.

Remark 13. Notice there is a gap between the two theorems relating regularity and decay. This cannot be avoided. Furthermore, we notice that the decay of $|\widehat{f}(\omega)|$ depends upon the *worst* singular behavior of f. Indeed as the function $f(t) = \mathbf{1}_{[-1,1]}(t)$ illustrates, the function is discontinuous and thus its Fourier decay is limited by Theorem 11. However, f has only two singular points. It is often much more desirable to characterize the local regularity of a function. However, the Fourier transform cannot do this since the sinusoids $e^{i\omega t}$ are global functions on $\mathbb{R}$. In order to remedy both of these points, we will need to introduce localized waveforms. We will see later that wavelets do the job.

Exercise 8. Show that the Fourier transform of

$$f(t) = e^{-(a-ib)t^2}, \quad a > 0$$

is

$$\widehat{f}(\omega) = \sqrt{\frac{\pi}{a - ib}} \exp\left(-\frac{a + ib}{4(a^2 + b^2)}\omega^2\right)$$

Exercise 9 (*Riemann-Lebesgue Lemma*). Prove that if $f \in \mathbf{L}^1(\mathbb{R})$, then $\lim_{|\omega|\to\infty} \widehat{f}(\omega) = 0$. Hint: Start with $f \in \mathbf{C}^1(\mathbb{R})$ that have compact support, and use a density argument. This approach uses the standard fact from real analysis that compactly supported $\mathbf{C}^\infty(\mathbb{R})$ functions are dense in $\mathbf{L}^1(\mathbb{R})$. However, if you have not seen this before, it is unsatisfying to use it here to prove the exercise. In this case, consider instead our Gaussian function:

$$g(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$$

which is normalized to have unit integral, $\int_{\mathbb{R}} g(u)\, du = 1$. Define dilations of g as:

$$g_\sigma(u) = \sigma^{-1} g(\sigma^{-1} u), \quad \sigma > 0$$

Notice that $\int_{\mathbb{R}} g_\sigma(u)\, du = 1$ for all $\sigma > 0$. Thus the family $\{g_\sigma\}_{\sigma>0}$ forms an approximate identity. You should now prove for yourself the fact that we stated in class, which is that for any $f \in \mathbf{L}^1(\mathbb{R})$,

$$\lim_{\sigma\to 0} \|f - f * g_\sigma\|_1 = 0 \tag{5}$$

The functions $\{f * g_\sigma\}_{\sigma>0}$ are not compactly supported, but they are in $\mathbf{C}^\infty(\mathbb{R}) \cap \mathbf{L}^1(\mathbb{R})$ (why?). That plus (5) is enough to use a density argument.

Exercise 10. Suppose that $f \in \mathbf{L}^1(\mathbb{R})$ and $f(t) \geq 0$. Verify that $|\widehat{f}(\omega)| \leq \widehat{f}(0)$.

References

[1] Stéphane Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, 3rd edition, 2008.

[2] Yves Meyer. *Wavelets and Operators*, volume 1. Cambridge University Press, 1993.

[3] Elias M. Stein and Rami Shakarchi. *Fourier Analysis: An Introduction*. Princeton Lectures in Analysis. Princeton University Press, 2003.

[4] Richard L. Wheeden and Antoni Zygmund. *Measure and Integral: An Introduction to Real Analysis*. Marcel Dekker, 1977.

[5] Elias M. Stein and Guido Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press, 1971.