2.4 Uncertainty Principle

Section 2.3.2 of A Wavelet Tour of Signal Processing.

The previous section motivates the following question. Can we construct a function f that is well localized in both time and frequency, and if so, how well localized can it be simultaneously in both domains? We know that a Dirac $\delta(t)$ is well localized in space, but $\widehat{\delta}(\omega) = 1$ for all ω, and similarly $e_\xi(t) = e^{i\xi t}$ is not well localized in space, but $\widehat{e_\xi}(\omega) = \delta(\omega - \xi)$. From the previous section, we know that $|\widehat{f}(\omega)|$ decays quickly as $\omega \to \infty$ only if f is very regular. But if f is very regular, it cannot have sharp transitions and thus cannot decay too fast in space as $t \to \infty$.

Similarly, to adjust the spread of a function f while keeping its total energy constant, we can dilate by a factor $s > 0$ with suitable normalization:

$$f_s(t) = s^{-1/2} f(s^{-1}t)$$

If $s < 1$, then the spread of f is decreased, while if $s > 1$ the spread of f is increased. Regardless, the normalization $s^{-1/2}$ insures that $\|f_s\|_2 = \|f\|_2$. The Fourier transform of f_s is:

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

We see that the dilation has the opposite effect on $\widehat{f}$. In particular, if $s < 1$, then the spread of $\widehat{f}$ is increased, while if $s > 1$, the spread of $\widehat{f}$ is decreased. We thus begin to see there is a trade-off between time and frequency localization.

Time and frequency localizations are limited by the (Heisenberg) uncertainty principle, which you may have seen in quantum mechanics as the uncertainty on the position and momentum of a free particle. We will use the framework of quantum mechanics to motivate the following discussion, although it will hold for general functions $f \in \mathbf{L}^2(\mathbb{R})$.

The state of a one-dimensional particle is described by a wave function $f \in \mathbf{L}^2(\mathbb{R})$. The probability density function for the location of this particle to be at t is

$$\frac{1}{\|f\|^2}|f(t)|^2$$

while the probability density function for its momentum to be ω is

$$\frac{1}{2\pi\|f\|^2}|\widehat{f}(\omega)|^2$$

It follows that the average location of the particle is given by

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

while its average momentum is:

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

The variance around the average location u is

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

and the variance around the average momentum is:

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

The variances measure our uncertainty as to the location and momentum of the particle. In particular, the large the variance, the less certain we are. As one may know from quantum mechanics, we cannot know the position and momentum of a particle simultaneously. The following theorem makes this statement precise

Theorem 14 (*Uncertainty Principle*)**.** *The temporal variance and the frequency variance of a function $f \in \mathbf{L}^2(\mathbb{R})$ must satisfy*

$$\sigma_t^2 \sigma_\omega^2 \geq \frac{1}{4}$$

We obtain equality if and only if there exists $(u, \xi, a, b) \in \mathbb{R}^2 \times \mathbb{C}^2$ such that

$$f(t) = ae^{i\xi t - b(t-u)^2}, \quad \mathrm{Real}(b) > 0 \tag{6}$$

Functions (6) are called Gabor functions.

Proof. The proof is relatively simple for functions $f \in \mathcal{S}(\mathbb{R})$, which are Schwartz class functions. The Schwartz class is an important class of functions to know, so we define it now. The space $\mathcal{S}(\mathbb{R})$ consists of all infinitely differentiable functions $f : \mathbb{R} \to \mathbb{C}$ such that $f^{(n)}(t)$ is rapidly decreasing for all $n \geq 0$, that is

$$\sup_{t\in\mathbb{R}} |t|^m |f^{(n)}(t)| < \infty, \quad \forall\, m, n \geq 0$$

An example of a Schwartz class function is the family of functions defined in (6). The Fourier transform, as defined for $\mathbf{L}^1(\mathbb{R})$ functions in (1), is also well defined for $f \in \mathcal{S}(\mathbb{R})$, and furthermore $\mathcal{F} : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$.

Now to the proof. First, note that if the time and frequency averages of f are u and ξ respectively, then the time and frequency averages of $e^{-i\xi t} f(t+u)$ are zero. Thus it is sufficient to prove the theorem for $u = \xi = 0$. First note that if we write $f(t) = f_1(t) + if_2(t)$, then $f'(t) = f_1'(t) + if_2'(t)$,

$$|f(t)|^2 = f_1(t)^2 + f_2(t)^2$$

and

$$\frac{d}{dt}|f(t)|^2 = 2f_1(t)f_1'(t) + 2f_2(t)f_2'(t) = f^*(t)f'(t) + f(t)f'^*(t)$$

We then have using integration by parts:

$$\begin{aligned}
\|f\|^2 &= \int_{-\infty}^{+\infty} |f(t)|^2\,dt \\
&= \underbrace{t|f(t)|^2\Big|_{-\infty}^{+\infty}}_{0\text{ b/c } f\in\mathcal{S}(\mathbb{R})} - \int_{-\infty}^{+\infty} t\frac{d}{dt}|f(t)|^2\,dt \\
&= -\int_{-\infty}^{+\infty} t[f^*(t)f'(t) + f(t)f'^*(t)]\,dt
\end{aligned}$$

Taking the absolute value of both sides yields and using Hölder's inequality (Cauchy-Schwarz) we have:

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

Now use the Plancheral formula (Corollary 7) and the identity $\mathcal{F}(f')(\omega) = i\omega\widehat{f}(\omega)$ to obtain

$$\left(\int_{-\infty}^{+\infty} |f'(t)|^2\,dt\right)^{\frac{1}{2}} = \frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{+\infty} \omega^2|\widehat{f}(\omega)|^2\right)^{\frac{1}{2}} = \|f\|\sigma_\omega$$

Thus we obtain:

$$\|f\|^2 \leq 2\|f\|\sigma_t\|f\|\sigma_\omega$$

from which the desired inequality follows.

For the second part, if $u = \xi = 0$, one can verify that equality holds for $f(t) = ae^{-bt^2}$. Now suppose equality holds. Then we must have equality when we applied the Cauchy-Schwarz inequality. But this can only happen if the two functions are equal, up to a constant, which in this case means that

$$f'(t) = \beta t f(t)$$

The solutions to this differential equation are $f(t) = ae^{\beta t^2/2}$. Setting $-b = \beta/2$ we obtain (6).

The proof can be extended to any $\mathbf{L}^2(\mathbb{R})$ function; see for example [6]. $\square$

The uncertainty principle does not preclude a function having compact support in both time an frequency. However, this is also impossible.

Theorem 15. *Let $f \in \mathbf{L}^1(R) \cup \mathbf{L}^2(\mathbb{R})$. If $f \neq 0$ has a compact support, then $\widehat{f}(\omega)$ cannot be zero on a whole interval. Similarly, if $\widehat{f} \neq 0$ has compact support, then $f(t)$ cannot be zero on a whole interval.*

Proof. We prove the second statement. Suppose that $\widehat{f}$ has compact support, which is included in the interval $[-b, b]$. Then using the Fourier inversion formula, we have

$$f(t) = \frac{1}{2\pi}\int_{-b}^{b} \widehat{f}(\omega)e^{i\omega t}\,d\omega \tag{7}$$

Suppose by contradiction that $f(t) = 0$ for all $t \in [c, d]$. Set $t_0 = (c+d)/2$ and calculate the n^{th} derivative of f at t_0 as:

$$0 = f^{(n)}(t_0) = \frac{1}{2\pi}\int_{-b}^{b} \widehat{f}(\omega)\frac{d}{dt}e^{i\omega t}\Big|_{t=t_0}\,d\omega = \frac{1}{2\pi}\int_{-b}^{b} \widehat{f}(\omega)(i\omega)^n e^{i\omega t_0}\,d\omega$$

Now expand $e^{i\omega t}$ as an infinite Taylor series around t_0:

$$\forall\, t \in \mathbb{R}, \quad e^{i\omega t} = \sum_{n=0}^{\infty} \frac{(i\omega)^n}{n!} e^{i\omega t_0} (t - t_0)^n$$

Now go back to (7) and plug in the Taylor series for $e^{i\omega t}$,

$$f(t) = \sum_{n=0}^{\infty} \frac{(t-t_0)^n}{n!} \underbrace{\frac{1}{2\pi} \int_{-b}^{b} \widehat{f}(\omega)(i\omega)^n e^{i\omega t_0}\, d\omega}_{0} = 0$$

But now we have $f(t) = 0$ for all $t \in \mathbb{R}$, which implies that $\widehat{f}(\omega) = 0$ for all $\omega \in \mathbb{R}$; but this is a contradiction. □

Exercise 11. Read Section 2.3 of *A Wavelet Tour of Signal Processing.*

Exercise 12. Read Section 2.4 of *A Wavelet Tour of Signal Processing.*

Exercise 13. For any $A > 0$, construct a function f such that $\sigma_t(f) > A$ and $\sigma_\omega(f) > A$.

Exercise 14. Let $\|f\|_V$ the total variation of f, which is defined in Section 2.3.3 of *A Wavelet Tour of Signal Processing.* Let $f_\xi = f * \phi_\xi$ with $\widehat{\phi}_\xi(\omega) = \mathbf{1}_{[-\xi,\xi]}(\omega)$. Suppose that f has bounded variation, meaning that $\|f\|_V < \infty$, and that f is continuous in a neighborhood of t_0. Prove that in a neighborhood of t_0, $f_\xi(t)$ converges uniformly to $f(t)$ when $\xi \to \infty$, i.e., prove that

$$\lim_{\xi \to \infty} \sup_{t \in (a,b)} |f(t) - f_\xi(t)| = 0, \quad \text{for some } a < t_0 < b$$

3 Discrete Revolution

Chapter 3 of A Wavelet Tour of Signal Processing.

3.1 Sampling Analog Signals

Section 3.1 of A Wavelet Tour of Signal Processing

Signals $f : \mathbb{R} \to \mathbb{C}$ must be discretized to be stored on a computer. In practice we can only keep a finite amount of information, which means that we can only keep a finite number of samples from f. We will return to this setting in a bit. For now we consider a discrete, countably infinite number of samples from f, given by:

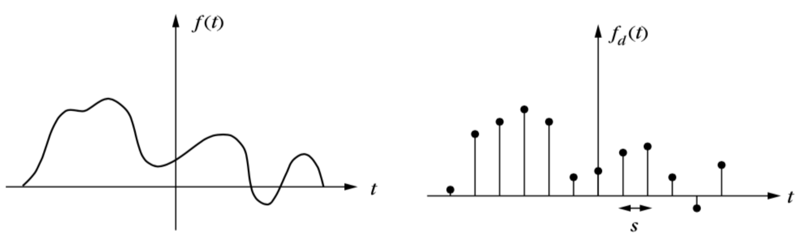
$$\text{Samples } = \{f(ns)\}_{n\in\mathbb{Z}}, \quad s^{-1} = \text{sampling rate} \tag{8}$$

In particular $s = 1$ means we sample every integer, $s = 2$ means we sample every other integer, while $s = 1/2$ means we sample every half integer, and so on.

Assume that f is continuous, so that (8) is well defined. We want to know when we can recover $f(t)$ for all $t \in \mathbb{R}$ from the samples $\{f(ns)\}_{n\in\mathbb{Z}}$. We represent these discrete samples as a sum of weighted Diracs:

$$f_d(t) = \sum_{n\in\mathbb{Z}} f(ns)\delta(t - ns)$$

The signal $f_d : \mathbb{R} \to \mathbb{C}$ is defined for all $t \in \mathbb{R}$ but only takes nonzero values at $t = ns$ for $n \in \mathbb{Z}$. It is thus a discrete sampling of f; see Figure 2.



(a) Function $f(t)$ (b) Discrete function samples $f_d(t)$

Figure 2: A continuous function and its discrete sampled version. Taken from Figure 3.1 of *A Wavelet Tour of Signal Processing.*

The Fourier transform of $f_d(t)$ is:

$$\widehat{f_d}(\omega) = \sum_{n\in\mathbb{Z}} f(ns)e^{-ins\omega}$$

Notice this is a Fourier series; we'll come back to this point later. We first compute $\widehat{f_d}(\omega)$ a second way, which will illuminate the relationship between $\widehat{f}(\omega)$ and $\widehat{f_d}(\omega)$.

Theorem 16. *The Fourier transform of $f_d(t)$ is:*

$$\widehat{f_d}(\omega) = \frac{1}{s}\sum_{k\in\mathbb{Z}} \widehat{f}\left(\omega - \frac{2k\pi}{s}\right)$$

Proof. Define the Dirac comb (see also (4)) as:

$$c(t) = \sum_{n\in\mathbb{Z}} \delta(t - ns)$$

We can rewrite $f_d(t)$ as the multiplication of $f(t)$ with $c(t)$:

$$f_d(t) = f(t)c(t)$$

Using the convolution theorem (Theorem 5), we have:

$$\widehat{f_d}(\omega) = \frac{1}{2\pi}\widehat{f} * \widehat{c}(\omega)$$

But the Poisson Formula (Theorem 10) proves:

$$\widehat{c}(\omega) = \frac{2\pi}{s}\sum_{k\in\mathbb{Z}} \delta\left(\omega - \frac{2\pi k}{s}\right)$$

The theorem then follows immediately. □

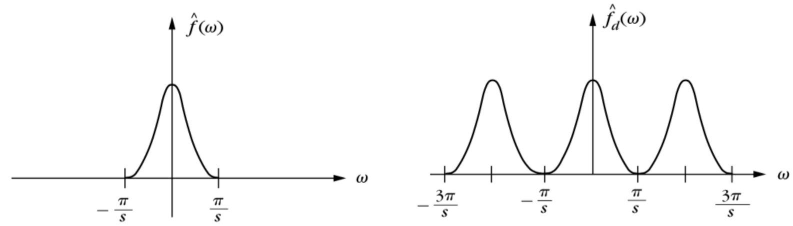
Theorem 16 proves that the Fourier transform $\widehat{f_d}(\omega)$ is obtained by making the Fourier transform $\widehat{f}(\omega)$ $2\pi/s$ periodic. Thus sampling f "periodizes" its frequency response. Figure 3 illustrates the point. The main point here is that if $\operatorname{supp}\widehat{f} \subseteq [-\pi/s, \pi/s]$, then $f(t)$ can be recovered from $f_d(t)$; if $\widehat{f}$ is supported outside of $[-\pi/s, \pi/s]$ then aliasing may occur, in which case we cannot recover $f(t)$ from $f_d(t)$. The next theorem makes precise the first point.

Theorem 17 (*Whittaker–Nyquist–Kotelnikov–Shannon Sampling Theorem*)**.** *If* $\operatorname{supp}\widehat{f} \subseteq [-\pi/s, \pi/s]$, *then*

$$f(t) = f_d * \phi_s(t) = \sum_{n\in\mathbb{Z}} f(ns)\phi_s(t - ns)$$

where

$$\phi_s(t) = \frac{\sin(\pi t/s)}{\pi t/s}$$



(a) Fourier transform $\widehat{f}(\omega)$ of the signal from Figure 2

(b) Fourier transform $\widehat{f_d}(omega)$ of the sampled signal from Figure 2

Figure 3: The Fourier transforms of $\widehat{f}(\omega)$ and $\widehat{f_d}(\omega)$. Taken from Figure 3.1 of *A Wavelet Tour of Signal Processing*.

Proof. If $n \neq 0$, then the support of $\widehat{f}(\omega - 2n\pi/s)$ does not intersect with $\widehat{f}(\omega)$ since $\widehat{f}(\omega) = 0$ for $|\omega| > \pi/s$. Thus by Theorem 16 (see also Figure 3)

$$\widehat{f_d}(\omega) = \frac{\widehat{f}(\omega)}{s}, \quad |\omega| \leq \frac{\pi}{s}$$

The Fourier transform of $\phi_s(t)$ is

$$\widehat{\phi}_s(\omega) = s\mathbf{1}_{[-\pi/s,\pi/s]}(\omega)$$

Therefore

$$\widehat{f}(\omega) = \widehat{\phi}_s(\omega)\widehat{f_d}(\omega)$$

Now apply the inverse Fourier transform both sides:

$$f(t) = \phi_s * f_d(t) = \phi_s * \sum_{n\in\mathbb{Z}} f(ns)\delta(t - ns) = \sum_{n\in\mathbb{Z}} f(ns)\phi_s(t - ns)$$

□

If the support of $\widehat{f}(\omega)$ is not included in $[-\pi/s, \pi/s]$ then *aliasing* can occur, which is what happens when the supports of $\widehat{f}(\omega - 2k\pi/s)$ overlap for several k. In this case $\widehat{f}(\omega) \neq \widehat{\phi}_s(\omega)\widehat{f_d}(\omega)$, and the sampling theorem (Theorem 17) does not apply and we cannot recover $f(t)$ from $f_d(t)$. Indeed, the Fourier transform of $f_d * \phi_s(t)$ may be very different than the Fourier transform of $f(t)$, in which case $f_d * \phi_s(t)$ will look very different than $f(t)$. See Figure 4 for an illustration.

A *bandlimited* signal is a function f such that $\operatorname{supp} \widehat{f} \subseteq [-R, R]$ for some $R > 0$. The sampling theorem (Theorem 17) proves that such signals can

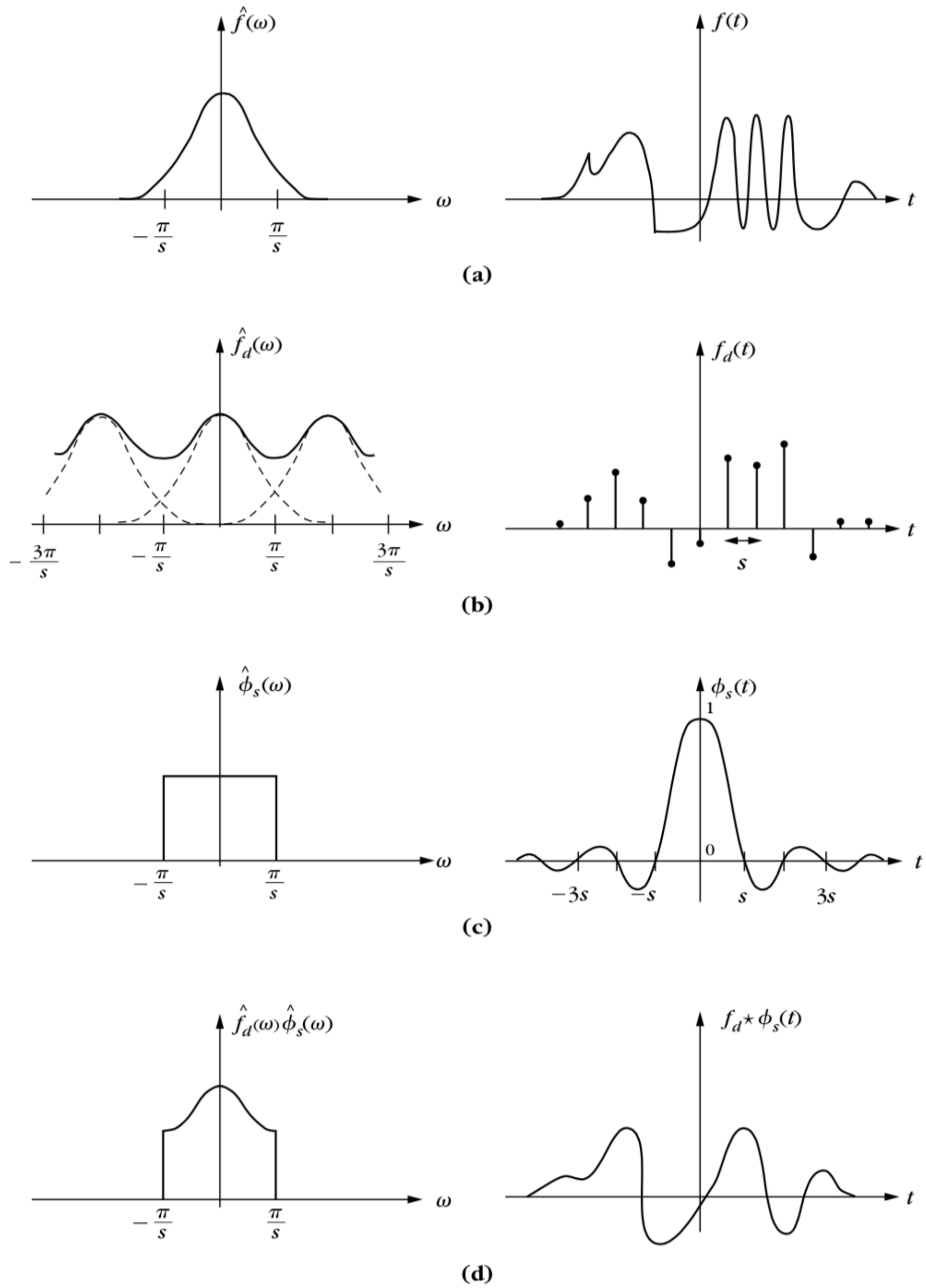


Figure 4: (a) Signal f and its Fourier transform $\widehat{f}$. (b) Aliasing produced by an overlapping of $\widehat{f}(\omega - 2k\pi/s)$ for different k, shown with dashed lines. (c) Low pass filter ϕ_s and its Fourier transform. (d) The filtering $f * \phi_s(t)$ which creates a low frequency signal that is different from f. Notice that non-differentiable singular points are smoothed out, and that the high frequency oscillations on the positive horizontal axis are replaced with a single bump.

be sampled with a discrete set of samples for an appropriate sampling rate $s = \pi/R$. However, by Theorem 11, such signals must necessarily be $\mathbf{C}^\infty$. We will want to be able to process other signals as well. We can do so by first filtering f with some filter h (or a family of filters), which computes $f * h(t)$. If $\operatorname{supp} \widehat{h} \subseteq [-R, R]$ then $f * h(t)$ is bandlimited as well, with the same frequency range. We can thus sample $f * h(t)$ according to Theorem 17. In general we are going to need more than one filter, and each filter will need to be localized in some part of the frequency axis. This will lead us to Gabor filters (windowed Fourier) and wavelets, amongst other filter families.

Exercise 15. Read Section 3.1 of *A Wavelet Tour of Signal Processing.*

3.2 Fourier Series

Section 3.2.2 of A Wavelet Tour of Signal Processing.

Let the sampling rate be $s = 1$ now, which gives samples $\{f(n)\}_{n\in\mathbb{Z}}$ of a signal $f(t)$. Previously we defined

$$f_d(t) = \sum_{n\in\mathbb{Z}} f(n)\delta(t-n)$$

and observed that

$$\widehat{f_d}(\omega) = \sum_{n\in\mathbb{Z}} f(n)e^{-in\omega}$$

This is a *Fourier series*. Clearly $\widehat{f_d}(\omega)$ is 2π periodic, and thus it is uniquely determined by its restriction to $[-\pi, \pi]$. This motivates defining Fourier series on $\boldsymbol{\ell}^1$ and $\boldsymbol{\ell}^2$, which will allow us to represent $f_d(t)$ as a sequence $a = (a[n])_{n\in\mathbb{Z}} \in \boldsymbol{\ell}^p$ with $a[n] = f(n)$. Define

$$\boldsymbol{\ell}^p = \left\{ a = (a[n])_{n\in\mathbb{Z}} : a[n] \in \mathbb{C} \text{ and } \sum_{n\in\mathbb{Z}} |a[n]|^p < \infty \right\}, \quad 0 < p < \infty$$

and

$$\boldsymbol{\ell}^\infty = \left\{ a = (a[n])_{n\in\mathbb{Z}} : a[n] \in \mathbb{C} \text{ and } \sup_{n\in\mathbb{Z}} |a[n]| < \infty \right\}$$

Define the Fourier transform of $a \in \boldsymbol{\ell}^p$ as:

$$\mathcal{F}(a)(\omega) = \widehat{a}(\omega) = \sum_{n\in\mathbb{Z}} a[n]e^{-in\omega}, \quad \omega \in [-\pi, \pi]$$

The Fourier transform of $a \in \boldsymbol{\ell}^1$ is a Fourier series; it is analogous to the Fourier transform of $f \in \mathbf{L}^1(\mathbb{R})$. The spaces $\mathbf{L}^p[-\pi,\pi]$ are defined the same as $\mathbf{L}^p(\mathbb{R})$, except that the domain $\mathbb{R}$ is replaced with $[-\pi,\pi]$, and we normalize the norm so that for $A \in \mathbf{L}^p[-\pi,\pi]$ we have

$$\|A\|_p = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |A(\omega)|^p \, d\omega \right)^{1/p}$$

For $\mathbf{L}^2[-\pi,\pi]$ we have the inner product defined as:

$$\langle A, B \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} A(\omega) B^*(\omega) \, d\omega$$

It is easy to see that $\mathcal{F} : \boldsymbol{\ell}^1 \to \mathbf{L}^\infty[-\pi,\pi]$ and with a little more work (see Theorem 18 below) that $\mathcal{F} : \boldsymbol{\ell}^2 \to \mathbf{L}^2[-\pi,\pi]$, which mirrors our results for the Fourier transform defined on $\mathbf{L}^1(\mathbb{R})$ and $\mathbf{L}^2(\mathbb{R})$. Further developing the parallel story, Theorem 18 below shows that the family of functions $\{e_n\}_{n\in\mathbb{Z}}$ with

$$e_n(\omega) = e^{-in\omega}$$

is an orthonormal basis for $\mathbf{L}^2[-\pi,\pi]$. It follows that $\mathcal{F} : \boldsymbol{\ell}^2 \to \mathbf{L}^2[-\pi,\pi]$ is a bijection, and hence invertible.

Theorem 18. *The family of functions $\{e_n\}_{n\in\mathbb{Z}}$ is a an orthonormal basis for $\mathbf{L}^2[-\pi,\pi]$.*

Proof. The proof that $\{e_n\}_{n\in\mathbb{Z}}$ are orthonormal is by direct calculation of

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\omega} e^{im\omega} \, d\omega = \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases}$$

Now we must show that linear expansions of $\{e_n\}_{n\in\mathbb{Z}}$ are dense in $\mathbf{L}^2[-\pi,\pi]$. This means we need to show the following: Let $A \in \mathbf{L}^2[-\pi,\pi]$, $N > 0$ and define the partial Fourier series of A as:

$$S_N(\omega) = \sum_{n=-N}^{N} \langle A, e_n \rangle e^{-in\omega}, \quad \langle A, e_n \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} A(\omega) e^{in\omega} \, d\omega$$

We need to show that for each $\varepsilon > 0$, there exists $N > 0$ such that

$$\|A - S_N\|_2 < \varepsilon$$

which will mean that $\lim_{N\to\infty} S_N = A$ (in the $\mathbf{L}^2$ sense), which we can write as

$$A(\omega) = \sum_{n\in\mathbb{Z}} \langle A, e_n\rangle e^{-in\omega}$$

To prove this we will use some facts about periodic functions; the proofs of these results can be found in [3]. To start, define a trigonometric polynomial $P(\omega)$ as any function

$$P(\omega) = \sum_{n\in\mathbb{Z}} a_n e^{-in\omega}$$

with only a finite number of coefficients a_n being non-zero. The degree of P is defined as the largest value $|n|$ such that $a_n \neq 0$. One fact we will need is that any function $\phi \in \mathbf{C}[-\pi,\pi]$ with $\phi(-\pi) = \phi(\pi)$ can be uniformly approximated by trigonometric polynomials. That is, for each such ϕ and each $\varepsilon > 0$ there exists P such that

$$|\phi(\omega) - P(\omega)| \leq \varepsilon, \quad \forall -\pi \leq \omega \leq \pi$$

A second fact we will need is that for $A \in \mathbf{L}^2[-\pi,\pi]$ and $\varepsilon > 0$, we can find a function $\phi \in \mathbf{C}[-\pi,\pi]$ with $\phi(-\pi) = \phi(\pi)$ such that

$$\|\phi\|_\infty \leq \|A\|_\infty$$

and

$$\|A - \phi\|_2 \leq \varepsilon^2$$

We also remark that $A \in \mathbf{L}^2[-\pi,\pi]$ implies that $A \in \mathbf{L}^p[-\pi,\pi]$ for any since the length of $[-\pi,\pi]$ is finite.

Now for the remainder of the proof. Since the family $\{e_n\}_{n\in\mathbb{Z}}$ is orthonormal, we must have

$$A - S_N \perp e_n, \quad \forall\, |n| \leq N$$

from which it follows that $A - S_N \perp P_N$, where P_N is any trigonometric polynomial of degree N. Taking $P_N = S_N$, this in turn gives:

$$\|A\|_2^2 = \|A - S_N + S_N\|_2^2 = \|A - S_N\|^2 + \|S_N\|^2$$

We also note that for any trigonometric polynomial P_N, we have

$$\|A - S_N\|_2 \leq \|A - P_N\|_2 \tag{9}$$

with equality only when $P_N = S_N$. Indeed:

$$A - P_N = A - S_N + \underbrace{(S_N - P_N)}_{\widetilde{P}_N}$$

which implies that

$$\|A - P_N\|_2^2 = \|A - S_N\|_2^2 + \|\widetilde{P}_N\|_2^2$$

from which the inequality (9) follows.

We now complete the proof. First consider a $\phi \in \mathbf{C}[-\pi, \pi]$ with $\phi(-\pi) = \phi(\pi)$. Given $\varepsilon > 0$, we can find a trigonometric polynomial P_M with degree M such that

$$|\phi(\omega) - P_M(\omega)| < \varepsilon, \quad \forall -\pi \leq \omega \leq \pi$$

Therefore:

$$\begin{aligned}\|\phi - P_M\|_2^2 &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |\phi(\omega) - P_M(\omega)|^2 \, d\omega \\ &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \varepsilon^2 \, d\omega \leq \varepsilon^2\end{aligned}$$

Thus we have $\|\phi - P_M\| \leq \varepsilon$. But since partial Fourier series are the best approximation (9), we then conclude that

$$\|\phi - S_N(\phi)\|_2 \leq \|\phi - P_M\|_2 \leq \varepsilon, \quad \forall N \geq M$$

Now let us return to the case of general $A \in \mathbf{L}^2[-\pi, \pi]$. For $\varepsilon > 0$, approximate A with a $\phi \in \mathbf{C}[-\pi, \pi]$, $\phi(-\pi) = \phi(\pi)$, such that $\|A - \phi\|_2 \leq \varepsilon$. Approximate ϕ with a trigonometric polynomial P_M, as before, to obtain:

$$\|A - P_M\|_2 \leq \|A - \phi\|_2 + \|\phi - P_M\|_2 \leq 2\varepsilon$$

Now again use the best approximation inequality (9) to conclude that:

$$\|A - S_N\|_2 \leq 2\varepsilon, \quad \forall N \geq M$$

□

Theorem 18 proves that any periodic function $A \in \mathbf{L}^2[-\pi, \pi]$ can be written as

$$A(\omega) = \sum_{n \in \mathbb{Z}} c_n e^{-in\omega} \tag{10}$$

with

$$c_n = \langle A, e_n \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} A(\omega) e^{in\omega} \, d\omega$$

In particular, if we start with $a \in \boldsymbol{\ell}^2$ and compute its Fourier series:

$$\widehat{a}(\omega) = \sum_{n \in \mathbb{Z}} a[n] e^{-in\omega}$$

then we must have

$$a[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \widehat{a}(\omega) e^{in\omega} \, d\omega$$

which is a type of Fourier inversion formula for Fourier series. Additionally, if $a[n] = f(n)$ for some signal f, then from the beginning of this section we see that $\widehat{f_d}(\omega) = \widehat{a}(\omega)$ and

$$f(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \widehat{f_d}(\omega) e^{in\omega} \, d\omega$$

which gives Fourier inversion for the samples $f(n)$ from $\widehat{f_d}(\omega)$. We have a similar distributional version, which requires defining one more distributional Fourier transform. Recall that for $\delta_\tau(t) = \delta(t - \tau)$, we defined the Fourier transform as $\widehat{\delta_\tau}(\omega) = e^{-i\omega\tau}$. Given this, if we set $e_\xi(t) = e^{i\eta t}$, it makes sense to define its Fourier transform as $\widehat{e_\xi}(\omega) = 2\pi\delta(\omega - \eta)$. Indeed, e_ξ is a perfect harmonic vibrating at frequency ξ. Using this fact, if we compute the inverse Fourier transform of $\widehat{f_d}(\omega)$ in the distributional sense, it is clear we will get back $f_d(t)$. More generally, if we start $A \in \mathbf{L}^2[-\pi, \pi]$ and write it as in (10), and compute the distributional inverse Fourier transform of A, we will get

$$\mathcal{F}^{-1}(A)(t) = \sum_{n \in \mathbb{Z}} \langle A, e_n \rangle \delta(t - n)$$

Of course this is equivalent to computing the Fourier series inversion of A and getting a sequence $(\langle A, e_n \rangle)_{n \in \mathbb{Z}} \in \boldsymbol{\ell}^2(\mathbb{R})$, but it is sometimes convenient to use one over the other.

We also have the following version of the Plancheral formula:

$$\|a\|^2 = \sum_{n \in \mathbb{Z}} |a[n]|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\widehat{a}(\omega)|^2 \, d\omega = \|\widehat{a}\|^2$$

Define the convolution of $a, b \in \boldsymbol{\ell}^1$ as:

$$a * b[n] = \sum_{m \in \mathbb{Z}} a[m]b[n-m]$$

We have a convolution theorem for $\boldsymbol{\ell}^1$ sequences as well:

Theorem 19. *Let $a, b \in \boldsymbol{\ell}^1$. Then $a * b \in \boldsymbol{\ell}^1$ and*

$$\widehat{a * b}(\omega) = \widehat{a}(\omega)\widehat{b}(\omega)$$

Exercise 16. Read Section 3.2 of *A Wavelet Tour of Signal Processing.*

Exercise 17. A rectifier computes $g(t) = |f(t)|$ for recovering the envelope of modulated signals.

(a) Show that if $f(t) = h(t)\sin(\omega_0 t)$ with $h \in \mathbf{L}^1(\mathbb{R})$, $h(t) \geq 0$ and $\omega_0 > 0$, then $g(t) = |f(t)|$ satisfies

$$\widehat{g}(\omega) = \frac{2}{\pi} \sum_{n=-\infty}^{+\infty} \frac{\widehat{h}(\omega - 2n\omega_0)}{4n^2 - 1}$$

Hint: Let $A(t)$ be a 2π periodic function. By Theorem 18 we can write

$$A(t) = \sum_{n \in \mathbb{Z}} c_n e^{int}$$

with

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} A(t) e^{-int} \, dt$$

Compute now the Fourier transform of A in the distributional sense. Do this for $A(t) = |\sin(t)|$ and apply the convolution theorem to $g(t)$ (along with the rescaling property of the Fourier transform).

(b) Suppose that $\widehat{h}(\omega) = 0$ for $|\omega| \geq \omega_0$. Find ϕ such that $h(t) = \phi * g(t)$.

Exercise 18. An *interpolation function* $f(t)$ satisfies $f(n) = \delta(n)$ for any $n \in \mathbb{Z}$.

(a) Prove that

$$\sum_{n \in \mathbb{Z}} \widehat{f}(\omega + 2n\pi) = 1 \quad \Longleftrightarrow \quad f \text{ is an interpolation function}$$

(b) Suppose that

$$f(t) = \sum_{n \in \mathbb{Z}} a[n]\theta(t-n), \quad a \in \boldsymbol{\ell}^1, \ \theta \in \mathbf{L}^2(\mathbb{R})$$

Find $\widehat{a}(\omega)$ as a function of $\widehat{\theta}(\omega)$ so that $f(t)$ is an interpolation function. Relate $\widehat{f}(\omega)$ to $\widehat{\theta}(\omega)$, and give a sufficient condition on $\widehat{\theta}$ to guarantee that $f \in \mathbf{L}^2(\mathbb{R})$.

Exercise 19. Let $g \in \boldsymbol{\ell}^1$ and set $h[n] = (-1)^n g[n]$. Relate $\widehat{h}(\omega)$ to $\widehat{g}(\omega)$. If g is a low pass filter (meaning that $\widehat{g}(\omega)$ is concentrated around 0), then what kind of filter is h? (i.e., where is its support concentrated?)

Exercise 20. Let $b \in \boldsymbol{\ell}^1$. A *decimation* of b computes a signal $a \in \boldsymbol{\ell}^1$ with $a[n] = b[Mn]$ for $M > 1$ ($M \in \mathbb{Z}$).

(a) Show that

$$\widehat{a}(\omega) = \frac{1}{M} \sum_{k=0}^{M-1} \widehat{b}(M^{-1}(\omega - 2k\pi))$$

(b) Give a sufficient condition on $\widehat{b}(\omega)$ to recover b from a and give the interpolation formula that recovers $b[n]$ from a.

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.

[6] John J. Benedetto and Matthew Dellatorre. Uncertainty principles and weighted norm inequalities. *Contemporary Mathematics*, 693:55–78, 2017.