Chapter 1

The Fourier Transform

1.1 Fourier transforms as integrals

There are several ways to define the Fourier transform of a function $f : \mathbb{R} \to \mathbb{C}$. In this section, we define it using an integral representation and state some basic uniqueness and inversion properties, without proof. Thereafter, we will consider the transform as being defined as a suitable limit of Fourier series, and will prove the results stated here.

Definition 1 *Let $f : \mathbb{R} \to \mathbb{R}$. The Fourier transform of $f \in L^1(\mathbb{R})$, denoted by $\mathcal{F}[f](.)$, is given by the integral:*

$$\mathcal{F}[f](x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) \exp(-ixt) dt$$

for $x \in \mathbb{R}$ for which the integral exists. *

We have the **Dirichlet condition** for inversion of Fourier integrals.

Theorem 1 *Let $f : \mathbb{R} \to \mathbb{R}$. Suppose that (1) $\int_{-\infty}^{\infty} |f| \, dt$ converges and (2) in any finite interval, f, f' are piecewise continuous with at most finitely many maxima/minima/discontinuities. Let $F = \mathcal{F}[f]$. Then if f is continuous at $t \in \mathbb{R}$, we have*

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(x) \exp(itx) dx.$$

*This definition also makes sense for complex valued f but we stick here to real valued f

Moreover, if f is discontinuous at $t \in \mathbb{R}$ and $f(t+0)$ and $f(t-0)$ denote the right and left limits of f at t, then

$$\frac{1}{2}[f(t+0)+f(t-0)] = \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\infty} F(x)\exp(itx)dx.$$

From the above, we deduce a uniqueness result:

Theorem 2 *Let $f, g : \mathbb{R} \to \mathbb{R}$ be continuous, f', g' piecewise continuous. If*

$$\mathcal{F}[f](x) = \mathcal{F}[g](x),\ \forall x$$

then

$$f(t) = g(t),\ \forall t.$$

Proof: We have from inversion, easily that

$$\begin{aligned} f(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{F}[f](x)\exp(itx)dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathcal{F}[g](x)\exp(itx)dx \\ &= g(t). \end{aligned}$$

□

Example 1 *Find the Fourier transform of $f(t) = \exp(-|t|)$ and hence using inversion, deduce that $\int_0^\infty \frac{dx}{1+x^2} = \frac{\pi}{2}$ and $\int_0^\infty \frac{x\sin(xt)}{1+x^2}dx = \frac{\pi\exp(-t)}{2}$, $t > 0$.*

Solution We write

$$\begin{aligned} F(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)\exp(-ixt)dt \\ &= \frac{1}{\sqrt{2\pi}} \left[\int_{-\infty}^{0} \exp(t(1-ix))dt + \int_0^\infty \exp(-t(1+ix)) \right] \\ &= \sqrt{\frac{2}{\pi}} \frac{1}{1+x^2}. \end{aligned}$$

Now by the inversion formula,

$$\begin{aligned} \exp(-|t|) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(x)\exp(ixt)dx \\ &= \frac{1}{\pi} \left[\int_0^\infty \frac{\exp(ixt)+\exp(-ixt)}{1+x^2} dt \right] \\ &= \frac{2}{\pi} \int_0^\infty \frac{\cos(xt)}{1+x^2} dx. \end{aligned}$$

Now this formula holds at $t = 0$, so substituting $t = 0$ into the above gives the first required identity. Differentiating with respect to t as we may for $t > 0$, gives the second required identity. □.

Proceeding in a similar way as the above example, we can easily show that

$$\mathcal{F}[\exp(-\frac{1}{2}t^2)](x) = \exp(-\frac{1}{2}x^2),\ x \in \mathbb{R}.$$

We will discuss this example in more detail later in this chapter.

We will also show that we can reinterpret Definition 1 to obtain the Fourier transform of any complex valued $f \in L^2(\mathbb{R})$, and that the Fourier transform is unitary on this space:

Theorem 3 *If $f, g \in L^2(\mathbb{R})$ then $\mathcal{F}[f], \mathcal{F}[g] \in L^2(\mathbb{R})$ and*

$$\int_{-\infty}^{\infty} f(t)\overline{g}(t)\ dt = \int_{-\infty}^{\infty} \mathcal{F}[f](x)\overline{\mathcal{F}[g](x)}\ dx.$$

This is a result of fundamental importance for applications in signal processing.

1.2 The transform as a limit of Fourier series

We start by constructing the Fourier series (complex form) for functions on an interval $[-\pi L, \pi L]$. The ON basis functions are

$$e_n(t) = \frac{1}{\sqrt{2\pi L}} e^{\frac{int}{L}}, \qquad n = 0, \pm 1, \cdots,$$

and a sufficiently smooth function f of period $2\pi L$ can be expanded as

$$f(t) = \sum_{n=-\infty}^{\infty} \left(\frac{1}{2\pi L} \int_{-\pi L}^{\pi L} f(x) e^{-\frac{inx}{L}} dx \right) e^{\frac{int}{L}}.$$

For purposes of motivation let us abandon periodicity and think of the functions f as differentiable everywhere, vanishing at $t = \pm\pi L$ and identically zero outside $[-\pi L, \pi L]$. We rewrite this as

$$f(t) = \sum_{n=-\infty}^{\infty} e^{\frac{int}{L}} \frac{1}{2\pi L} \hat{f}(\frac{n}{L})$$

which looks like a Riemann sum approximation to the integral

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\lambda) e^{i\lambda t} d\lambda \tag{1.2.1}$$

to which it would converge as $L \to \infty$. (Indeed, we are partitioning the λ interval $[-L, L]$ into $2L$ subintervals, each with partition width $1/L$.) Here,

$$\hat{f}(\lambda) = \int_{-\infty}^{\infty} f(t) e^{-i\lambda t} dt. \tag{1.2.2}$$

Similarly the Parseval formula for f on $[-\pi L, \pi L]$,

$$\int_{-\pi L}^{\pi L} |f(t)|^2 dt = \sum_{n=-\infty}^{\infty} \frac{1}{2\pi L} |\hat{f}(\frac{n}{L})|^2$$

goes in the limit as $L \to \infty$ to the *Plancherel identity*

$$2\pi \int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |\hat{f}(\lambda)|^2 d\lambda. \tag{1.2.3}$$

Expression (1.2.2) is called the *Fourier integral* or *Fourier transform* of f. Expression (1.2.1) is called the *inverse Fourier integral* for f. The Plancherel identity suggests that the Fourier transform is a one-to-one norm preserving map of the Hilbert space $L^2[-\infty, \infty]$ onto itself (or to another copy of itself). We shall show that this is the case. Furthermore we shall show that the pointwise convergence properties of the inverse Fourier transform are somewhat similar to those of the Fourier series. Although we could make a rigorous justification of the the steps in the Riemann sum approximation above, we will follow a different course and treat the convergence in the mean and pointwise convergence issues separately.

A second notation that we shall use is

$$\mathcal{F}[f](\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\lambda t} dt = \frac{1}{\sqrt{2\pi}} \hat{f}(\lambda) \tag{1.2.4}$$

$$\mathcal{F}^*[g](t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\lambda) e^{i\lambda t} d\lambda \tag{1.2.5}$$

Note that, formally, $\mathcal{F}^*[\hat{f}](t) = \sqrt{2\pi} f(t)$. The first notation is used more often in the engineering literature. The second notation makes clear that $\mathcal{F}$ and $\mathcal{F}^*$ are linear operators mapping $L^2[-\infty, \infty]$ onto itself in one view, and $\mathcal{F}$ mapping the *signal space* onto the *frequency space* with $\mathcal{F}^*$ mapping the frequency space onto the signal space in the other view. In this notation the Plancherel theorem takes the more symmetric form

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |\mathcal{F}[f](\lambda)|^2 d\lambda.$$

Examples:

1. The box function (or rectangular wave)

$$\Pi(t) = \begin{cases} 1 & \text{if } -\pi < t < \pi \\ \frac{1}{2} & \text{if } t = \pm\pi \\ 0 & \text{otherwise.} \end{cases} \tag{1.2.6}$$

Then, since $\Pi(t)$ is an even function and $e^{-i\lambda t} = \cos(\lambda t) + i\sin(\lambda t)$, we have

$$\hat{\Pi}(\lambda) = \sqrt{2\pi}\mathcal{F}[\Pi](\lambda) = \int_{-\infty}^{\infty} \Pi(t)e^{-i\lambda t}dt = \int_{-\infty}^{\infty} \Pi(t)\cos(\lambda t)dt$$

$$= \int_{-\pi}^{\pi} \cos(\lambda t)dt = \frac{2\sin(\pi\lambda)}{\lambda} = 2\pi \text{ sinc } \lambda.$$

Thus sinc λ is the Fourier transform of the box function. The inverse Fourier transform is

$$\int_{-\infty}^{\infty} \text{sinc}(\lambda)e^{i\lambda t}d\lambda = \Pi(t), \tag{1.2.7}$$

as follows from (**??**). Furthermore, we have

$$\int_{-\infty}^{\infty} |\Pi(t)|^2 dt = 2\pi$$

and

$$\int_{-\infty}^{\infty} |\text{ sinc }(\lambda)|^2 d\lambda = 1$$

from (**??**), so the Plancherel equality is verified in this case. Note that the inverse Fourier transform converged to the midpoint of the discontinuity, just as for Fourier series.

2. A truncated cosine wave.

$$f(t) = \begin{cases} \cos 3t & \text{if } -\pi < t < \pi \\ -\frac{1}{2} & \text{if } t = \pm\pi \\ 0 & \text{otherwise.} \end{cases}$$

Then, since the cosine is an even function, we have

$$\hat{f}(\lambda) = \sqrt{2\pi}\mathcal{F}[f](\lambda) = \int_{-\infty}^{\infty} f(t)e^{-i\lambda t}dt = \int_{-\pi}^{\pi} \cos(3t)\cos(\lambda t)dt$$

$$= \frac{2\lambda\sin(\lambda\pi)}{9-\lambda^2}.$$

3. A truncated sine wave.

$$f(t) = \begin{cases} \sin 3t & \text{if } -\pi \le t \le \pi \\ 0 & \text{otherwise.} \end{cases}$$

Since the sine is an odd function, we have

$$\hat{f}(\lambda) = \sqrt{2\pi}\mathcal{F}[f](\lambda) = \int_{-\infty}^{\infty} f(t)e^{-i\lambda t}dt = -i\int_{-\pi}^{\pi} \sin(3t)\sin(\lambda t)dt$$

$$= \frac{-6i\sin(\lambda\pi)}{9-\lambda^2}.$$

4. A triangular wave.

$$f(t) = \begin{cases} 1+t & \text{if } -1 \le t \le 0 \\ -1 & \text{if } 0 \le t \le 1 \\ 0 & \text{otherwise.} \end{cases} \tag{1.2.8}$$

Then, since f is an even function, we have

$$\hat{f}(\lambda) = \sqrt{2\pi}\mathcal{F}[f](\lambda) = \int_{-\infty}^{\infty} f(t)e^{-i\lambda t}dt = 2\int_{0}^{1} (1-t)\cos(\lambda t)dt$$

$$= \frac{2-2\cos\lambda}{\lambda^2}.$$

NOTE: The Fourier transforms of the discontinuous functions above decay as $\frac{1}{\lambda}$ for $|\lambda| \to \infty$ whereas the Fourier transforms of the continuous functions decay as $\frac{1}{\lambda^2}$. The coefficients in the Fourier series of the analogous functions decay as $\frac{1}{n}$, $\frac{1}{n^2}$, respectively, as $|n| \to \infty$.

1.2.1 Properties of the Fourier transform

Recall that

$$\mathcal{F}[f](\lambda) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{-i\lambda t}dt = \frac{1}{\sqrt{2\pi}}\hat{f}(\lambda)$$

$$\mathcal{F}^*[g](t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} g(\lambda)e^{i\lambda t}d\lambda$$

We list some properties of the Fourier transform that will enable us to build a repertoire of transforms from a few basic examples. Suppose that f, g belong to $L^1[-\infty,\infty]$, i.e., $\int_{-\infty}^{\infty}|f(t)|dt < \infty$ with a similar statement for g. We can state the following (whose straightforward proofs are left to the reader):

1. $\mathcal{F}$ and $\mathcal{F}^*$ are linear operators. For $a, b \in C$ we have

$$\mathcal{F}[af + bg] = a\mathcal{F}[f] + b\mathcal{F}[g], \quad \mathcal{F}^*[af + bg] = a\mathcal{F}^*[f] + b\mathcal{F}^*[g].$$

2. Suppose $t^n f(t) \in L^1[-\infty, \infty]$ for some positive integer n. Then

$$\mathcal{F}[t^n f(t)](\lambda) = i^n \frac{d^n}{d\lambda^n}\{\mathcal{F}[f](\lambda)\}.$$

3. Suppose $\lambda^n f(\lambda) \in L^1[-\infty, \infty]$ for some positive integer n. Then

$$\mathcal{F}^*[\lambda^n f(\lambda)](t) = i^n \frac{d^n}{dt^n}\{\mathcal{F}^*[f](t)\}.$$

4. Suppose the nth derivative $f^{(n)}(t) \in L^1[-\infty, \infty]$ and piecewise continuous for some positive integer n, and f and the lower derivatives are all continuous in $(-\infty, \infty)$. Then

$$\mathcal{F}[f^{(n)}](\lambda) = (i\lambda)^n \mathcal{F}[f](\lambda)\}.$$

5. Suppose nth derivative $f^{(n)}(\lambda) \in L^1[-\infty, \infty]$ for some positive integer n and piecewise continuous for some positive integer n, and f and the lower derivatives are all continuous in $(-\infty, \infty)$. Then

$$\mathcal{F}^*[f^{(n)}](t) = (-it)^n \mathcal{F}^*[f](t).$$

6. The Fourier transform of a translation by real number a is given by

$$\mathcal{F}[f(t-a)](\lambda) = e^{-i\lambda a}\mathcal{F}[f](\lambda).$$

7. The Fourier transform of a scaling by positive number b is given by

$$\mathcal{F}[f(bt)](\lambda) = \frac{1}{b}\mathcal{F}[f](\frac{\lambda}{b}).$$

8. The Fourier transform of a translated and scaled function is given by

$$\mathcal{F}[f(bt-a)](\lambda) = \frac{1}{b}e^{-i\lambda a/b}\mathcal{F}[f](\frac{\lambda}{b}).$$

Examples

- We want to compute the Fourier transform of the rectangular box function with support on $[c, d]$:

$$R(t) = \begin{cases} 1 & \text{if } c < t < d \\ \frac{1}{2} & \text{if } t = c, d \\ 0 & \text{otherwise.} \end{cases}$$

  Recall that the box function

$$\Pi(t) = \begin{cases} 1 & \text{if } -\pi < t < \pi \\ \frac{1}{2} & \text{if } t = \pm\pi \\ 0 & \text{otherwise.} \end{cases}$$

  has the Fourier transform $\hat{\Pi}(\lambda) = 2\pi \text{ sinc } \lambda$. but we can obtain R from Π by first translating $t \longrightarrow s = t - \frac{(c+d)}{2}$ and then rescaling $s \longrightarrow \frac{2\pi}{d-c}s$:

$$R(t) = \Pi(\frac{2\pi}{d-c}t - \pi\frac{c+d}{d-c}).$$

$$\hat{R}(\lambda) = \frac{4\pi^2}{d-c}e^{i\pi\lambda(c+d)/(d-c)}\text{sinc}(\frac{2\pi\lambda}{d-c}). \qquad (1.2.9)$$

  Furthermore, from (**??**) we can check that the inverse Fourier transform of $\hat{R}$ is R, i.e., $\mathcal{F}^*(\mathcal{F})R(t) = R(t)$.

- Consider the truncated sine wave

$$f(t) = \begin{cases} \sin 3t & \text{if } -\pi \le t \le \pi \\ 0 & \text{otherwise} \end{cases}$$

  with

$$\hat{f}(\lambda) = \frac{-6i\sin(\lambda\pi)}{9 - \lambda^2}.$$

  Note that the derivative f' of $f(t)$ is just $3g(t)$ (except at 2 points) where $g(t)$ is the truncated cosine wave

$$g(t) = \begin{cases} \cos 3t & \text{if } -\pi < t < \pi \\ -\frac{1}{2} & \text{if } t = \pm\pi \\ 0 & \text{otherwise.} \end{cases}$$

  We have computed

$$\hat{g}(\lambda) = \frac{2\lambda\sin(\lambda\pi)}{9 - \lambda^2}.$$

  so $3\hat{g}(\lambda) = (i\lambda)\hat{f}(\lambda)$, as predicted.

- Reversing the example above we can differentiate the truncated cosine wave to get the truncated sine wave. The prediction for the Fourier transform doesn't work! Why not?

1.2.2 Fourier transform of a convolution

The following property of the Fourier transform is of particular importance in signal processing. Suppose f, g belong to $L^1[-\infty, \infty]$.

Definition 2 *The convolution of f and g is the function $f * g$ defined by*

$$(f * g)(t) = \int_{-\infty}^{\infty} f(t-x)g(x)dx.$$

Note also that $(f * g)(t) = \int_{-\infty}^{\infty} f(x)g(t-x)dx$, as can be shown by a change of variable.

Lemma 1 $f * g \in L^1[-\infty, \infty]$ *and*

$$\int_{-\infty}^{\infty} |f * g(t)|dt = \int_{-\infty}^{\infty} |f(x)|dx \int_{-\infty}^{\infty} |g(t)|dt.$$

Sketch of proof:

$$\int_{-\infty}^{\infty} |f * g(t)|dt = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} |f(x)g(t-x)|dx \right) dt$$

$$= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} |g(t-x)|dt \right) |f(x)|dx = \int_{-\infty}^{\infty} |g(t)|dt \int_{-\infty}^{\infty} |f(x)|dx.$$

□

Theorem 4 *Let $h = f * g$. Then*

$$\hat{h}(\lambda) = \hat{f}(\lambda)\hat{g}(\lambda).$$

Sketch of proof:

$$\hat{h}(\lambda) = \int_{-\infty}^{\infty} f * g(t) e^{-i\lambda t} dt = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x)g(t-x)dx \right) e^{-i\lambda t} dt$$

$$= \int_{-\infty}^{\infty} f(x)e^{-i\lambda x} \left(\int_{-\infty}^{\infty} g(t-x)e^{-i\lambda(t-x)}dt \right) dx = \int_{-\infty}^{\infty} f(x)e^{-i\lambda x}dx \ \hat{g}(\lambda)$$

$$= \hat{f}(\lambda)\hat{g}(\lambda).$$

□

1.3 L^2 convergence of the Fourier transform

In this book our primary interest is in Fourier transforms of functions in the Hilbert space $L^2[-\infty,\infty]$. However, the formal definition of the Fourier integral transform,

$$\mathcal{F}[f](\lambda)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}f(t)e^{-i\lambda t}dt \tag{1.3.10}$$

doesn't make sense for a general $f\in L^2[-\infty,\infty]$. If $f\in L^1[-\infty,\infty]$ then f is absolutely integrable and the integral (1.3.10) converges. However, there are square integrable functions that are not integrable. (Example: $f(t)=\frac{1}{1+|t|}$.) How do we define the transform for such functions?

We will proceed by defining $\mathcal{F}$ on a dense subspace of $f\in L^2[-\infty,\infty]$ where the integral makes sense and then take Cauchy sequences of functions in the subspace to define $\mathcal{F}$ on the closure. Since $\mathcal{F}$ preserves inner product, as we shall show, this simple procedure will be effective.

First some comments on integrals of L^2 functions. If $f,g\in L^2[-\infty,\infty]$ then the integral $(f,g)=\int_{-\infty}^{\infty}f(t)\overline{g}(t)dt$ necessarily exists, whereas the integral (1.3.10) may not, because the exponential $e^{-i\lambda t}$ is not an element of L^2. However, the integral of $f\in L^2$ over any finite interval, say $[-N,N]$ does exist. Indeed for N a positive integer, let $\chi_{[-N,N]}$ be the indicator function for that interval:

$$\chi_{[-N,N]}(t)=\begin{cases}1 & \text{if } -N\le t\le N\\ 0 & \text{otherwise.}\end{cases} \tag{1.3.11}$$

Then $\chi_{[-N,N]}\in L^2[-\infty,\infty]$ so $\int_{-N}^{N}f(t)dt$ exists because

$$\int_{-N}^{N}|f(t)|dt=|(|f|,\chi_{[-N,N]})|\le ||f||_{L^2}||\chi_{[-N,N]}||_{L^2}=||f||_{L^2}\sqrt{2N}<\infty$$

Now the space of step functions is dense in $L^2[-\infty,\infty]$, so we can find a convergent sequence of step functions $\{s_n\}$ such that $\lim_{n\to\infty}||f-s_n||_{L^2}=0$. Note that the sequence of functions $\{f_N=f\chi_{[-N,N]}\}$ converges to f pointwise as $N\to\infty$ and each $f_N\in L^2\cap L^1$.

Lemma 2 *$\{f_N\}$ is a Cauchy sequence in the norm of $L^2[-\infty,\infty]$ and $\lim_{n\to\infty}||f-f_n||_{L^2}=0$.*

Proof: Given $\epsilon>0$ there is step function s_M such that $||f-s_M||<\frac{\epsilon}{2}$. Choose N so large that the support of s_M is contained in $[-N,N]$, i.e.,

$s_M(t)\chi_{[-N,N]}(t) = s_M(t)$ for all t. Then $||s_M - f_N||^2 = \int_{-N}^{N} |s_M(t) - f(t)|^2 dt \le \int_{-\infty}^{\infty} |s_M(t) - f(t)|^2 dt = ||s_M - f||^2$, so

$$||f - f_N|| - ||(f - s_M) + (s_M - f_N)|| \le ||f - s_M|| + ||s_M - f_N|| \le 2||f - s_M|| < \epsilon.$$

□

Here we will study the linear mapping $\mathcal{F} : L^2[-\infty, \infty] \to \hat{L}^2[-\infty, \infty]$ from the signal space to the frequency space. We will show that the mapping is *unitary*, i.e., it preserves the inner product and is 1-1 and onto. Moreover, the map $\mathcal{F}^* : \hat{L}^2[-\infty, \infty] \to L^2[-\infty, \infty]$ is also a unitary mapping and is the inverse of $\mathcal{F}$:

$$\mathcal{F}^*\mathcal{F} = I_{L^2}, \qquad \mathcal{F}\mathcal{F}^* = I_{\hat{L}^2}$$

where $I_{L^2}, I_{\hat{L}^2}$ are the identity operators on L^2 and $\hat{L}^2$, respectively. We know that the space of step functions is dense in L^2. Hence to show that $\mathcal{F}$ preserves inner product, it is enough to verify this fact for step functions and then go to the limit. Once we have done this, we can define $\mathcal{F}f$ for any $f \in L^2[-\infty, \infty]$. Indeed, if $\{s_n\}$ is a Cauchy sequence of step functions such that $\lim_{n\to\infty} ||f - s_n||_{L^2} = 0$, then $\{\mathcal{F}s_n\}$ is also a Cauchy sequence (indeed, $||s_n - s_m|| = ||\mathcal{F}s_n - \mathcal{F}s_m||$) so we can define $\mathcal{F}f$ by $\mathcal{F}f = \lim_{n\to\infty} \mathcal{F}s_n$. The standard methods of Section 1.3 show that $\mathcal{F}f$ is uniquely defined by this construction. Now the truncated functions f_N have Fourier transforms given by the convergent integrals

$$\mathcal{F}[f_N](\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-N}^{N} f(t) e^{-i\lambda t} dt$$

and $\lim_{N\to\infty} ||f - f_N||_{L^2} = 0$. Since $\mathcal{F}$ preserves inner product we have $||\mathcal{F}f - \mathcal{F}f_N||_{L^2} = ||\mathcal{F}(f - f_N)||_{L^2} = ||f - f_N||_{L^2}$, so $\lim_{N\to\infty} ||\mathcal{F}f - \mathcal{F}f_N||_{L^2} = 0$. We write

$$\mathcal{F}[f](\lambda) = \text{l.i.m.}_{N\to\infty} \mathcal{F}[f_N](\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-N}^{N} f(t) e^{-i\lambda t} dt$$

where 'l.i.m.' indicates that the convergence is in the mean (Hilbert space) sense, rather than pointwise.

We have already shown that the Fourier transform of the rectangular box function with support on $[c, d]$:

$$R_{c,d}(t) = \begin{cases} 1 & \text{if } c < t < d \\ \frac{1}{2} & \text{if } t = c, d \\ 0 & \text{otherwise.} \end{cases}$$

is

$$\mathcal{F}[R_{c,d}](\lambda) = \frac{4\pi^2}{\sqrt{2\pi}(d - c)} e^{i\pi\lambda(c+d)/(d-c)} \text{sinc}(\frac{2\pi\lambda}{d - c}).$$

and that $\mathcal{F}^*(\mathcal{F})R_{c,d}(t) = R_{c,d}(t)$. (Since here we are concerned only with convergence in the mean the value of a step function at a particular point is immaterial. Hence for this discussion we can ignore such niceties as the values of step functions at the points of their jump discontinuities.)

Lemma 3

$$(R_{a,b}, R_{c,d})_{L^2} = (\mathcal{F}R_{a,b}, \mathcal{F}R_{c,d})_{\hat{L}^2}$$

for all real numbers $a \leq b$ and $c \leq d$.

Proof:

$$(\mathcal{F}R_{a,b}, \mathcal{F}R_{c,d})_{\hat{L}^2} = \int_{-\infty}^{\infty} \mathcal{F}[R_{a,b}](\lambda)\overline{\mathcal{F}}[R_{c,d}](\lambda)d\lambda$$

$$= \lim_{N\to\infty} \int_{-N}^{N} \left(\mathcal{F}[R_{a,b}](\lambda) \int_c^d \frac{e^{i\lambda t}}{\sqrt{2\pi}}dt\right) d\lambda$$

$$= \lim_{N\to\infty} \int_c^d \left(\int_{-N}^{N} \mathcal{F}[R_{a,b}](\lambda)\frac{e^{i\lambda t}}{\sqrt{2\pi}}d\lambda\right) dt.$$

Now the inside integral is converging to $R_{a,b}$ as $N \to \infty$ in both the pointwise and L^2 sense, as we have shown. Thus

$$(\mathcal{F}R_{a,b}, \mathcal{F}R_{c,d})_{\hat{L}^2} = \int_c^d R_{a,b}dt = (R_{a,b}, R_{c,d})_{L^2}.$$

□

Since any step functions u, v are finite linear combination of indicator functions R_{a_j,b_j} with complex coefficients, $u = \sum_j \alpha_j R_{a_j,b_j}$, $v = \sum_k \beta_k R_{c_k,d_k}$ we have

$$(\mathcal{F}u, \mathcal{F}v)_{\hat{L}^2} = \sum_{j,k} \alpha_j\overline{\beta}_k(\mathcal{F}R_{a_j,b_j}, \mathcal{F}R_{c_k,d_k})_{\hat{L}^2}$$

$$= \sum_{j,k} \alpha_j\overline{\beta}_k(R_{a_j,b_j}, R_{c_k,d_k})_{L^2} = (u, v)_{L^2}.$$

Thus $\mathcal{F}$ preserves inner product on step functions, and by taking Cauchy sequences of step functions, we have the

Theorem 5 *(Plancherel Formula) Let $f, g \in L^2[-\infty, \infty]$. Then*

$$(f, g)_{L^2} = (\mathcal{F}f, \mathcal{F}g)_{\hat{L}^2}, \qquad ||f||^2_{L^2} = ||\mathcal{F}f||^2_{\hat{L}^2}$$

In the engineering notation this reads

$$2\pi \int_{-\infty}^{\infty} f(t)\overline{g}(t)dt = \int_{-\infty}^{\infty} \hat{f}(\lambda)\overline{\hat{g}}(\lambda)d\lambda.$$

Theorem 6 *The map* $\mathcal{F}^* : \hat{L}^2[-\infty,\infty] \to L^2[-\infty,\infty]$ *has the following properties:*

1. *It preserves inner product, i.e.,*

$$(\mathcal{F}^*\hat{f}, \mathcal{F}^*\hat{g})_{L^2} = (\hat{f}, \hat{g})_{\hat{L}^2}$$

   for all $\hat{f}, \hat{g} \in \hat{L}^2[-\infty,\infty]$.

2. $\mathcal{F}^*$ *is the adjoint operator to* $\mathcal{F} : L^2[-\infty,\infty] \to \hat{L}^2[-\infty,\infty]$, *i.e.,*

$$(\mathcal{F}f, \hat{g})_{\hat{L}^2} = (f, \mathcal{F}^*\hat{g})_{L^2},$$

   for all $f \in L^2[-\infty,\infty]$, $\hat{g} \in \hat{L}^2[-\infty,\infty]$.

Proof:

1. This follows immediately from the facts that $\mathcal{F}$ preserves inner product and $\overline{\mathcal{F}}[\overline{f}](\lambda) = \mathcal{F}^*[f](\lambda)$.

2.
$$(\mathcal{F}R_{a,b}, R_{c,d})_{\hat{L}^2} = (R_{a,b}, \mathcal{F}^*R_{c,d})_{L^2}$$

   as can be seen by an interchange in the order of integration. Then using the linearity of $\mathcal{F}$ and $\mathcal{F}^*$ we see that

$$(\mathcal{F}u, v)_{\hat{L}^2} = (u, \mathcal{F}^*v)_{L^2},$$

   for all step functions u, v. Since the space of step functions is dense in $\hat{L}^2[-\infty,\infty]$ and in $L^2[-\infty,\infty]$

□

Theorem 7

1. *The Fourier transform* $\mathcal{F} : L^2[-\infty,\infty] \to \hat{L}^2[-\infty,\infty]$ *is a unitary transformation, i.e., it preserves the inner product and is 1-1 and onto.*

2. *The adjoint map* $\mathcal{F}^* : \hat{L}^2[-\infty,\infty] \to L^2[-\infty,\infty]$ *is also a unitary mapping.*

3. $\mathcal{F}^*$ *is the inverse operator to* $\mathcal{F}$*:*

$$\mathcal{F}^*\mathcal{F} = I_{L^2}, \qquad \mathcal{F}\mathcal{F}^* = I_{\hat{L}^2}$$

   where $I_{L^2}, I_{\hat{L}^2}$ *are the identity operators on* L^2 *and* $\hat{L}^2$, *respectively.*

Proof:

1. The only thing left to prove is that for every $\hat{g} \in \hat{L}^2[-\infty, \infty]$ there is a $f \in L^2[-\infty, \infty]$ such that $\mathcal{F}f = \hat{g}$, i.e., $\mathcal{R} \equiv \{\mathcal{F}f : f \in L^2[-\infty, \infty]\} = \hat{L}^2[-\infty, \infty]$. Suppose this isn't true. Then there exists a nonzero $\hat{h} \in \hat{L}^2[-\infty, \infty]$ such that $\hat{h} \perp \mathcal{R}$, i.e., $(\mathcal{F}f, \hat{h})_{\hat{L}^2} = 0$ for all $f \in L^2[-\infty, \infty]$. But this means that $(f, \mathcal{F}^*\hat{h})_{L^2} = 0$ for all $f \in L^2[-\infty, \infty]$, so $\mathcal{F}^*\hat{h} = \Theta$. But then $||\mathcal{F}^*\hat{h}||_{L^2} = ||\hat{h}||_{\hat{L}^2} = 0$ so $\hat{h} = \Theta$, a contradiction.

2. Same proof as for 1.

3. We have shown that $\mathcal{F}\mathcal{F}^* R_{a,b} = \mathcal{F}^*\mathcal{F} R_{a,b} = R_{a,b}$ for all indicator functions $R_{a,b}$. By linearity we have $\mathcal{F}\mathcal{F}^* s = \mathcal{F}^*\mathcal{F} s = s$ for all step functions s. This implies that

$$(\mathcal{F}^*\mathcal{F}f, g)_{L^2} = (f, g)_{L^2}$$

for all $f, g \in L^2[-\infty, \infty]$. Thus

$$([\mathcal{F}^*\mathcal{F} - I_{L^2}]f, g)_{L^2} = 0$$

for all $f, g \in L^2[-\infty, \infty]$, so $\mathcal{F}^*\mathcal{F} = I_{L^2}$. An analogous argument gives $\mathcal{F}\mathcal{F}^* = I_{\hat{L}^2}$.

□

1.4 The Riemann-Lebesgue Lemma and pointwise convergence

Lemma 4 *(Riemann-Lebesgue) Suppose f is absolutely Riemann integrable in $(-\infty, \infty)$ (so that $f \in L^1[-\infty, \infty]$), and is bounded in any finite subinterval $[a, b]$, and let α, β be real. Then*

$$\lim_{\alpha \to +\infty} \int_{-\infty}^{\infty} f(t) \sin(\alpha t + \beta) dt = 0.$$

Proof: Without loss of generality, we can assume that f is real, because we can break up the complex integral into its real and imaginary parts.

1. The statement is true if $f = R_{a,b}$ is an indicator function, for

$$\int_{-\infty}^{\infty} R_{a,b}(t) \sin(\alpha t + \beta) dt = \int_a^b \sin(\alpha t + \beta) dt = \frac{-1}{\alpha} \cos(\alpha t + \beta)|_a^b \to 0$$

as $\alpha \to +\infty$.

2. The statement is true if f is a step function, since a step function is a finite linear combination of indicator functions.

3. The statement is true if f is bounded and Riemann integrable on the finite interval $[a,b]$ and vanishes outside the interval. Indeed given any $\epsilon > 0$ there exist two step functions $\overline{s}$ (Darboux upper sum) and $\underline{s}$ (Darboux lower sum) with support in $[a,b]$ such that $\overline{s}(t) \geq f(t) \geq \underline{s}(t)$ for all $t \in [a,b]$ and $\int_a^b |\overline{s} - \underline{s}| < \frac{\epsilon}{2}$. Then

$$\int_a^b f(t)\sin(\alpha t + \beta)dt =$$

$$\int_a^b [f(t) - \underline{s}(t)]\sin(\alpha t + \beta)dt + \int_a^b \underline{s}(t)\sin(\alpha t + \beta)dt.$$

Now

$$|\int_a^b [f(t) - \underline{s}(t)]\sin(\alpha t + \beta)dt| \leq \int_a^b |f(t) - \underline{s}(t)|dt \leq \int_a^b |\overline{s} - \underline{s}| < \frac{\epsilon}{2}$$

and (since $\underline{s}$ is a step function, by choosing α sufficiently large we can ensure

$$|\int_a^b \underline{s}(t)\sin(\alpha t + \beta)dt| < \frac{\epsilon}{2}.$$

Hence

$$|\int_a^b f(t)\sin(\alpha t + \beta)dt| < \epsilon$$

for α sufficiently large.

4. The statement of the lemma is true in general. Indeed

$$|\int_{-\infty}^{\infty} f(t)\sin(\alpha t + \beta)dt| \leq |\int_{-\infty}^{a} f(t)\sin(\alpha t + \beta)dt|$$

$$+|\int_a^b f(t)\sin(\alpha t + \beta)dt| + |\int_b^{\infty} f(t)\sin(\alpha t + \beta)dt|.$$

Given $\epsilon > 0$ we can choose a and b such the first and third integrals are each $< \frac{\epsilon}{3}$, and we can choose α so large the the second integral is $< \frac{\epsilon}{3}$. Hence the limit exists and is 0.

□

The sinc function has a delta-function property:

Lemma 5 *Let $c > 0$, and $F(x)$ a function on $[0,c]$. Suppose*

- *$F(x)$ is piecewise continuous on $[0,c]$*
- *$F'(x)$ is piecewise continuous on $[0,c]$*
- *$F'(+0)$ exists.*

Then

$$\lim_{L\to\infty}\int_0^\delta \frac{\sin Lx}{x}F(x)dx = \frac{\pi}{2}F(+0).$$

Proof: We write

$$\int_0^c \frac{\sin kx}{x}F(x)dx = F(+0)\int_0^c \frac{\sin Lx}{x}dx + \int_0^c \frac{F(x)-F(+0)}{x}\sin Lx\ dx.$$

Set $G(x) = \frac{F(x)-F(+0)}{x}$ for $x \in [0,\delta]$ and $G(x) = 0$ elsewhere. Since $F'(+0)$ exists it follows that $G \in L^2$. hence, by the Riemann-Lebesgue Lemma, the second integral goes to 0 in the limit as $L \to \infty$. Hence

$$\lim_{L\to\infty}\int_0^c \frac{\sin Lx}{x}F(x)dx = F(+0)\lim_{L\to\infty}\int_0^c \frac{\sin Lx}{x}dx$$

$$= F(+0)\lim_{L\to\infty}\int_0^{Lc}\frac{\sin u}{u}du = \frac{\pi}{2}F(+0).$$

For the last equality we have used our evaluation (**??**) of the integral of the sinc function. □

Theorem 8 *Let f be a complex valued function such that*

1. *$f(t)$ is absolutely Riemann integrable on $(-\infty,\infty)$.*
2. *$f(t)$ is piecewise continuous on $(-\infty,\infty)$, with only a finite number of discontinuities in any bounded interval.*
3. *$f'(t)$ is piecewise continuous on $(-\infty,\infty)$, with only a finite number of discontinuities in any bounded interval.*
4. *$f(t) = \frac{f(t+0)+f(t-0)}{2}$ at each point t.*

Let

$$\hat{f}(\lambda) = \int_{-\infty}^{\infty} f(t)e^{-i\lambda t}dt$$

be the Fourier transform of f. Then

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}(\lambda)e^{i\lambda t}d\lambda \qquad (1.4.12)$$

for every $t \in (-\infty,\infty)$.

Proof: For real $L > 0$ set

$$f_L(t) = \int_{-L}^{L} \hat{f}(\lambda)e^{i\lambda t}d\lambda = \frac{1}{2\pi}\int_{-L}^{L}\left[\int_{-\infty}^{\infty} f(x)e^{-i\lambda x}dx\right]e^{i\lambda t}d\lambda$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\left[\int_{-L}^{L} e^{i\lambda(t-x)}d\lambda\right]dx = \int_{-\infty}^{\infty} f(x)\Delta_L(t-x)dx,$$

where

$$\Delta_L(x) = \frac{1}{2\pi}\int_{-L}^{L} e^{i\lambda x}d\lambda = \begin{cases} \frac{L}{\pi} & \text{if } x = 0 \\ \frac{\sin Lx}{\pi x} & \text{otherwise.} \end{cases}$$

Here we have interchanged the order of integration, which we can since the integral is absolutely convergent. Indeed

$$\int_{-L}^{L}\left|\int_{-\infty}^{\infty} f(x)e^{-i\lambda x}e^{i\lambda t}\right|dx\ d\lambda \le \int_{-L}^{L}\int_{-\infty}^{\infty}|f(x)|dx\ d\lambda < \infty.$$

We have for any $c > 1$,

$$f_L(t) - f(t) = \int_0^{\infty}\Delta_L(x)[f(t+x) + f(t-x)]dx - f(t)$$

$$= \int_0^c \{\frac{f(t+x) + f(t-x)}{\pi x}\}\sin Lx\ dx - f(t)$$

$$+ \int_c^{\infty}[f(t+x) + f(t-x)]\frac{\sin Lx}{\pi x}\ dx$$

Now choose $\epsilon > 0$. Since

$$\left|\int_c^{\infty}[f(t+x) + f(t-x)]\frac{\sin Lx}{\pi x}\ dx\right| \le \frac{1}{\pi}\int_c^{\infty}|f(t+x) + f(t-x)|\ dx$$

and f is absolutely integrable, by choosing c sufficiently large we can make

$$\left|\int_c^{\infty}[f(t+x) + f(t-x)]\frac{\sin Lx}{\pi x}\ dx\right| < \frac{\epsilon}{2}.$$

On the other hand, by applying Lemma 5 to the expression in curly brackets we see that for this c and sufficiently large L we can achieve

$$|\int_0^c\{\frac{f(t+x)}{\pi x}\}\sin Lx\ dx - \frac{f(t+0)}{2} + \int_0^c\{\frac{f(t-x)}{\pi x}\}\sin Lx\ dx - \frac{f(t-0)}{2}| < \frac{\epsilon}{2}.$$

Thus for any $\epsilon > 0$ we can assure $|f_L(t) - f(t)| < \epsilon$ by choosing L sufficiently large. Hence $\lim_{L\to\infty} f_L(t) = f(t)$. □

Note: Condition 4 is just for convenience; redefining f at the discrete points where there is a jump discontinuity doesn't change the value of any of the integrals. The inverse Fourier transform converges to the midpoint of a jump discontinuity, just as does the Fourier series.

Exercise 1 *Assuming that the improper integral* $\int_0^\infty (\sin x/x)dx = I$ *exists, establish its value (??) by first using the Riemann - Lebesgue lemma for Fourier series to show that*

$$I = \lim_{k\to\infty} \int_0^{(k+1/2)\pi} \frac{\sin x}{x} dx = \lim_{k\to\infty} \int_0^\pi D_k(u)du$$

where $D_k(u)$ *is the Dirchlet kernel function. Then use Lemma* **??**.

Exercise 2 *Define the right-hand derivative* $f'_R(t)$ *and the left-hand derivative* $f'_L(t)$ *of* f *by*

$$f'_R(t) = \lim_{u\to t+} \frac{f(u) - f(t+0)}{u-t}, \quad f'_L(t) = \lim_{u\to t-} \frac{f(u) - f(t-0)}{u-t},$$

respectively, as in Exercise **??** *Show that in the proof of Theorem 8 we can drop the requirements 3 and 4, and the righthand side of (1.4.12) will converge to* $\frac{f(t+0)+f(t-0)}{2}$ *at any point* t *such that both* $f'_R(t)$ *and* $f'_L(t)$ *exist.*

Exercise 3 *Let* $a > 0$. *Use the Fourier transforms of* $\mathrm{sinc}(x)$ *and* $\mathrm{sinc}^2(x)$, *together with the basic tools of Fourier transform theory, such as Parseval's equation, substitution,* $\cdots$ *to show the following. (Use only rules from Fourier transform theory. You shouldn't do any detailed computation such as integration by parts.)*

- $\int_{-\infty}^\infty \left(\frac{\sin ax}{x}\right)^3 dx = \frac{3a^2\pi}{4}$
- $\int_{-\infty}^\infty \left(\frac{\sin ax}{x}\right)^4 dx = \frac{2a^3\pi}{3}$

Exercise 4 *Show that the* n*-translates of sinc are orthonormal:*

$$\int_{-\infty}^\infty \mathrm{sinc}(x-n)\cdot\mathrm{sinc}(x-m)\ dx = \begin{cases} 1 & \text{for } n = m \\ 0 & \text{otherwise,} \end{cases} \quad n, m = 0, \pm, 1, \cdots$$

Exercise 5 *Let*

$$f(x) = \begin{cases} 1 & -2 \le t \le -1 \\ 1 & 1 \le t \le 2 \\ 0 & \text{otherwise,} \end{cases}$$

- *Compute the Fourier transform* $\hat{f}(\lambda)$ *and sketch the graphs of* f *and* $\hat{f}$.
- *Compute and sketch the graph of the function with Fourier transform* $\hat{f}(-\lambda)$
- *Compute and sketch the graph of the function with Fourier transform* $\hat{f}'(\lambda)$

- *Compute and sketch the graph of the function with Fourier transform* $\hat{f} * \hat{f}(\lambda)$
- *Compute and sketch the graph of the function with Fourier transform* $\hat{f}(\frac{\lambda}{2})$

Exercise 6 *Deduce what you can about the Fourier transform $\hat{f}(\lambda)$ if you know that $f(t)$ satisfies the dilation equation*

$$f(t) = f(2t) + f(2t-1).$$

Exercise 7 *Just as Fourier series have a complex version and a real version, so does the Fourier transform. Under the same assumpitions as Theorem 8 set*

$$\hat{C}(\alpha) = \frac{1}{2}[\hat{f}(\alpha) + \hat{f}(-\alpha)], \quad \hat{S}(\alpha) = \frac{1}{2i}[-\hat{f}(\alpha) + \hat{f}(-\alpha)], \ \alpha \geq 0,$$

and derive the expansion

$$f(t) = \frac{1}{\pi}\int_0^\infty \Big(\hat{C}(\alpha)\cos\alpha t + \hat{S}(\alpha)\sin\alpha t\Big)\, d\alpha, \tag{1.4.13}$$

$$\hat{C}(\alpha) = \int_{-\infty}^{\infty} f(s)\cos\alpha s\ ds, \quad \hat{S}(\alpha) = \int_{-\infty}^{\infty} f(s)\sin\alpha s\ ds.$$

Show that the transform can be written in a more compact form as

$$f(t) = \frac{1}{\pi}\int_0^\infty d\alpha \int_{-\infty}^{\infty} f(s)\cos\alpha(s-t)\ ds.$$

Exercise 8 *There are also Fourier integral analogs of the Fourier cosine series and the Fourier sine series. Let $f(t)$ be defined for all $t \geq 0$ and extend it to an even function on the real line, defined by*

$$F(t) = \begin{cases} f(t) & \text{if } t \geq 0, \\ f(-t) & \text{if } t < 0. \end{cases}$$

By applying the results of Exercise 7 show that, formally,

$$f(t) = \frac{2}{\pi}\int_0^\infty \cos\alpha t\ d\alpha \int_0^\infty f(s)\cos\alpha s\ ds, \quad t \geq 0. \tag{1.4.14}$$

Find conditions on $f(t)$ such that this pointwise expansion is rigorously correct.

Exercise 9 *Let $f(t)$ be defined for all $t > 0$ and extend it to an odd function on the real line, defined by*

$$G(t) = \begin{cases} f(t) & \text{if } t > 0, \\ -f(-t) & \text{if } t < 0. \end{cases}$$

By applying the results of Exercise 7 show that, formally,

$$f(t) = \frac{2}{\pi} \int_0^\infty \sin \alpha t \; d\alpha \int_0^\infty f(s) \sin \alpha s \; ds, \quad t > 0. \tag{1.4.15}$$

Find conditions on $f(t)$ such that this pointwise expansion is rigorously correct.

1.5 Relations between Fourier series and Fourier integrals: sampling

For the purposes of Fourier analysis we have been considering signals $f(t)$ as arbitrary $L^2[-\infty, \infty]$ functions. In the practice of signal processing, however, one can treat only a finite amount of data. Typically the signal is digitally sampled at regular or irregular discrete time intervals. Then the processed sample alone is used to reconstruct the signal. If the sample isn't altered, then the signal should be recovered exactly. How is this possible? How can one reconstruct a function $f(t)$ exactly from discrete samples? The answer, Of course, this is not possible for arbitrary functions $f(t)$. The task isn't hopeless, however, because the signals employed in signal processing, such as voice or images, are not arbitrary. The human voice for example is easily distinguished from static or random noise. One distinguishing characteristic is that the frequencies of sound in the human voice are mostly in a narrow frequency band. In fact, any signal that we can acquire and process with real hardware must be restricted to some finite frequency band. In this section we will explore Shannon-Whittaker sampling, one way that the special class of signals restricted in frequency can be sampled and then reproduced exactly. This method is of immense practical importance as it is employed routinely in telephone, radio and TV transmissions, radar, etc. In later chapters we will study other special structural properties of signal classes, such as sparsity, that can be used to facilitate their processing and efficient reconstruction.

Definition 3 *A function f is said to be frequency band-limited if there exists a constant $\Omega > 0$ such that $\hat{f}(\lambda) = 0$ for $|\lambda| > \Omega$. The frequency $\nu = \frac{\Omega}{2\pi}$ is called the Nyquist frequency and 2ν is the Nyquist rate.*

Theorem 9 *(Shannon-Whittaker Sampling Theorem) Suppose f is a function such that*

1. *f satisfies the hypotheses of the Fourier convergence theorem 8.*
2. *$\hat{f}$ is continuous and has a piecewise continuous first derivative on its domain.*
3. *There is a fixed $\Omega > 0$ such that $\hat{f}(\lambda) = 0$ for $|\lambda| > \Omega$.*

Then f is completely determined by its values at the points $t_j = \frac{j\pi}{\Omega}$, $j = 0, \pm 1, \pm 2, \cdots$:

$$f(t) = \sum_{-\infty}^{\infty} f(\frac{j\pi}{\Omega})\frac{\sin(\Omega t - j\pi)}{\Omega t - j\pi},$$

and the series converges uniformly on $(-\infty, \infty)$.

(NOTE: The theorem states that for a frequency band-limited function, to determine the value of the function at all points, it is sufficient to sample the function at the Nyquist rate, i.e., at intervals of $\frac{\pi}{\Omega}$. The method of proof is obvious: compute the Fourier series expansion of $\hat{f}(\lambda)$ on the interval $[-\Omega, \Omega]$.)

Proof: We have

$$\hat{f}(\lambda) = \sum_{k=-\infty}^{\infty} c_k e^{\frac{i\pi k\lambda}{\Omega}}, \qquad c_k = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} \hat{f}(\lambda)e^{-\frac{i\pi k\lambda}{\Omega}}d\lambda,$$

where the convergence is uniform on $[-\Omega, \Omega]$. This expansion holds only on the interval: $\hat{f}(\lambda)$ vanishes outside the interval.

Taking the inverse Fourier transform we have

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\lambda)e^{i\lambda t}d\lambda = \frac{1}{2\pi}\int_{-\Omega}^{\Omega} \hat{f}(\lambda)e^{i\lambda t}d\lambda$$

$$= \frac{1}{2\pi}\int_{-\Omega}^{\Omega} \sum_{k=-\infty}^{\infty} c_k e^{\frac{i(\pi k + t\Omega)\lambda}{\Omega}}d\lambda$$

$$= \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} c_k \int_{-\Omega}^{\Omega} e^{\frac{i(\pi k + t\Omega)\lambda}{\Omega}}d\lambda = \sum_{k=-\infty}^{\infty} c_k \frac{\Omega\sin(\Omega t + k\pi)}{\pi(\Omega t + k\pi)}.$$

Now

$$c_k = \frac{1}{2\Omega}\int_{-\Omega}^{\Omega} \hat{f}(\lambda)e^{-\frac{i\pi k\lambda}{\Omega}}d\lambda = \frac{1}{2\Omega}\int_{-\infty}^{\infty} \hat{f}(\lambda)e^{-\frac{i\pi k\lambda}{\Omega}}d\lambda = \frac{\pi}{\Omega}f(-\frac{\pi k}{\Omega}).$$

Hence, setting $k = -j$,

$$f(t) = \sum_{j=-\infty}^{\infty} f(\frac{j\pi}{\Omega}) \frac{\sin(\Omega t - j\pi)}{\Omega t - j\pi}.$$

□

Exercise 10 *Suppose $f(t)$ satisfies the conditions of Theorem 9. Derive the Parseval formula*

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \frac{\pi}{\Omega} \sum_{k=-\infty}^{\infty} |f(\frac{\pi k}{\Omega})|^2.$$

There is a trade-off in the choice of Ω. Choosing it as small as possible reduces the sampling rate, hence the amount of data to be processed or stored. However, if we increase the sampling rate, i.e. *oversample*, the series converges more rapidly. Moreover, sampling at exactly the Nyquist rate leads to numerical instabilities in the reconstruction of the signal. This difficulty is related to the fact that the reconstruction is an expansion in sinc $(\Omega t/\pi - j) = (\sin(\Omega t - j\pi)/(\Omega t - j\pi)$. The sinc function is frequency band-limited, but its Fourier transform is discontinuous, see (**??**), (1.2.7). This causes the sinc function to decay slowly in time, like $1/(\Omega t - j\pi)$. Summing over j yields the, divergent, harmonic series:$\sum_{j=-\infty}^{\infty} |\text{sinc } (\Omega t/\pi - j)|$. Thus a small error ϵ for each sample can lead to arbitrarily large reconstruction error. Suppose we could replace sinc (t) in the expansion by a frequency band-limed function $g(t)$ such that $\hat{g}(\lambda)$ was infinitely differentiable. Since all derivatives $\hat{g}^{(n)}(\lambda)$ have compact support it follows from Section 1.2.1 that $t^n g(t)$ is square integrable for all positive integers n. Thus $g(t)$ decays faster than $|t|^{-n}$ as $|t| \to \infty$. This fast decay would prevent the numerical instability.

Exercise 11 *Show that the function*

$$h(\lambda) = \begin{cases} exp(\frac{1}{1-\lambda^2}) & \text{if } -1 < \lambda < 1 \\ 0 & \text{if } |\lambda| \geq 1, \end{cases}$$

is infinitely differentiable with compact support. In particular compute the derivatives $\frac{d^n}{d\lambda^n} h(\pm 1)$ for all n.

In order to employ such functions $g(t)$ in place of the sinc function it will be necessary to oversample. Oversampling will provide us with redundant information but also flexibility in the choice of expansion function, and improved convergence properties. We will now take samples $f(j\pi/a\Omega)$ where $a > 1$. (A typical choice is $a = 2$.) Recall that the support of $\hat{f}$ is contained in the interval $[-\Omega,\Omega]\subset[-a\Omega,a\Omega]$. We choose $g(t)$ such that 1) $\hat{g}(\lambda)$ is arbitrarily differentiable, 2) its support is contained in the interval $[-a\Omega,a\Omega]$, and 3) $\hat{g}(\lambda)=1$ for $\lambda\in[-\Omega,\Omega]$. Note that there are many possible functions g that could satisfy these requirements. Now we repeat the major steps of the proof of the sampling theorem, but for the interval $[-a\Omega,a\Omega]$. Thus

$$\hat{f}(\lambda)=\sum_{k=-\infty}^{\infty}c_k e^{\frac{i\pi k\lambda}{a\Omega}},\qquad c_k=\frac{1}{2a\Omega}\int_{-a\Omega}^{a\Omega}\hat{f}(\lambda)e^{-\frac{i\pi k\lambda}{a\Omega}}d\lambda.$$

At this point we insert $\hat{g}$ by noting that $\hat{f}(\lambda)=\hat{f}(\lambda)\hat{g}(\lambda)$, since $\hat{g}(\lambda)=1$ on the support of $\hat{f}$. Thus,

$$\hat{f}(\lambda)=\sum_{k=-\infty}^{\infty}c_k\hat{g}(\lambda)e^{\frac{i\pi k\lambda}{a\Omega}},\tag{1.5.16}$$

where from property 6 in Section 1.2.1, $\hat{g}(\lambda)e^{\frac{i\pi k\lambda}{a\Omega}}$, is the Fourier transform of $g(t+\pi k/a\Omega)$. Taking the inverse Fourier transform of both sides (OK since the series on the right converges uniformly) we obtain

$$f(t)=\sum_{j=-\infty}^{\infty}f(\frac{j\pi}{a\Omega})g(t-\pi j/a\Omega).\tag{1.5.17}$$

Since $|g(t)t^n|\rightarrow 0$ as $|t|\rightarrow\infty$ for any positive integer n this series converges very rapidly and is not subject to instabilities.

Remark: We should note that, theoretically, it isn't possible to restrict a finite length signal in the time domain $f(t)$ to a finite frequency interval. Since the support of f is bounded, the Fourier transform integral $\hat{f}(\lambda)=\int_{-\infty}^{\infty}f(t)e^{-i\lambda t}dt$ converges for all complex λ and defines $\hat{f}(\lambda)$ as an analytic function for all points in the complex plane. A well known property of functions analyic in the entire complex plane is that if they vanish along a segment of a curve, say an interval on the real axis, then they must vanish everywhere. Thus a finite length time signal cannot be frequency bounded unless it is identically 0. For practical transmission of finite signals they must be truncated in the frequency domain (but in such a way that most of the information content of the original signal is retained.)

Exercise 12 *Construct a function* $\hat{g}(\lambda)$ *which 1) is arbitrarily differentiable, 2) has support contained in the interval* $[-4,4]$*, and 3)* $\hat{g}(\lambda)=1$ *for* $\lambda\in[-1,1]$*. Hint: Consider the convolution* $\frac{1}{2c}R_{[-2,2]}*h_2(\lambda)$ *where* $R_{[-2,2]}$ *is the rectangular box function on the interval* $[-2,2]$*,* $h_2(\lambda)=h(\lambda/2)$*, h is defined in Exercise 11, and* $c=\int_{-\infty}^{\infty}h(\lambda)d\lambda$*.*

1.6 Relations between Fourier series and Fourier integrals: aliasing

Another way to compare the Fourier transform with Fourier series is to periodize a function. The periodization of a function $f(t)$ on the real line is the function

$$P[f](t) = \sum_{m=-\infty}^{\infty} f(t + 2\pi m) \qquad (1.6.18)$$

Then it is easy to see that $P[f]$ is 2π-periodic: $P[f](t) = P[f](t + 2\pi)$, assuming that the series converges. However, this series will not converge in general, so we need to restrict ourselves to functions that decay suffiuciently rapidly at infinity. We could consider functions with compact support, say infinitely differentiable. Another useful but larger space of functions is the Schwartz class. We say that $f \in L^2[-\infty, \infty]$ belongs to the *Schwartz class* if f is infinitely differentiable everywhere, and there exist constants $C_{n,q}$ (depending on f) such that $|t^n \frac{d^q}{dt^q} f| \le C_{n,q}$ on R for each $n, q = 0, 1, 2, \cdots$. Then the projection operator P maps an f in the Schwartz class to a continuous function in $L^2[0, 2\pi]$ with period 2π. (However, periodization can be applied to a much larger class of functions, e.g. functions on $L^2[-\infty, \infty]$ that decay as $\frac{c}{t^2}$ as $|t| \to \infty$.) Assume that f is chosen appropriately so that its periodization is a continuous function. Thus we can expand $P[f](t)$ in a Fourier series to obtain

$$P[f](t) = \sum_{n=-\infty}^{\infty} c_n e^{int}$$

where

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} P[f](t) e^{-int} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-int} dx = \frac{1}{2\pi} \hat{f}(n)$$

where $\hat{f}(\lambda)$ is the Fourier transform of $f(t)$. Then,

$$\sum_{n=-\infty}^{\infty} f(t + 2\pi n) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{int}, \qquad (1.6.19)$$

and we see that $P[f](t)$ tells us the value of $\hat{f}$ at the integer points $\lambda = n$, but not in general at the non-integer points. (For $t = 0$, equation (1.6.19) is known as the *Poisson summation formula*. If we think of f as a signal, we see that **periodization** (1.6.18) of f results in a loss of information. However, if f vanishes outside of $[0, 2\pi)$) then $P[f](t) \equiv f(t)$ for $0 \le t < 2\pi$ and

$$f(t) = \sum_n \hat{f}(n) e^{int}, \quad 0 \le t < 2\pi$$

without error.)

Exercise 13 *Let* $f(t) = \frac{a}{t^2+a^2}$ *for* $a > 0$.

- *Show that* $\hat{f}(t) = \pi e^{-a|\lambda|}$. *Hint: It is easier to work backwards.*
- *Use the Poisson summation formula to derive the identity*

$$\sum_{n=-\infty}^{\infty} \frac{1}{n^2+a^2} = \frac{\pi}{a}\frac{1+e^{-2\pi a}}{1-e^{-2\pi a}}.$$

  What happens as $a \to 0+$ *? Can you obtain the value of* $\sum_{n=1}^{\infty} \frac{1}{n^2}$ *from this?*

1.7 The Fourier integral and the uncertainty relation of quantum mechanics

The uncertainty principle gives a limit to the degree that a function $f(t)$ can be simultaneously localized in time as well as in frequency. To be precise, we introduce some notation from probability theory.

We recall some basic definitions from probability theory. A continuous probability distribution for the random variable t on the real line R is a continuous function $\rho(t)$ on R such that $0 \le \rho(t) \le 1$ and $\int_{-\infty}^{\infty} \rho(t)\, dt = 1$. We also require that $\int_{-\infty}^{\infty} p(t)\rho(t)\, dt$ converges for any polynomial $p(t)$. Here $\int_{t_1}^{t_2} \rho(t)\, dt$ is interpreted as the probability that a sample t taken from R falls in the interval $t_1 \le t \le t_2$. The expectation (or mean) $\bar{t}$ of the distribution is $\bar{t} = E_\rho(t) \equiv \int_{-\infty}^{\infty} t\rho(t)\, dt$ and the standard deviation $\sigma \ge 0$ is defined by

$$\sigma^2 = \int_{-\infty}^{\infty} (t-\bar{t})^2 \rho(t)\, dt = E_\rho\left((t-\bar{t})^2\right).$$

Here σ is a measure of the concentration of the distribution about its mean. The most famous continuous distribution is the normal (or Gaussian) distribution function

$$\rho_0(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(t-\mu)^2/2\sigma^2} \tag{1.7.20}$$

where μ is a real parameter and $\sigma > 0$. This is just the bell curve, centered about $t = \mu$. In this case $E_{\rho_0}(t) = \mu$ and $\sigma^2 = E_{\rho_0}(t^2)$. The standard notation for the normal distribution with mean μ and standard deviation σ is $N(\mu, \sigma)$.

Every nonzero continuous $f \in L^2[-\infty, \infty]$ defines a probability distribution function $\rho(t) = \frac{|f(t)|^2}{||f||^2}$, i.e., $\rho(t) \ge 0$ and $\int_{-\infty}^{\infty} \rho(t) dt = 1$. For convenience we will normalize the function, so that $||f|| = 1$.

Definition 4 • *The mean of the distribution defined by the normalized f is*

$$\bar{t} = \int_{-\infty}^{\infty} t|f(t)|^2 dt.$$

• *The dispersion of f about $a \in R$ is*

$$\Delta_a f = \int_{-\infty}^{\infty} (t-a)^2 |f(t)|^2 dt.$$

($\Delta_{\bar{t}} f$ is called the variance of f, and $\sqrt{\Delta_{\bar{t}} f}$ the standard deviation.

The dispersion of f about a is a measure of the extent to which the graph of f is concentrated at a. If $f = \delta(x-a)$ the "Dirac delta function", the dispersion is zero. The constant $f(t) \equiv 1$ has infinite dispersion. (However there are no such L^2 functions.) Similarly we can define the dispersion of the Fourier transform of f about some point $\alpha \in R$:

$$\Delta_\alpha \hat{f} = \int_{-\infty}^{\infty} (\lambda - \alpha)^2 |\hat{f}(\lambda)|^2 d\lambda.$$

Note: $\int_{-\infty}^{\infty} |\hat{f}(\lambda)|^2 d\lambda = 1$. Also it makes no difference which definition of the Fourier transform that we use, $\hat{f}$ or $\mathcal{F}f$, because the normalization gives the same probability measure.

Example 2 *Let $f_s(t) = (\frac{2s}{\pi})^{1/4} e^{-st^2}$ for $s > 0$, the Gaussian distribution. From the fact that $\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$ we see that $||f_s|| = 1$. The normed Fourier transform of f_s is $\hat{f}_s(\lambda) = (\frac{2}{s\pi})^{1/4} e^{\frac{-\lambda^2}{4s}}$. By plotting some graphs one can see informally that as s increases the graph of f_s concentrates more and more about $t = 0$, i.e., the dispersion $\Delta_0 f_s$ decreases. However, the dispersion of $\hat{f}_s$ increases as s increases. We can't make both values, simultaneously, as small as we would like. Indeed, a straightforward computation gives*

$$\Delta_0 f_s = \frac{1}{4s}, \qquad \Delta_0 \hat{f}_s = s,$$

so the product of the variances of f_s and $\hat{f}_s$ is always $\frac{1}{4}$, no matter how we choose s.

Theorem 10 *(Heisenberg inequality, Uncertainty theorem) If $f(t) \neq 0$ and $tf(t)$ belong to $L^2[-\infty, \infty]$ then $\Delta_a f \Delta_\alpha \hat{f} \geq \frac{1}{4}$ for any $a, \alpha \in R$.*

Sketch of proof: We will give the proof under the added assumptions that $f'(t)$ exists everywhere and also belongs to $L^2[-\infty,\infty]$. (In particular this implies that $f(t)\to 0$ as $t\to\pm\infty$.) The main ideas occur there.

We make use of the canonical commutation relation of quantum mechanics, the fact that the operations of multiplying a function $f(t)$ by t, $(Tf(t) = tf(t))$ and of differentiating a function $(Df(t) = f'(t))$ don't commute: $DT - TD = I$. Thus

$$\frac{d}{dt}\left[tf(t)\right] - t\left[\frac{d}{dt}f(t)\right] = f(t).$$

Now it is easy from this to check that

$$(\frac{d}{dt} - i\alpha)\left[(t-a)f(t)\right] - (t-a)\left[(\frac{d}{dt} - i\alpha)f(t)\right] = f(t)$$

also holds, for any $a,\alpha\in R$. (The a,α dependence just cancels out.) This implies that

$$\left((\frac{d}{dt} - i\alpha)[(t-a)f(t)], f(t)\right) - \left((t-a)[(\frac{d}{dt}\right.$$

$$-i\alpha)f(t)], f(t)) = (f(t), f(t)) = ||f||^2.$$

Integrating by parts in the first integral, we can rewrite the identity as

$$-\left([(t-a)f(t)], [(\frac{d}{dt} - i\alpha)f(t)]\right) - \left([(\frac{d}{dt} - i\alpha)f(t)], [(t-a)f(t)]\right) = ||f||^2.$$

The Schwarz inequality and the triangle inequality now yield

$$||f||^2 \le 2||(t-a)f(t)||\cdot||(\frac{d}{dt} - i\alpha)f(t)||. \qquad (1.7.21)$$

From the list of properties of the Fourier transform in Section 1.2.1 and the Plancherel formula, we see that $||(\frac{d}{dt} - i\alpha)f(t)|| = \frac{1}{\sqrt{2\pi}}||(\lambda-\alpha)\hat{f}(\lambda)||$ and $||f|| = \frac{1}{\sqrt{2\pi}}||\hat{f}||$. Then, dividing by $||f||$ and squaring, we have

$$\Delta_a f \Delta_\alpha \hat{f} \ge \frac{1}{4}.$$

□

Note: Normalizing to $a=\alpha=0$ we see that the Schwarz inequality becomes an equality if and only if $2stf(t) + \frac{d}{dt}f(t) = 0$ for some constant s. Solving this differential equation we find $f(t) = c_0 e^{-st^2}$ where c_0 is the integration constant, and we must have $s>0$ in order for f to be square integrable. Thus the Heisenberg inequality becomes an equality only for Gaussian distributions.

Exercise 14 *Let*

$$f(t) = \begin{cases} 0 & \text{if } t < 0, \\ \sqrt{2}e^{-t} & \text{if } t \geq 0. \end{cases}$$

Compute $\Delta_a f \Delta_\alpha \hat{f}$ *for any* $a, \alpha \in R$ *and compare with Theorem 10.*

These considerations suggest that for proper understanding of signal analysis we should be looking in the two dimensional time-frequency space (phase space), rather than the time domain or the frequency domains alone. In subsequent chapters we will study tools such as windowed Fourier transforms and wavelet transforms that probe the full phase space. The Heisenberg inequality also suggests that probabilistic methods have an important role to play in signal analysis, and we shall make this clearer in later sections.

1.8 Digging deeper. Probabilistic tools.

The notion of a probability distribution can easily be extended to n dimensions. A continuous probability distribution (multivariate distribution) for the vector random variables $t = (t_1, t_2, \cdots, t_n) \in R^n$ is a continuous function $\rho(t) = \rho(t_1, \cdots, t_n)$ on R^n such that $0 \leq \rho(t) \leq 1$ and $\int_{R^n} \rho(t)\, dt_1 \cdots dt_n = 1$. We also require that $\int_{-\infty}^{\infty} p(t)\rho(t)\, dt$ converges for any polynomial $p(t_1, \cdots, t_n)$. If S is an open subset of R^n and $\chi_S(t)$ is the charactistic function of S, i.e.,

$$\chi_S(t) = \begin{cases} 1 & \text{if } t \in S \\ 0 & \text{if} t \notin S, \end{cases}$$

then $\int_S \rho(t)\, dt_1 \cdots dt_n$ is interpreted as the probability that a sample t taken from R^n lies in the set S. The expectation (or mean) μ_i of random variable t_i is

$$\mu_i = E_\rho(t_i) \equiv \int_{R^n} t_i \rho(t)\, dt_1 \cdots\, dt_n, \quad i = 1, 2, \cdots, n$$

and the standard deviation $\sigma_i \geq 0$ is defined by

$$\sigma_i^2 = \int_{R^n} (t_i - \mu_i)^2 \rho(t)\, dt_1 \cdots dt_n = E_\rho\left((t_i - \mu_i)^2\right).$$

The **covariance matrix** of the distribution is the $n \times n$ symmetric matrix

$$C(i, j) = E\left((t_i - \mu_i)(t_j - \mu_j)\right), \quad 1 \leq i, j \leq n.$$

Note that $\sigma_i^2 = C(i, i)$, In general, the expectation of any function $f(t_1, \cdots, t_t)$ of the random variables is defined as $E\left(f(t_1, \cdots, t_n)\right)$.

Exercise 15 *Show that the eigenvalues of a covariance matrix are nonnegative.*

One of the most important multivariate distributions is the multivariate normal distribution

$$\rho(t) = \frac{1}{\sqrt{(2\pi)^n \det(C)}} \exp\left[-\frac{1}{2}(t-\mu)C^{-1}(t-\mu)^{\rm tr}\right]. \qquad (1.8.22)$$

Here μ is the column vector with components $\mu_1, \cdots, \mu_n$ where $\mu_i = E(t_i)$ and C is an $n \times n$ nonsingular symmetric matrix.

Exercise 16 *Show that C is in fact the covariance matrix of the distribution (1.8.22). This takes some work and involves making an orthogonal change of coordinates where the new coordinate vectors are the orthonormal eigenvectors of C.*

Important special types of multivariate distributions are those in which the random variables are independently distributed, i.e., there are 1 variable probability distributions $\rho_1(t_1), \cdots, \rho_n(t_n)$ such that $\rho(t) = \Pi_{i=1}^n \rho_i(t_1)$. If, further, $\rho_1(\tau) = \cdots = \rho_n(\tau)$ for all τ we say that the random variables are independently and identically distributed (iid).

Exercise 17 *If the random variables are independently distributed, show that*

$$C(t_i, t_j) = \sigma_i^2 \delta_{ij}.$$

Example 3 *If C is a diagonal matrix in the multivariate normal distribution (1.8.22) then the random variables $t_1, \cdots,$ are independently distributed. If $C = \sigma^2 I$ where I is the $n \times n$ identity matrix and $\mu_i = \mu$ for all i then the distribution is iid. where each random variable is is distributed according to the normal distribution $N(\mu, \sigma^2)$, i.e., the Gaussian distribution (1.7.20) with mean μ and variance σ^2.*

An obvious way to construct an iid multivariate distribution is to take a random sample $T_1, \cdots, T_n$ from of values of a single random variable t with probability distribution $\rho(t)$. Then the multivariate distribution function for the vector random variables $T = (T_1, \cdots T_n)$ is the function $\rho(T) = \Pi_{i=1}^n \rho(T_i)$. It follows that for any integers $k_1, \cdots, k_n$ we have

$$E(T_1^{k_1} \cdots T_n^{k_n}) = \Pi_{i=1}^n E(T_i^{k_i}).$$

Note that if t has mean μ and standard deviation σ^2 then $E(T_i) = \mu$, $E((T_i - \mu)^2) = \sigma^2$ and $E(T_i^2) = \sigma^2 + \mu^2$.

Exercise 18 *Show that* $E(T_iT_j) = \mu^2 + \sigma^2\delta_{ij}$.

Now we define the **sample mean** of the random sample as

$$\bar{T} = \frac{1}{n}\sum_{i=1}^{n} T_i.$$

The sample mean is itself a random variable with expectation

$$E(\bar{T}) = E\left(\frac{1}{n}\sum_i T_i\right) = \frac{1}{n}\sum_i E(T_i) = \frac{1}{n}\sum_i \mu = \mu \tag{1.8.23}$$

and variance

$$E\left((\bar{T}-\mu)^2\right) = \frac{1}{n^2}\sum_{i,j=1}^{n} E(T_iT_j) - \frac{2\mu}{n}\sum_{i=1}^{n} E(T_i) + E(\mu^2) \tag{1.8.24}$$

$$= \frac{1}{n^2}(\sum_{i,j}(\mu^2 + \delta_{ij}\sigma^2) - \frac{2n\mu^2}{n} + \mu^2 = \mu^2 + \frac{n}{n^2}\sigma^2 - \mu^2 = \frac{\sigma^2}{n}.$$

Thus the sample mean has the same mean as a random variable as t, but its variance is less than the variance of t by the factor $1/n$. This suggests that the distribution of the sample mean is increasingly "peaked" around the mean as n grows. Thus if the original mean is unknown, we can obtain better and better estimates of it by taking many samples. This idea lies behind the Law of Large Numbers that we will prove later.

A possible objection to the argument presented in the previous paragraph is that we already need to know the mean of the distribution to compute the varience of the sample mean. This leads us to the definition of the **sample variance**:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(T_i - \bar{T})^2, \tag{1.8.25}$$

where $\bar{T}$ is the sample mean. Here S is the **sample standard deviation** (The sample mean and sample standard deviation are typically reported as outcomes for an exam in a high enrollment undergraduate math course.) We will explain the factor $n-1$, rather than n.

Theorem 11 *The expectation of the sample variance* S^2 *is* $E(S^2) = \sigma^2$.

Proof:

$$E(S^2) = \frac{1}{n-1}E(\sum_i (T_i - \bar{T})^2) = \frac{1}{n-1}\left[\sum_i E(T_i^2) - 2E(\bar{T}\sum_i T_i) + nE(\bar{T}^2)\right]$$

$$= \frac{1}{n-1}\left[n(\mu^2+\sigma^2) - \frac{2}{n}\sum_{i,j} E(T_iT_j) + \frac{1}{n}\sum_{i,j} E(T_iT_j)\right]$$

$$= \frac{1}{n-1}[(n-2n+n)\mu^2 + (n-2+1)\sigma^2] = \sigma^2. \quad \square$$

If we had used n as the denominator of the sample variance then the expectation would not have been σ^2.

We continue to explore the extent to which the probability distribution function of a random variable peaks around its mean. For this we need to make use of a form of the Markov inequality, a simple but powerful result that applies to a wide variety of probability distributions.

Theorem 12 *Let x be a nonnegative random variable with continuous probability distribution function $p(x)$ (so that $p(x) = 0$ for $x < 0$) and let d be a positive constant. Then*

$$\Pr(x \geq d) \leq \frac{1}{d}E(x). \tag{1.8.26}$$

PROOF: The theorem should be interpreted as saying that the probability that a random selection of the variable x is $\geq d$ is $\leq \frac{1}{d}E(x)$. Now

$$E(x) = \int_0^\infty xp(x)\ dx = \int_0^d xp(x)\ dx + \int_d^\infty xp(x)\ dx$$

$$\geq \int_0^d xp(x)\ dx + d\int_d^\infty p(x)\ dx \geq d\int_d^\infty p(x)\ dx = d\ \Pr(X \geq d).$$

$\square$

Corollary 1 *Chebyshev's inequality. Let t be a random variable with expected value μ and finite variance σ^2. Then for any real number $\alpha > 0$,*

$$\Pr(|t-\mu| \geq \alpha\sigma) \leq \frac{1}{\alpha^2}.$$

Proof: We apply the Markov inequality to the random variable $x = (t-\mu)^2$ with $d = (\sigma\alpha)^2$. Thus

$$\Pr((t-\mu)^2 \geq (\sigma\alpha)^2) \leq \frac{1}{(\sigma\alpha)^2}E((t-\mu)^2) = \frac{1}{(\alpha)^2}.$$

This is equivalent to the statement of the Chebyshev inequality. $\square$

Example 4 *Setting $\alpha = 1/\sqrt{2}$ we see that at least half of the sample values will lie in the interval $(\mu - \sigma/\sqrt{2}, \mu + \sigma/\sqrt{2})$.*

Corollary 2 *Law of Large Numbers. Let $\rho(t)$ be a probability distribution with mean μ and standard deviation σ'. Take a sequence of independent random samples from this population: $T_1, \cdots, T_n, \cdots$ and let $\bar{T}^{(n)} = \frac{1}{n}\sum_{i=1}^{n} T_i$ be the sample mean of the first n samples. Then for any $\epsilon > 0$ we have $\lim_{n\to\infty} \Pr(|\bar{T}^{(n)} - \mu| > \epsilon) = 0$.*

Proof: From equations (1.8.23) and (1.8.24) we have $E(\bar{T}^{(n)}) = \mu$, $E\left((\bar{T}^{(n)} - \mu)^2\right) = \sigma'^2/n$. Applying the Chebyshev inequality to the random variable $\bar{T}^{(n)}$ with $\sigma = \sigma'/\sqrt{n}$, $\alpha = \sqrt{n}\epsilon/\sigma'$ we obtain

$$\Pr(\left|\bar{T}^{(n)} - \mu\right| \geq \epsilon) \leq \frac{\sigma'^2}{n\epsilon^2}.$$

Thus the probability that the sample mean differs by more than ϵ from the distribution meam gets smaller as n grows and approaches 0. In particular $\lim_{n\to\infty} \Pr(|\bar{T}^{(n)} - \mu| > \epsilon) = 0$. □

This form of the Law of Large Numbers tells us that for any fixed $\epsilon > 0$ and sufficiently large sample size, we can show that the sample average will differ by less than ϵ from the mean *with high probability*, but not with certainty. It shows us that the sample distribution is more and more sharply peaked around the mean as the sample size grows. With modern computers than can easily generate large random samples and compute sample means, this insight forms the basis for many practical applications, as we shall see. decreasing magnitude down the diagonal.

1.9 Additional Exercises

Exercise 19 *Find the Fourier transform of the following functions (a sketch may help!). Also write down the inversion formula for each, taking account of where they are discontinuous.*

(i) Let $A, T > 0$. Let f be the rectangular pulse

$$f(t) = \begin{cases} A, & t \in [0, T] \\ 0, & t \notin [0, T] \end{cases}$$

(ii) Let $A, T > 0$. Let f be the two-sided pulse

$$f(t) = \begin{cases} -A, & t \in [-T, 0] \\ A, & t \in (0, T] \\ 0, & t \notin [-T, T] \end{cases}$$

(iii) Let f be the triangular pulse

$$f(t) = \begin{cases} t+1, & t \in [-1,0] \\ 1-t, & t \in (0,1] \\ 0, & t \notin [-1,1] \end{cases}$$

Deduce that

$$\int_0^\infty \frac{\sin^2(x/2)}{x^2} dx = \frac{\pi}{4}.$$

(iv) Let $a > 0$ and

$$f(t) := \begin{cases} \sin(at), & |t| \leq \pi/a \\ 0, & \text{else} \end{cases}$$

(v) Let

$$f(t) := \begin{cases} 0, & t < 0 \\ \exp(-t), & t \geq 0 \end{cases}$$

Deduce that

$$\int_0^\infty \frac{\cos(xt) + x\sin(xt)}{1+x^2} dx = \begin{cases} \pi\exp(-t), & t > 0 \\ \pi/2, & t = 0 \\ 0, & t < 0 \end{cases}$$

(vi) Let $a, b \in \mathbb{R}$ and $f(t) := \exp^{-|at+b|}$, $t \in \mathbb{R}$.

(vii) Let $f(t) := (t^2 - 1)\exp^{-t^2/2}$, $t \in \mathbb{R}$.

Exercise 20 *Prove the following: If f is even,*

$$\mathcal{F}[f](x) = \sqrt{\frac{2}{\pi}} \int_0^\infty f(t)\cos(xt)dt$$

and if f is odd,

$$\mathcal{F}[f](x) = \sqrt{\frac{2}{\pi}} \int_0^\infty f(t)\sin(xt)dt.$$

Exercise 21 *The Fourier Cosine ($\mathcal{F}_c[f](.)$) and Fourier Sine ($\mathcal{F}_s[f](.)$) of $f : \mathbb{R} \to \mathbb{R}$ are defined as follows:*

$$\mathcal{F}_c[f](x) := \sqrt{\frac{2}{\pi}} \int_0^\infty f(t)\cos(xt)dt.$$

$$\mathcal{F}_s[f](x) := \sqrt{\frac{2}{\pi}} \int_0^\infty f(t)\sin(xt)dt.$$

Find the Fourier Cosine and Sine transform of the following funcions:

$$f(t) := \begin{cases} 1, & t \in [0, a] \\ 0, & t > a \end{cases}$$

$$f(t) := \begin{cases} \cos(at), & t \in [0, a] \\ 0, & t > a \end{cases}$$

Exercise 22 *Let $f, g, h : \mathbb{R} \to \mathbb{R}$. Let $a, b \in \mathbb{R}$. The following question deals with (convolution *): Show that:*

(i) ** is linear:*

$$(af + bg) * h = a(f * h) + b(g * h).$$

(ii) ** is commutative:*

$$f * g = g * f.$$

(iii) ** is associative:*

$$(f * g) * h = f * (g * h).$$

Exercise 23 *Let $g : \mathbb{R} \to \mathbb{R}$. Find a function H such that for all x,*

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} g(t)dt = (H * g)(x).$$

(H is called the Heaviside function).

Exercise 24 *Let $f, g : \mathbb{R} \to \mathbb{R}$. Let f' exist. Assuming the convergence of the relevent integrals below, show that*

$$(f * g)'(x) = f'(x) * g(x).$$

Exercise 25 *For $a \in \mathbb{R}$, let*

$$f_a(t) = \begin{cases} 0, & t < a \\ 1, & t \geq a \end{cases}$$

Compute $f_a * f_b$ for $a, b \in \mathbb{R}$. Deduce that

$$(f_a * f_{-a})(x) = \frac{x f_0(x)}{\sqrt{2\pi}}.$$

Does $f_a * (1 - f_b)$ exist? For $a \in \mathbb{R}$, let

$$g_a(t) := \begin{cases} 0, & t < 0 \\ \exp(-at), & t \geq 0 \end{cases}$$

Compute $g_a * g_b$.

Exercise 26 *Fourier transforms are useful in "deconvolution" or solving "convolution integral equations". Suppose, that we are given functions g, h and are given that*

$$f * g = h.$$

Our task is to find f in terms of g, h.

(i) *Show that*

$$\mathcal{F}[f] = \mathcal{F}[h]/\mathcal{F}[g]$$

and hence, if we can find a function k such that

$$\mathcal{F}[h]/\mathcal{F}[g] = \mathcal{F}[k]$$

then $f = k$.

(ii) *As an example, suppose that*

$$f * \exp(-t^2/2) = (1/2)t\exp(-t^2/4).$$

Find f.

Exercise 27 (i) *We recall that the* **Laplace transform** *of a function $f : [0, \infty) \to \mathbb{R}$ is defined as*

$$\mathcal{L}[f](p) = \int_0^\infty f(t)\exp(-pt)dt$$

whenever the right hand side makes sense. Show formally, that if we set

$$g(x) := \begin{cases} f(x), & x \geq 0 \\ 0, & x < 0 \end{cases}$$

then

$$\mathcal{L}[f](p) := \sqrt{2\pi}\mathcal{F}[g](-ip).$$

(ii) *Let $h : \mathbb{R} \to \mathbb{R}$ and define:*

$$h_+(x) := \begin{cases} h(x), & x \geq 0 \\ 0, & x < 0 \end{cases}$$

and

$$h_-(x) := \begin{cases} h(-x), & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Show that $h(x) = h_+(x) + h_-(x)$ and express $\mathcal{F}[h]$ in terms of $\mathcal{L}[h_+]$ and $\mathcal{L}[h_-]$.