Fourier Transforms

Gary D. Knott
Civilized Software Inc.
12109 Heritage Park Circle
Silver Spring MD 20906
phone:301-962-3711
email:knott@civilized.com
URL:www.civilized.com

February 12, 2024

1 Fourier Series

Around 1740, Daniel Bernoulli, Jean D'Alembert, and Leonhard Euler separately realized that, under poorly-understood conditions, a real-valued periodic period-p function, $y(t) = y(t + p)$, for $t \in \mathcal{R}$ with p a fixed constant in $\mathcal{R}^+$, could be expressed as the potentially-infinite sum, *i.e.*, a potentially-infinite *series*, of sinusoidal oscillations of various frequencies, amplitudes, and phaseshifts, so that

$$y(t) = \sum_{h=0}^{\infty} M(h/p) \cdot cos\left(2\pi(h/p)t + \phi(h/p)\right).$$

(Here $\mathcal{R}$ denotes the set of real numbers, and $\mathcal{R}^+$ denotes the set of positive real numbers.) This series is called the real spectral-decomposition-form *Fourier* series of the period-p function y because of Jean Baptiste Joseph Fourier's book on heat transfer that explored the use of trigonometric series in representing the solutions of differential equations (submitted to the Paris Academy of Sciences in 1807. [Sti86])

The term $M(h/p) \cdot cos(2\pi(h/p)t + \phi(h/p))$ is a cosine *oscillation* of *period* p/h, *frequency* h/p, *amplitude* $M(h/p)$, and *phaseshift* $\phi(h/p)$. The period-p function y determines and is determined by the amplitude function M and the phase function ϕ, which are both defined on the frequency values $\{0, 1/p, 2/p, \ldots\}$.

It is convenient to use Euler's relation $e^{i\theta} = cos(\theta) + i\,sin(\theta)$ to develop the mathematical theory of Fourier series for complex-valued functions of a real argument, rather than just real-valued functions. In this case, we can express the period-p function y in terms of an associated discrete complex-valued function $y^\wedge$, which contains the amplitude and phase functions combined together. The complex-valued function $y^\wedge$ is defined on the discrete set $\{\ldots, -2/p, -1/p, 0, 1/p, 2/p, \ldots\}$. This function $y^\wedge$ will be introduced below; it is called the *Fourier transform* of y.

Consider

$$x(t) = \sum_{h=-\infty}^{\infty} C_h e^{2\pi i(h/p)t} := \lim_{k\to\infty} \sum_{h=-k}^{k} C_h e^{2\pi i(h/p)t},$$

where $\ldots, C_{-1}, C_0, C_1, \ldots$ are complex numbers and $p > 0$. If $|C_h| \to 0$ as $|h| \to \infty$ fast enough so that $\sum_{h=-\infty}^{\infty} |C_h|^2 < \infty$, then the series defining $x(t)$ converges in the least-squares sense [DM72] and also converges pointwise almost everywhere [Edw67]; it is the *Fourier series* for the periodic complex function x of period p. The complex numbers $\ldots, C_{-1}, C_0, C_1, \ldots$ are called the *Fourier coefficients* of the function x. (Let $S_k(t) = \sum_{h=-k}^{k} C_h e^{2\pi i(h/p)t}$; S_k is the k-th partial sum of the series $x(t) = \sum_{h=-\infty}^{\infty} C_h e^{2\pi i(h/p)t}$. To say $\sum_{h=-\infty}^{\infty} C_h e^{2\pi i(h/p)t}$ converges in the least-squares sense means $\int_{-p/2}^{p/2} |S_n(t) - S_m(t)|^2 dt \to 0$ as $n \to \infty$ and $m \to \infty$ jointly.)

Note that, for $h \neq 0$, $e^{2\pi i(h/p)t}$ is a period-$(p/|h|)$ periodic function of t, and for $h = 0$, $e^{2\pi i(h/p)t}$ is constant and hence is a periodic function for every positive period. The term $C_h e^{2\pi i(h/p)t}$ is called a *complex oscillation with the frequency* h/p *and the complex amplitude* C_h. For h a non-zero integer, the frequency h/p is called a *harmonic frequency* of the *fundamental frequency* $1/p$. (In general, the frequency hq is a harmonic frequency of the frequency q when h is a non-zero integer.) A periodic function with the frequency h/p has the period $p/|h|$ and any integral multiple of $p/|h|$ is also a period. The Fourier series of x is thus the sum of a constant term, C_0, a pair of complex oscillation terms with the fundamental frequencies $\pm 1/p$, and all the complex oscillations at all the harmonic frequencies of the fundamental frequencies. (Of course, the amplitudes of some or all of these terms might be zero, in which case the corresponding harmonics are missing.)

Exercise 1.1: Show that $e^{2\pi i(h/p)t}$ is a period-$p/|h|$ function.

Solution 1.1: Since $cos(\alpha \pm 2\pi = cos(\alpha)$ and $sin(\alpha \pm 2\pi = sin(\alpha)$, we have $cos(2\pi(h/p)(t + p/|h|)) = cos(2\pi(h/p)t \pm 2\pi) = cos(2\pi(h/p)t)$ and $sin(2\pi(h/p)(t + p/|h|)) = sin(2\pi(h/p)t \pm 2\pi) = sin(2\pi(h/p)t)$, and thus the functions $cos(2\pi(h/p)t)$ and $sin(2\pi(h/p)t)$ are period-$p/|h|$ functions of t, and $e^{2\pi i(h/p)t} = cos(2\pi(h/p)t) + i\, sin(2\pi(h/p)t)$, $e^{2\pi i(h/p)t}$ is a period-$p/|h|$ function of t. Alternately,

$$e^{2\pi i(h/p)(t+p/|h|)} = e^{2\pi i(h/p)t} \cdot e^{2\pi i\, sign(h)} = e^{2\pi i(h/p)t} \cdot 1 = e^{2\pi i(h/p)t}$$

since $e^{2\pi i} = 1$ and $e^{2\pi i(-1)} = 1/e^{2\pi i} = 1$.

Exercise 1.2: Show that e^{it} is a period=2π periodic function of t, and show that e^t is a period-$2\pi i$ periodic function of t.

Exercise 1.3: Show that $|C_h e^{2\pi i(h/p)t}| = |C_h|$.

The Fourier series for x can be manipulated to produce the spectral decomposition-form Fourier series of x. The spectral decomposition of x consists of an *amplitude spectrum function* M, and a *phase spectrum function* ϕ, where the amplitude spectrum function and the phase spectrum function both appear in the spectral decomposition-form Fourier series of x.

When x is real, this is

$$x(t) = \sum_{h=0}^{\infty} M(h/p) \cdot cos\left(2\pi(h/p)t + \phi(h/p)\right).$$

When x is real, the functions $M(h/p)$ and $\phi(h/p)$ are defined for $h \geq 0$ as:

$$M(h/p) = \frac{\sqrt{A_h^2 + B_h^2}}{1 + \delta_{h,0}}$$

and

$$\phi(h/p) = \text{atan2}(-B_h, A_h),$$

where

$$A_h = C_h + C_{-h} \quad \text{and} \quad B_h = i(C_h - C_{-h}).$$

The function $\text{atan2}(y, x)$ is defined to be the angle θ in radians lying in $[-\pi, \pi)$ formed by the vectors $(1, 0)$ and (x, y) in the xy-plane, with $\text{atan2}(0, 0) := \pi/2$. When $x > 0$, $\text{atan2}(y, x) = tan^{-1}(y/x)$. When $y > 0$, $\text{atan2}(y, x) \in (0, \pi)$, when $y < 0$, $\text{atan2}(y, x) \in (-\pi, 0)$, and $\text{atan2}(0, x) = \begin{cases} 0 & \text{if } x > 0 \\ -\pi & \text{if } x < 0 \end{cases}$. Note that $\text{atan2}(y, x) = -\text{atan2}(-y, x)$ for $y \neq 0$.

Recall that the Kroneker delta function $\delta_{h,k}$ is defined by

$$\delta_{h,k} = \begin{cases} 1 & \text{if } h = k; \\ 0 & \text{otherwise.} \end{cases}$$

The period-p function $x(t)$ is real-valued if and only if $C_{-h} = C_h^*$ for all h, where C_h^* is the complex conjugate of C_h. In this case A_h and B_h are real with $A_h = 2\,\text{Re}(C_h)$ and $B_h = -2\,\text{Im}(C_h)$, and M and ϕ are real-valued with $M(h/p) = 2 \cdot |C_h|$ for $h > 0, M(0) = |C_0|$, and $\phi(h/p) = \text{atan2}\left(\text{Im}(C_h), \text{Re}(C_h)\right)$. Also, with x real, $cos(\phi(0)) = sign(C_0)$ and $M(h/p) \geq 0$.

For $h = 0, 1, 2, \ldots$, the function of t, $M(h/p) \cdot cos\left(2\pi(h/p)t + \phi(h/p)\right)$, is periodic with period p/h, and is a sinusoidal oscillation of frequency h/p. Thus $x(t)$ is a sum of oscillations of periods ∞, p, $p/2$, $p/3$, $\ldots$ (and frequencies 0, $1/p$, $2/p$, $\ldots$), where the oscillation of frequency h/p has the phase shift $\phi(h/p)$ and the amplitude $M(h/p)$. The frequencies $0, \frac{1}{p}, \frac{2}{p}, \ldots$ are all harmonic frequencies of the fundamewntal frequency $\frac{1}{p}$, so the corresponding periods $\infty, p, p/2, \ldots$ all evenly divide the fundametal period p. The limit $\sum_{h=0}^{\infty} M(h/p) \cdot cos\left(2\pi(h/p)t + \phi(h/p)\right)$ is thus a periodic function with period p. Frequency is measured in cycles per t-unit; if t-units are seconds, then the frequency h/p denotes h/p cycles per second, or h/p *Hertz.*

We can also write

$$x(t) = \frac{A_0}{2} + \sum_{h=1}^{\infty} \left(A_h\, cos(2\pi(h/p)t) + B_h\, sin(2\pi(h/p)t)\right).$$

Exercise 1.4: Let x be a real period-p function with $x(t) = \sum_{h=\infty}^{\infty} C_h e^{2\pi i(h/p)t} = \sum_{h=0}^{\infty} M(h/p) \cdot cos\left(2\pi(h/p)t + \phi(h/p)\right)$. Show that, for $h > 0$,

$$C_h e^{2\pi i(h/p)t} + C_{-h} e^{-2\pi i(h/p)t} = M(h/p) \cdot cos\left(2\pi(h/p)t + \phi(h/p)\right) \text{ with } C_h^* = C_{-h}.$$

Solution 1.4: Let $C_h = \alpha_h + i\beta_h$ with $\alpha_h,\ \beta_h \in \mathcal{R}$. Then $\alpha_h = \alpha_{-h}$ and $\beta_h = -\beta_{-h}$ since x is real. Assume $C_h \neq 0$ and let we have $v = C_h e^{2\pi i(h/p)t} + C_{-h}e^{-2\pi i(h/p)t}$. Then $v = \alpha_h(e^{2\pi i(h/p)t} + e^{-2\pi i(h/p)t}) + i\beta_h(e^{2\pi i(h/p)t} - e^{-2\pi i(h/p)t}) = \alpha_h 2\, cos(2\pi(h/p)t) - \beta_h 2\, sin(2\pi(h/p)t)$, since $cos(\theta) = \frac{1}{2}\left[e^{i\theta} + e^{-i\theta}\right]$ and $sin(\theta) = -\frac{i}{2}\left[e^{i\theta} - e^{-i\theta}\right]$.

Now, $A_h := C_h + C_{-h} = 2\alpha_h$ and $B_h := i(C_h - C_{-h}) = -2\beta_h$, and so, $v = A_h\, cos(2\pi(h/p)t) + B_h\, sin(2\pi(h/p)t) = \left[A_h^2 + B_h^2\right]^{1/2}\left[\frac{A_h}{\left[A_h^2 + B_h^2\right]^{1/2}} \cdot cos(2\pi(h/p)t) + \frac{B_h}{\left[A_h^2 + B_h^2\right]^{1/2}} \cdot sin(2\pi(h/p)t)\right]$.

Let $\theta = \text{atan2}(-B_h, A_h) = \text{atan2}(\beta_h, \alpha_h)$. Then $\frac{A_h}{\left[A_h^2 + B_h^2\right]^{1/2}} = cos(\theta)$ and $\frac{B_h}{\left[A_h^2 + B_h^2\right]^{1/2}} = sin(-\theta) = -\,sin(\theta)$. Thus, $v = \left[A_h^2 + B_h^2\right]^{1/2}\left[cos(\theta)\, cos(2\pi(h/p)t) - sin(\theta)\, sin(2\pi(h/p)t)\right]$.

Also, for $h > 0$, $M(h/p) = \left[A_h^2 + B_h^2\right]^{1/2} = 2\left[\alpha_h^2 + \beta_h^2\right]^{1/2}$ and $\phi(h/p) = \theta$, so $v = M(h/p) \cdot cos\left(2\pi(h/p)t + \phi(h/p)\right)$.

And for $C_h = 0 = C_{-h}^*$ with $h > 0$, we have $A_0 = B_0 = 0$ and $\phi(0) = \text{atan2}(0, 0) = \frac{\pi}{2}$, and it is immediate that $0 = M(h/p) \cdot cos\left(2\pi(h/p)t + \phi(h/p)\right)$.

(Also, if $h = 0$, then $C_0 + C_0 = \left[A_h^2 + B_h^2\right]^{1/2} \cdot cos(\phi(0))$, but $M(0) = 2|C_0|/2$ and $cos(\phi(0)) = sign(C_0)$, so $C_0 e^{2\pi i(h/p)0} = C_0 = M(0)\, cos(\phi(0))$.)

When x is complex, the same spectral decomposition form applies. Michael O'Conner has shown that we can give extended definitions of M and ϕ, with M and ϕ depending upon h/p and an additional parameter ε, so that

$$\tilde{x}(t, \varepsilon) := \sum_{0 \le h \le \infty} M(h/p) \cdot cos\left(2\pi(h/p)t + \phi(h/p)\right)$$

uniformly approximates $x(t)$ a.e., *i.e.*, for $-\infty < t < \infty$, $|x(t) - \tilde{x}(t)| < 2\varepsilon$ a.e. for any chosen real value $\varepsilon > 0$. ("a.e." stands for "almost-everywhere", meaning everywhere, except possibly on a set of measure 0.) In order to have $\tilde{x}$ approximate x uniformly it suffices to define $M(h/p)$ and $\phi(h/p)$ so that, a.e.

$$\left| M(h/p) \cdot cos\left(2\pi(h/p)t + \phi(h/p)\right) - \left(C_h e^{2\pi i(h/p)t} + C_{-h}e^{-2\pi i(h/p)t}\right)\right| < \varepsilon/2^{|h|}.$$

In fact, except when exactly one of C_h and C_{-h} is zero, $M(h/p)$ and $\phi(h/p)$ will be defined so that

$$M(h/p) \cdot cos\left(2\pi(h/p)t + \phi(h/p)\right) = C_h e^{2\pi i(h/p)t} + C_{-h}e^{-2\pi i(h/p)t}.$$

We define

$$\phi(h/p) := \begin{cases} \text{atan2}(-B_h, A_h) & \text{if } A_h = 0 \text{ or } A_h \neq \pm iB_h, \\ \text{atan2}(-B_h - i\varepsilon/2^{|h|}, A_h + \varepsilon/2^{|h|}) & \text{if } 0 \neq A_h = iB_h, \text{ and} \\ \text{atan2}(-B_h + i\varepsilon/2^{|h|}, A_h + \varepsilon/2^{|h|}) & \text{if } 0 \neq A_h = -iB_h, \end{cases}$$

where

$$\text{atan2}(y,x) = \begin{cases} \pi/2 & \text{if } x = 0 \text{ and } \text{Re}(y) \geq 0, \\ -\pi/2 & \text{if } x = 0 \text{ and } \text{Re}(y) < 0, \\ tan^{-1}(y/x) & \text{if } x \neq 0 \text{ and } \text{Re}(x) \geq 0, \\ tan^{-1}(y/x) - \pi & \text{if } \text{Re}(x) < 0 \text{ and } 0 \leq \text{Re}(tan^{-1}(y/x)) < \pi/2, \\ tan^{-1}(y/x) + \pi & \text{otherwise,} \end{cases}$$

and where

$$tan^{-1}(z) = \frac{1}{2i} \log\left(\frac{i-z}{i+z}\right),$$

with $\text{Im}(\log(w)) \in [-\pi, \pi)$, and with $tan^{-1}(i) = \infty \cdot i$ and $tan^{-1}(-i) = (3/4)\pi - \infty i$.

Now when just one of C_h and C_{-h} is 0, we define

$$M(h/p) := \begin{cases} A_h/\left(cos(\phi(h/p)) \cdot (1 + \delta_{h,0})\right) & \text{when } cos(\phi(h/p)) \neq 0 \text{ and} \\ -B_h/\left(sin(\phi(h/p)) \cdot (1 + \delta_{h,0})\right) & \text{when } cos(\phi(h/p)) = 0, \end{cases}$$

and when both C_h and C_{-h} are non-zero, we define

$$M(h/p) := \left[C_h e^{2\pi i(h/p)t} + C_{-h} e^{-2\pi i(h/p)t}\right] / \, cos\left(2\pi(h/p)t + \phi(h/p)\right),$$

and when $C_h = C_{-h} = 0$, we define $M(h/p) := 0$.

These definitions of M, ϕ, and atan2 coincide with the definitions of M, ϕ and atan2 given above for the case where $x(t)$ is real.

Our purpose here is to introduce Fourier transforms and summarize some of their properties for impatient readers. It is not our purpose to properly and carefully justify every assertion; to do so would involve us in a thicket of technical issues: computing limits, establishing bounds, interchanging infinite sums and improper integrals, etc. Moreover, in some cases, the proximate arguments are too long, or depend on results that themselves require lengthy discussion to explain. These are not unimportant issues, but they are to be sought elsewhere [DM72], [Edw67].

[Major gaps:

1. Proof that $\sum_{h=-\infty}^{\infty} C_h e^{2\pi i(h/p)t}$ converges in the $L^2([0,p])$-norm if $\sum_j |C_j|^2$ converges.

2. Proof that $\{e^{2\pi i(j/p)t}\}$ is a countable approximating basis of $L^2([0,p])$.

3. Proof that $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(r) e^{-2\pi isr} \, dr \, e^{2\pi ist} \, ds$, exists and equals $x(t)$ a.e. when $x \in L^2(\mathcal{R})$.

]

2 Fourier Transforms

If $x(t) = \displaystyle\sum_{h=-\infty}^{\infty} C_h e^{2\pi i(h/p)t}$, then multiplying both sides by $e^{-2\pi i(j/p)t}$, and integrating along the real axis over one period from $-p/2$ to $p/2$ yields

$$\begin{aligned}\int_{-p/2}^{p/2} x(t)e^{-2\pi i(j/p)t}\,dt &= \int_{-p/2}^{p/2}\sum_{h=-\infty}^{\infty} C_h e^{2\pi i(h/p)t}\cdot e^{-2\pi i(j/p)t}\,dt\\ &= \sum_{h=-\infty}^{\infty} C_h \int_{-p/2}^{p/2} e^{2\pi i((h-j)/p)t}\,dt\\ &= \sum_{h=-\infty}^{\infty} C_h p\delta_{h,j}\\ &= pC_j,\end{aligned}$$

so $C_j = (1/p)\int_{-p/2}^{p/2} x(t)e^{-2\pi i(j/p)t}\,dt$. This is because the *orthogonality relation*

$$\int_{-p/2}^{p/2} e^{2\pi i(h/p)t}\cdot e^{-2\pi i(j/p)t}\,dt = \begin{cases} p & \text{if } h=j\\ 0 & \text{otherwise}\end{cases}$$

holds for $h, j \in \mathcal{Z}$, where $\mathcal{Z}$ denotes the set of integers, *i.e.*, the period-p functions $e^{2\pi i(h/p)t}\,dt$ and $e^{-2\pi i(j/p)t}\,dt$ are *orthogonal* to one-another when $h \neq j$. When $h = j$, we have $\int_{-p/2}^{p/2} 1dt = p$, and when $h \neq j$, we have

$$\begin{aligned}\int_{-p/2}^{p/2} e^{2\pi i((h-j)/p)t}\,dt &= (e^{2\pi i(h-j)t/p})/(2\pi i(h-j)/p)\Big|_{t=-p/2}^{t=p/2}\\ &= \left[e^{\pi i(h-j)} - e^{-\pi i(h-j)}\right]/(2\pi i(h-j)/p)\\ &= e^{-\pi i(h-j)}\left[e^{2\pi i(h-j)} - 1\right]/(2\pi i(h-j)/p)\\ &= \left[e^{2\pi i(h-j)} - 1\right]/\left[e^{\pi i(h-j)}(2\pi i(h-j)/p)\right]\\ &= [1-1]/\left[e^{\pi i(h-j)}(2\pi i(h-j)/p)\right]\\ &= 0.\end{aligned}$$

We thus have $x(t) = \sum_{h=-\infty}^{\infty}\left[\frac{1}{p}\int_{-p/2}^{p/2} x(r)e^{-2\pi i(h/p)r}\,dr\right]\cdot e^{2\pi i(h/p)t}$. (This holds when the integral and summation in the computation of pC_j above can be done in either order; this is the case when x has a Fourier series that converges in the least-squares sense.)

Note, since $e^{2\pi i(h/p)t}$ is a period-p periodic complex-valued function as well as a period-$p/|h|$ function for $h \in \mathcal{Z}$, $x(t) = \sum_{h=-\infty}^{\infty} C_h e^{2\pi i(h/p)t}$ is a period-p periodic complex-valued function defined for $-\infty < t < \infty$.

Lebesgue integration is employed throughout, since whenever a Riemann integral exists, the corresponding Lebesgue integral exists and has the same value, and the Lebesgue integral exists in cases where the Riemann integral does not [DM72]. Moreover, $\lim_{n\to\infty}\int_{-p/2}^{p/2} f_n = \int_{-p/2}^{p/2}\lim_{n\to\infty} f_n$ when $\lim_{n\to\infty} f_n$ converges, here $f_1, f_2, \ldots$ are measurable functions, and either $0 \le f_1 \le f_2 \le \cdots$ or $|f_n| \le g$ for $n = 1, 2, \ldots$ where g is an integrable function. (A real function f is measurable if the sets $\{t \mid a \le f(t) < b\}$ are measurable for all choices $a < b$. A complex function g is measurable if $\mathrm{Re}(g(t))$ and $\mathrm{Im}(g(t))$ are measurable functions.)

Exercise 2.1: Show that $\int_{-p/2}^{p/2} e^{2\pi i(h/p)t} e^{-2\pi i(j/p)t}\, dt = sin(\pi(h-j))/(\pi(h-j)/p)$ for $h, j \in \mathcal{R}$ with $h \neq j$.

Let $x(t)$ be a complex-valued periodic function of period p, defined on $-\infty < t < \infty$, that possesses a Fourier series expansion that converges in the least-squares sense: that is, x is *sufficiently-nice* so that $x(t) = \sum_{h=-\infty}^{\infty} C_h e^{2\pi i(h/p)t}$ where $\ldots, C_{-1}, C_0, C_1, \ldots$ are complex numbers such that $|C_h| \to 0$ as $|h| \to \infty$ fast enough so that $\sum_h |C_h|^2 < \infty$. The *Fourier transform* of x is:

$$x^\wedge(s) := (1/p)\int_{-p/2}^{p/2} x(t)e^{-2\pi i s t}\, dt,$$

for $s = \ldots,\ -2/p,\ -1/p,\ 0,\ 1/p,\ 2/p,\ \ldots$. The (period-$p$) Fourier transform of x is a discrete function that produces the Fourier coefficients of the period-p function x, *i.e.*, $x^\wedge(h/p) = C_h$ for $h = \ldots, -1, 0, 1, \ldots$. The Fourier transform of the period-p function x is a *regular* discrete function defined on the regular mesh $\ldots,\ -2/p,\ -1/p,\ 0,\ 1/p,\ 2/p,\ \ldots$ *i.e.*, a mesh with *stepsize* $1/p$. (Although named for Fourier, the Fourier transform is attributed to Pierre Laplace [Sti86].) You can think of $\wedge$ as an operator that, when applied to the function x, produces the function $(1/p)\int_{-p/2}^{p/2} x(t)e^{-2\pi i s t}\, dt$. For $h \in \mathcal{Z}$, the value $x^\wedge(h/p)$ is just the Fourier coefficient of x for the complex-oscillation term of frequence h/p. The period-p functions that have a Fourier transform thus include all those period-p functions that possess least-squares-convergent Fourier series expansions.

The *inverse Fourier transform* of $x^\wedge$ is:

$$x^{\wedge\vee}(t) := \sum_{h=-\infty}^{\infty} x^\wedge(h/p)e^{2\pi i(h/p)t} = x(t) \quad \text{a.e.}$$

This sum is the complex Fourier series of x. Indeed, we may define x to be a sufficiently-nice periodic function precisely when $x^{\wedge\vee}$ converges a.e. to x. Then the statement that a sufficiently-nice periodic function is equal a.e. to its Fourier series is the Fourier inversion theorem for a sufficiently-nice periodic function. The Fourier inversion theorem also shows that the Fourier coefficients C_h are uniquely determined by the sufficiently-nice function $x(t)$, in the sense that if any other sufficiently-nice function y has the same Fourier coefficients as x, then $y = x$ almost everywhere, and conversely.

The Fourier transform $x^\wedge$ of a sufficiently-nice period-p function x is restricted to the domain consisting of the multiples of $1/p$, and $x^\wedge(h/p)$ is the complex amplitude of the complex oscillation

$e^{2\pi i(h/p)t}$ of frequency h/p in the Fourier series for x. Thus $x^{\wedge\vee}$ is a sum of complex oscillations of frequencies $\ldots$, $-1/p$, 0, $1/p$, $2/p$, $\ldots$, which are multiples of the fundamental frequency $1/p$, where the complex values $\ldots$, $x^{\wedge}(-1/p)$, $x^{\wedge}(0)$, $x^{\wedge}(1/p)$, $x^{\wedge}(2/p)$, $\ldots$ are the associated complex amplitudes. Moreover the Fourier transform of a sufficiently-nice period-p function x defined on the real line is a discrete *decreasing* function defined at $\ldots$, $-1/p$, 0, $1/p$, $2/p$, $\ldots$, *i.e.*, with stepsize $1/p$; $1/p$ is the fundamental frequency of x. (A decreasing function $f(s)$ satisfies $|f(s)| \to 0$ as $|s| \to \infty$.)

We use the term "Fourier transform" carefully in conjunction with the term "Fourier integral transform", which is a distinct concept obtained by considering the Fourier transform integral for an arbitrary suitable integrable function, f, in the limit as $p \to \infty$, resulting in the integral $\displaystyle\int_{-\infty}^{\infty} f(t)e^{-2\pi ist}dt$ over the entire real line.

Note that the integral form for $x^{\wedge}(s)$ is computable when s is not an integer multiple of $1/p$. But, when s is not an integer multiple of $1/p$, $e^{-2\pi ist}$ is not of period p, so when s is an integral multiple of $1/q$ with $q \in \mathcal{R}^+$ rather than an integral multiple of $1/p$, we, in effect, are computing the value at s of an inner product of x extended with zero and $e^{-2\pi ist}$ in a space of period-q functions, and not in the space inhabited by periodic functions of period p. For example, given s, we could take the period q to be a positive integral multiple of $(\lfloor |s| \cdot p \rfloor + 1)/s$, The function of s given by the integral $\int_{-p/2}^{p/2} x(t)e^{-2\pi ist}\,dt$ defined for arbitrary values of s has another natural meaning as the *Fourier integral transform* of the function which coincides with x in the interval $[-p/2, p/2]$, and is zero outside. Hence, when we want to avoid such interpretations for the Fourier transform $x^{\wedge}$ of a period-p function x, we must take care to only compute $x^{\wedge}(s)$ for $s \in \{\ldots, -2/p, -1/p, 0, 1/p, 2/p, \ldots\}$; sometimes we may "forcibly" define $x^{\wedge}(s)$ to be zero at all points s, where s is not an integer multiple of $1/p$. Such a function, which we consider to be only defined on at most a countable set that has no Cauchy sequences as subsets, is said to have discrete support, and is called a *discrete function.* (In general the *support set* of a function x, defined on $\mathcal{R}$, is the set $\{t \in \mathcal{R} \mid x(t) \neq 0\}$.) Sometimes, however, we will want to compute $x^{\wedge}(s)$ where s is *not* necessarily a multiple of $1/p$; we will take care to identify these special situations.

A sufficiently-nice complex-valued periodic or almost-periodic function, x, has a Fourier transform which is a discrete decreasing function ($|x^{\wedge}(s)| \to 0$ as $|s| \to \infty$). The idea of expressing a periodic or almost-periodic function as a sum of oscillations can be applied to other kinds of functions as well. Indeed the extension of the concept of a Fourier transform to various domains of functions is a central theme of the theory of Fourier transforms. A discrete periodic function has a Fourier transform defined by a Riemann sum which results in another discrete periodic function; this is the discrete Fourier transform. A rapidly-decreasing (and therefore non-periodic) function, x, has a Fourier transform which is again a rapidly-decreasing function. This is the Fourier integral transform, $\int_{-\infty}^{\infty} x(t)e^{-2\pi ist}\,dt$, which is an integral over the entire real line. Finally, a merely polynomial-dominated measurable function has a Fourier-Stieljes transform which is the difference of two positive measures on $[-\infty, \infty]$ (or alternately, a linear functional on the space of rapidly-decreasing functions.)

All of these forms of Fourier transforms apply to different domains of functions, and as a function in one domain is approximated by a function in another, their respective Fourier transforms are related. In each case an "inverse transform" formula exists which recovers a function of appropriate type from its Fourier transform. These relations constitute some of the mathematical substance of the theory of Fourier transforms.

Some of the various Fourier transforms can be unified by introducing the so-called generalized functions, but this theory is difficult to appreciate, since periodic and discrete functions are not generalized functions, but are only represented by certain generalized functions and some generalized functions are not, in fact, functions, but merely the synthetic limits of sequences of progressively more sharply-varying functions; nevertheless this theory is computationally powerful and operationally easy to employ. It is best considered after studying each of the more elementary Fourier transforms separately.

Note that when x is periodic of period p and s is an integer multiple of $1/p$, $x(t)e^{-2\pi ist}$ is periodic of period p, so $x^\wedge(s)$ can be obtained by integrating over any interval of size p. Thus, for all a, $x^\wedge(s) = (1/p)\int_a^{a+p} x(t)e^{-2\pi ist}\,dt$ with s an integer multiple of $1/p$. Often we may prefer the interval $[0,p]$ instead of the symmetric interval $[-p/2, p/2]$, used earlier in the definition of the Fourier transform.

Occasionally, when, for example, we are simultaneously concerned with the Fourier transform of a function, x, of period p and a function, y, of period q, we shall use the notation $\wedge(p)$ and $\vee(p)$ to explicitly indicate the transform and inverse transform operators which involve the period parameter p, and which we usually apply to functions of period p; thus $x^{\wedge(p)\vee(p)} = x$.

Recall that a sufficiently-nice function is one where $x^{\wedge\vee}(t) = x(t)$, except possibly on a set of measure zero. A precise alternate characterization of sufficiently-nice is not known, but, for example, a function of bounded variation is sufficiently-nice. We shall generally write $f(t) = g(t)$, even though $f(t)$ may be different from $g(t)$ on a set of measure zero.

Knowing the Fourier transform of a sufficiently-nice periodic function, x, is equivalent to knowing its Fourier series; properties of the Fourier series of a periodic function correspond to related properties of the Fourier transform, $x^\wedge(s)$.

Exercise 2.2: Let $x(t)$ and $y(t)$ be real-valued continuous period-p periodic functions. The planar curve $c = \{(x(t), y(t)) \mid 0 \le t < p\}$ is a closed curve due to the periodicity of x and y.

Show that c is the "sum" of the ellipses

$$E_j = \{(M_x(j)\; cos(2\pi(j/p)t + \phi_x(j)), M_y(j)\; cos(2\pi(j/p)t + \phi_y(j))) \mid 0 \le t < p\}$$

where

$$\begin{aligned} M_x(j) &= (2-\delta_{j,0})|x^\wedge(j/p)|, \\ \phi_x(j) &= \text{atan2}(\text{Im}(x^\wedge(j/p), \text{Re}(x^\wedge(j/p))), \\ M_y(j) &= (2-\delta_{j,0})|y^\wedge(j/p)|, \text{ and} \\ \phi_y(j) &= \text{atan2}(\text{Im}(y^\wedge(j/p), \text{Re}(y^\wedge(j/p))), \end{aligned}$$

for $j = 0, 1, 2, \ldots$, in the following sense.

Define $\tilde{c}[t] = (x(t), y(t))$ and define $\tilde{E}_j[t] = (\, M_x(j)\, cos(2\pi(j/p)t + \phi_x(j)),\, M_y(j)\, cos(2\pi(j/p)t + \phi_y(j))\,)$.

Then show that $\tilde{c}[t] = \tilde{E}_0[t] + \tilde{E}_1[t] + \cdots$. Also show that $E_0 = \{\tilde{E}_0[0]\}$ and the set $E_j = \{\tilde{E}_j[t] \mid 0 \le t < p/j\}$ for $j = 1, 2, \ldots$. (Note E_j, as a multiset, consists of several "circuits" of an ellipse for $j > 1$.)

2.1 Geometric Interpretation

On a subset of the sufficiently-nice period-p functions, the Fourier transform is elegantly interpreted as a unitary linear transformation on an infinite-dimensional inner-product vector space; such a space is called a *Hilbert* space.

Let Q be the interval $[0, p]$, and let $L^2(Q)$ be the set of complex-valued functions, x, defined on Q such that $\int_0^p |x(t)|^2\, dt < \infty$.

Define the inner product in $L^2(Q)$ as:

$$(x, y) := \int_0^p x(t)y(t)^*\, dt$$

and define the norm as:

$$\|x\| := (x, x)^{1/2}.$$

When necessary, we shall write $(x, y)_{L^2(Q)}$ and $\|x\|_{L^2(Q)}$ to precisely specify this inner product and norm on $L^2(Q)$.

Both Schwarz's inequality: $\|(x, y)\| \le \|x\| \cdot \|y\|$, and the triangle inequality (also known as Minkowski's inequality): $\|x + y\| \le \|x\| + \|y\|$, hold for $x, y \in L^2(Q)$.

$L^2(Q)$ is an infinite-dimensional Hilbert space over the complex numbers, $\mathcal{C}$. $L^2(Q)$ is also a metric space with the distance function $\|x - y\|$ for $x, y \in L^2(Q)$, and $L^2(Q)$ is complete (Cauchy sequences of functions in $L^2(Q)$ converge to functions in $L^2(Q)$ with respect to the just-specified metric) and separable (any function in $L^2(Q)$ is arbitrarily close to a function in a distinguished countable subset D of $L^2(Q)$, *i.e.*, D is dense in $L^2(Q)$.) Separability is, in some sense, the most crucial property of $L^2(Q)$ - it guarantees that a kind of basis set for $L^2(Q)$ exists which is "non-trival" in that its cardinality is less than the cardinality of $L^2(Q)$ itself, given that we accomodate the representation of functions in $L^2(Q)$ with respect to such a basis set by including limits (in norm) of sequences of finite linear combinations of basis functions, as well as the finite combinations themselves. Such a dense countable set of functions, D, is called an *approximating basis* of $L^2(Q)$. In actuality, the Fourier basis of period-p complex exponentials was seen to be an approximating basis for a class of functions that was later determined to be the $L^2(Q)$-functions, so that $L^2(Q)$ was seen to be separable a priori.

Let the functions $\ldots$, e_{-2}, e_{-1}, e_0, e_1, e_2, $\ldots$ form an *orthogonal* approximating basis for $L^2(Q)$. This means that for $x \in L^2(Q)$, there exist complex numbers $\ldots$, C_{-2}, C_{-1}, C_0, C_1, C_2, $\ldots$ such that $\lim_{k\to\infty} \left\| x - \sum_{j=-k}^{k} C_j e_j \right\| = 0$; we write $x = \sum_j C_j e_j$ to indicate this. Note however, x is not, in general, equal to a linear combination of a finite number of the orthogonal basis functions, but is approximated arbitrarily closely a.e. by such linear combinations; this is the notion of an orthogonal approximating basis, rather than a basis in the strict vector-space sense. (The infinite-dimensional vector space $L^2(Q)$ does not possess a strict basis, since infinite sums must be accomodated.) Having an orthogonal approximating basis means that

$$(e_j, e_k) = \begin{cases} \|e_j\|^2 & \text{if } j = k \\ 0 & \text{otherwise.} \end{cases}$$

$L^2(Q)$ has other orthogonal approximqting bases, and $L^2(Q)$ also has numerous non-orthogonal approximating bases, $\langle \ldots, \phi_{-2}, \phi_{-1}, \phi_0, \phi_1, \ldots \rangle$; however, with such a basis we cannot take a *fixed* sequence of "coordinate components $\ldots, C_{-2}, C_{-1}, C_0, C_1, \ldots$ as defining $x \in L^2(Q)$, but can only claim that the finite linear combinations of $\ldots, \phi_{-2}, \phi_{-1}, \phi_0, \phi_1, \ldots$ form a dense subset of $L^2(Q)$ whose completion is $L^2(Q)$.

An orthogonal approximating basis can always be constructed from an arbitrary approximating basis by means of the Gram-Schmidt process. The function $(x, e_j/\|e_j\|)e_j/\|e_j\|$ is the orthogonal projection of x in the direction e_j, of length $(x, e_j/\|e_j\|)$.

$C_j = (x, e_j)/\|e_j\|^2$ is called the j-th *Fourier coefficient* of x, and $x = \sum_j C_j e_j$. (C_j is also called the j-th *component* of x with respect to the Fourier basis $\langle \ldots, e_{-1}/\|e_{-1}\|, e_0/\|e_0\|, e_1/\|e_1\|, \ldots \rangle$.)

If we also have $y = \sum_k D_k e_k$, then note that, due to the orthogonality of the approximating basis functions,

$$\begin{aligned} (x, y) &= \left(\sum_j C_j e_j, \sum_k D_k e_k \right) &= \int_Q \sum_j \sum_k C_j e_j D_k^* e_k^* \\ &= \sum_j \sum_k C_j D_k^* (e_j, e_k) &= \sum_j C_j D_j^* \|e_j\|^2. \end{aligned}$$

This is known as Parseval's identity. As a special case we have $\|x\|^2 = \sum_j |C_j|^2 \|e_j\|^2$; this is usually called Plancherel's identity; it is essentially the Pythagorean theorem in infinite-dimensional space. These identities are variously labeled with the names of Rayleigh, Parseval and Plancherel.

Let $\mathcal{Z}$ denote the integers $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$, and let $\mathcal{Z}/p$ denote the numbers $\{\ldots, -2/p, -1/p, 0, 1/p, 2/p, \ldots\}$, where p is a positive real number. Let $L^2(\mathcal{Z}/p)$ be the set of complex-valued functions, f, on $\mathcal{Z}/p$ such that $\sum_h |f(h/p)|^2 < \infty$. Introduce the inner product $(f, g) = p \sum_h f(h/p) g(h/p)^*$ for $f, g \in L^2(\mathcal{Z}/p)$, and the associated norm $\|f\| = (f, f)^{1/2}$. With this norm, $L^2(\mathcal{Z}/p)$ is a complete separable infinite-dimensional Hilbert space over the complex numbers, $\mathcal{C}$. When necessary, we shall write $(f, g)_{L^2(\mathcal{Z}/p)}$ and $\|f\|_{L^2(\mathcal{Z}/p)}$ to denote this inner product and norm on $L^2(\mathcal{Z}/p)$.

Now fix $e_j(t) = e^{2\pi i (j/p) t}$. Note $\|e_j\| = p^{1/2}$. These complex exponential functions form a particular orthogonal approximating basis for $L^2(Q)$ (this is proven in [DM72].) The mapping $x \to x^\wedge$, where $x \in L^2(Q)$ and $x^\wedge \in L^2(\mathcal{Z}/p)$, defined by

$$x^\wedge(j/p) := (x, e_j/\|e_j\|^2) = (1/p) \int_Q x(t) e_j(t)^* \, dt = (1/p) \int_Q x(t) e^{-2\pi i (j/p) t} \, dt$$

is a one-to-one linear transformation of $L^2(Q)$ into $L^2(\mathcal{Z}/p)$. Note $\wedge$ is a linear operator: $(\alpha x + \beta y)^\wedge = \alpha x^\wedge + \beta y^\wedge$. The Reisz-Fischer theorem states that, in fact, $\wedge$ maps $L^2(Q)$ *onto* $L^2(\mathcal{Z}/p)$. The discrete function $x^\wedge$ is the Fourier transform of x; it is just the representation of x in the coordinate system given by the orthonormal approximating basis $\langle \ldots, e_{-1}/\|e_{-1}\|, e_0/\|e_0\|, e_1/\|e_1\|, \ldots\rangle$. In fact, the mapping $\wedge$ is an isomorphism between $L^2(Q)$ and $L^2(\mathcal{Z}/p)$ as Hilbert spaces. In particular, inner products are preserved: $(x,y)_{L^2(Q)} = (x^\wedge, y^\wedge)_{L^2(\mathcal{Z}/p)}$ (Parseval's identity), and hence lengths defined by the respective norms in $L^2(Q)$ and $L^2(\mathcal{Z}/p)$ are preserved also: $\|x\|_{L^2(Q)} = \|x^\wedge\|_{L^2(\mathcal{Z}/p)}$ (Plancherel's identity). An inner-product-preserving one-to-one linear transformation is a unitary transformation, and since no reflection is involved, the mapping $\wedge$ can be characterized as a kind of rotation of $L^2(Q)$ onto $L^2(\mathcal{Z}/p)$ which is, since $\wedge$ is an isomorphism, just a representation of $L^2(Q)$ coordinatized with respect to the orthonormal approximating basis $\langle \ldots, e_{-1}/\|e_{-1}\|, e_0/\|e_0\|, e_1/\|e_1\|, \ldots\rangle$.

Note that *every* separable infinite-dimensional Hilbert space is isomorphic to $L^2(\mathcal{Z}/p)$, since a countable orthonormal approximating basis is guaranteed to exist, and expressing an element, x, in terms of this orthonormal approximating basis yields a coefficient sequence in $L^2(\mathcal{Z}/p)$ which then corresponds to x.

Since $\wedge$ is an isomorphism between $L^2(Q)$ and $L^2(\mathcal{Z}/p)$, the inverse mapping $\vee$ is also an isomorphism between $L^2(\mathcal{Z}/p)$ and $L^2(Q)$.

Note that if we were to choose $e_j(t) = e^{2\pi i (j/p)t}/\sqrt{p}$ then the factor $1/p$ in the Fourier transform integral would be redistributed into both the operators $\wedge(p)$ and $\vee(p)$ equally. Also note that there are many choices for the approximating basis functions $\ldots, e_{-1}, e_0, e_1, \ldots$, and for each such choice we have an associated generalized Fourier series expansion for the functions in $L^2(Q)$. For certain choices, we obtain classical orthogonal polynomial expansions, and for other choices, we obtain particular so-called *wavelet* expansions.

2.2 Almost-Periodic Functions

A larger class of functions than $\cup_{0<p<\infty} L^2(Q)$ admit a trigonometric series representation. This is the class, A, of almost-periodic functions. A complex-valued function, x, defined on the real line is almost-periodic if, for $\varepsilon > 0$, there exists $p > 0$ such that for all $t \in [0,p]$, $|x(t+a) - x(t)| < \varepsilon$ for at least one point a in every interval of length p. Such functions have a Fourier transform, $x^\wedge(s) = (1/p)\int_{-\infty}^{\infty} x(t)e^{-2\pi i s t}\,dt$, which differs from zero on an, at most countable, discrete set of points $\ldots, s_{-2}, s_{-1}, s_0, s_1, s_2, \ldots$. The inverse Fourier transform of $x^\wedge$ is

$$x^{\wedge\vee}(t) = \int_{s\in(-\infty,\infty)} x^\wedge(s)e^{2\pi i s t}\,dS(x^\wedge) = \sum_{h=-\infty}^{\infty} x^\wedge(s_h)e^{2\pi i s_h t},$$

where $S(x^\wedge)$ is the Dirac comb measure on the support set of $x^\wedge$, where

$$S(x^\wedge)[\{\alpha\}] = \begin{cases} 1 & \text{if } \alpha \in \{\ldots, s_{-1}, s_0, s_1, \ldots\} \\ 0 & \text{otherwise,} \end{cases}$$

and, in general for $U \subseteq \mathcal{R}$, $S(x^\wedge)[U] = card(\{U \cap \{\ldots, s_{-1}, s_0, s_1, \ldots\})$.

The class of functions A, when completed by adding all missing limit points, is a Hilbert space, but it is not separable, so no countable basis exists. In spite of this, each element x has a unique countable decomposition as specified above.

The inner product in A is

$$(x,y)_A := \lim_{p\to\infty}(1/p)\int_{-p/2}^{p/2} x(t)y(t)^*\,dt$$

and the norm $\|x\|_A := (x,x)_A^{1/2}$.

Plancherel's identity holds. For $x \in A : \|x\|_A^2 = \sum_{h=-\infty}^{\infty} |x^\wedge(s_h)|^2$, where $\ldots$, s_{-1}, s_0, s_1, $\ldots$ are the points at which $|x^\wedge(s)| > 0$.

2.3 Convergence

Let $S_k(t) = \sum_{h=-k}^{k} x^\wedge(h/p)e^{2\pi i(h/p)t}$. S_k is the k-th partial sum of $x^{\wedge\vee}$. For $x \in L^2(Q)$, the k-th partial sum of the Fourier series of $x(t)$, $S_k(t)$, has the property that $\|x - S_k\| \le \|x - A_k\|$, for any choice of coefficients a_{-k}, $\ldots$, a_{-1}, a_0, a_1, $\ldots$, a_k, where $A_k(t) = \sum_{j=-k}^{k} a_j e^{2\pi i(j/p)t}$.

Also, we have Bessel's inequality: $\|x\| \ge \|S_k\|$, so that $\langle S_k\rangle$ is a Cauchy sequence, and hence $x^{\wedge\vee}$ converges in the $L^2(Q)$ norm. Since each member $x^\wedge$, of $L^2(\mathcal{Z}/p)$, corresponds to $x^{\wedge\vee}$ in $L^2(Q)$, we see that $x^{\wedge\vee}$ converges in the $L^2(Q)$ norm when $\sum_{h=-\infty}^{\infty} |x^\wedge(h/p)|^2 < \infty$.

> **Exercise 2.3:** Is the Fourier series of x unique in the sense that no other sum of complex oscillations, no matter what their frequencies, can converge to $x^{\wedge\vee}$?

If $x^{\wedge\vee}$ converges at the point t, then $x^{\wedge\vee}(t) = x(t)$, unless x is discontinuous at t. In general, if $x^{\wedge\vee}$ converges at t, then $x^{\wedge\vee}(t) = (x(t-) + x(t+))/2$ where $x(t+)$ denotes $\lim_{\varepsilon\downarrow 0} x(t+\varepsilon)$ and $x(t-)$ denotes $\lim_{\varepsilon\downarrow 0} x(t-\varepsilon)$. In particular, if x has an isolated jump discontinuity at t, then $x^{\wedge\vee}(t)$ lies halfway between $x(t-)$ and $x(t+)$.

> **Exercise 2.4:** How can the sum of an infinite number of continuous functions converge to a discontinuous function?

If x is n-fold continuously-differentiable for $n \ge 0$ (*i.e.*, $x = x^{(0)}$, $x^{(1)}$, $\ldots$, $x^{(n)}$ exist, and $x^{(1)}$, $\ldots$, $x^{(n-1)}$ are continuous and periodic, where $x^{(j)}$ denotes the j-th derivative of x) and $x^{\wedge\vee}$ converges pointwise almost everywhere in Q then $x^\wedge(h/p) = o(1/|h|^{n+1})$ or more specifically, $|x^\wedge(h/p)| \le \alpha(1+|h|)^{-(n+1)}$, where α is a fixed real constant value. In particular, if $x \in L^2(Q)$ is continuous, then $x^{\wedge\vee}$ converges uniformly to x almost everywhere in Q.

In general, the smoother x is (*i.e.*, the more continuous derivatives x has,) the more rapidly its Fourier coefficients approach 0. The converse is also true; if $|x^\wedge(h/p)| \le \alpha(1+|h|)^{-(n+2)}$ for all $h \in \mathcal{Z}$ and some constant α, then x is n-fold continuously-differentiable.

Suppose x has an isolated jump discontinuity at t_0. Then

$$\lim_{k\to\infty} S_k(t_0 + p/(4k)) \quad = \quad x(t_0+) + d(x(t_0+) - x(t_0-)) \quad \text{and}$$

$$\lim_{k\to\infty} S_k(t_0 - p/(4k)) \quad = \quad x(t_0-) + d(x(t_0-) - x(t_0+)),$$

where $d = \dfrac{2}{\pi}\displaystyle\int_0^\pi (sin(u)/u)\,du - 1 = .0894899\ldots$. This $8.94\cdots$-percent overshoot/undershoot on each side of a jump discontinuity is known as Gibbs' phenomenon. Convergence of $\langle S_k \rangle$ is never uniform in the neighborhood of such a discontinuity.

Exercise 2.5: What is $\dfrac{1}{2}\lim\limits_{k\to\infty}[S_k(t_0 + p/(4k)) + S_k(t_0 - p/(4k))]$?

Let M_n denote an increasing mesh $0 \le t_1 < t_2 < \ldots < t_n \le p$. The *variation* of x on Q is

$$V_Q(x) := \lim_{n\to\infty} \sup_{M_n} \sum_{j=1}^{n} |x(t_j) - x(t_j - 1)|$$

$V_Q(x)$ is a measure of the length of the "curve" which is the graph of $|x|$ on Q. The function x is of bounded variation, *i.e.*, $V_Q(x) < \infty$, if and only if both $\mathrm{Re}(x)$ and $\mathrm{Im}(x)$ are separately expressible as the difference of two bounded increasing functions.

If x is of bounded variation, then $x^{\wedge\vee}$ converges pointwise almost everywhere in Q. Moreover the partial sums, $S_k(t) = \sum_{h=-k}^{k} x^\wedge(h/p)e^{2\pi i(h/p)t}$ satisfy $|S_k| < c_1 \cdot V_Q(x) + c_2$ where c_1 and c_2 are constants. In 1966, Lennart Carleson published a paper proving that if $x \in L^2(Q)$, then $x^{\wedge\vee}$ converges pointwise almost everywhere, even if x is not of bounded variation. [Edw67]

In fact, the Fourier transform is defined on the larger space, $L^1(Q)$, of period-p functions, x, such that $\int_Q |x| < \infty$; moreover, if a period-p function x has a Fourier transform $x^\wedge$, then $x \in L^1(Q)$; thus $L^1(Q)$ is the proper "maximal" domain for $\wedge$ within the class of period-p periodic functions. But the corresponding inverse transform $x^{\wedge\vee}$ may not converge pointwise a.e. to x. The Fourier series for a function $x \in L^1(Q)$ is, however, Féjer-summable to x. A series $\sum_j a_j(t)$ is Féjer-summable to the function $x(t)$ if

$$\lim_{k\to\infty} \frac{1}{k} \sum_{0\le n<k} \sum_{-n\le j\le n} a_j(t) = x(t).$$

As just mentioned, there are decreasing discrete complex-valued functions f, defined on $\mathcal{Z}/p$, that are the Fourier transforms of functions in $L^1(Q)$, whose inverse Fourier transforms $f^\vee$ do not belong to $L^1(Q)$. For example,

$$f(h/p) = \begin{cases} 0 & \text{if } |h| \le 1, \\ -i \cdot sign(h)/(2\log|h/p|) & \text{otherwise.} \end{cases}$$

Such functions f have $|f^\vee| = \infty$, essentially because f does not decrease fast enough. Nevertheless, in a certain sense, f has an inverse Fourier transform $f^\vee$, which is just not computable by the usual recipe.

Let $S(Q)$ denote the class of sufficiently-nice period-p functions where $x \in S(Q)$ implies $x^{\wedge\vee} = x$. Then $L^2(Q) \subset S(Q) \subset L^1(Q)$. Let $S(\mathcal{Z}/p) = \{y^\wedge \mid y \in S(Q)\}$, let $G(Q)$ be the class of period-p functions which are the pointwise sums of at least conditionally-convergent series of the form

$\sum_j f(j/p)e^{2\pi i(j/p)t}$, and let $G(\mathcal{Z}/p)$ be the corresponding class of discrete functions, f, on $\mathcal{Z}/p$. Then $S(Q) \subset G(Q)$ and $S(\mathcal{Z}/p) \subset G(\mathcal{Z}/p)$.

The Fourier transform $\wedge$ is an invertable linear map from $L^2(Q)$ onto $L^2(\mathcal{Z}/p)$. We can enlarge the domain of $\wedge$ to the set $S(Q)$ with the range $S(\mathcal{Z}/p)$, keeping $\wedge$ invertable. We can further enlarge the domain of $\wedge$ from $S(Q)$ to the set $L^1(Q)$, but the added functions do not possess inverse Fourier transforms (*i.e.*, do not have pointwise-convergent a.e. Fourier series.)

Similarly, we can enlarge the domain of $\vee$, $S(\mathcal{Z}/p)$, to the set $G(\mathcal{Z}/p)$, but the added functions define Fourier series which, although convergent, do not possess Fourier transforms. In other words, $\vee : G(\mathcal{Z}/p) \to G(Q)$, but there are functions in $G(Q)$ that cannot be mapped back to $G(\mathcal{Z}/p)$ by the Fourier transform recipe. Note that $S(Q) = L^1(Q) \cap G(Q)$.

Finally, there is the class $J(Q)$ of period-p functions, which are Féjer sums of series of the form $\sum_j f(j/p)e^{2\pi i(j/p)t}$, and the associated class $J(\mathcal{Z}/p)$ of discrete functions f for which we have $L^1(Q) \subset J(Q)$ and $G(Q) \subset J(Q)$. There are no known simple descriptive characterizations of the sets $S(Q), G(Q)$, and $J(Q)$, or of the sets $S(\mathcal{Z}/p), G(\mathcal{Z}/p)$, and $J(\mathcal{Z}/p)$.

2.4 Structural Relations

- $x^\wedge(0) = (1/p)\int_{-p/2}^{p/2} x(t)\,dt =$ the mean value of x on $[-p/2, p/2]$. $x^\wedge(0)$ is called the d.c. (direct current) component of x.
- $x^\wedge$ has discrete support, with $x^\wedge(s)$ defined when $s = h/p$ where h is an integer.
- The operators $\wedge$ and $\vee$ are linear operators, so that

$$\begin{aligned}(ax(t) + by(t))^\wedge(s) &= ax^\wedge(s) + by^\wedge(s) \quad \text{and} \\ (ax^\wedge(s) + by^\wedge(s))^\vee &= ax^{\wedge\vee}(t) + by^{\wedge\vee}(t).\end{aligned}$$

- x is real if and only if $x^\wedge(-s) = x^\wedge(s)^*$ where a^* denotesw the complex conjugate of a, *i.e.*, $x^\wedge$ is *Hermitian*. (As a matter of notation, we may write $y^*(s)$ to mean the same thing as $y(s)^*$; both expressions denote the application of the complex-conjugate operation to the output of the function y.) We write $x^R(s)$ to denote $x(-s)$, where "R" denotes the *reversal* operation (the term "reflection" is also used.) Then to say $x^\wedge$ is Hermitian is to say x is real and $(x^\wedge)^R = x^{\wedge R} = x^{\wedge *}$.

Exercise 2.6: Show that, for $x \in L^2(Q)$, $x = x^*$, if and only if $x^{\wedge R} = x^{\wedge *}$.

Solution 2.6:

$$\begin{aligned}x^{\wedge *}(s) &= \left[\frac{1}{p}\int_{-p/2}^{p/2} x(t)e^{-2\pi i s t}\,dt\right]^* \\ &= \frac{1}{p}\int_{-p/2}^{p/2} x^*(t)e^{2\pi i s t}\,dt \\ &= x^{*\wedge}(-s) \\ &= x^{*\wedge R}(s).\end{aligned}$$

And,

$$x^{\wedge R}(s) = \left[\frac{1}{p}\int_{-p/2}^{p/2} x(t)e^{-2\pi ist}\,dt\right]^R = \frac{1}{p}\int_{-p/2}^{p/2} x(t)e^{2\pi ist}\,dt.$$

Thus, if $x = x^*$ then $x^{\wedge R} = x^{\wedge *} = x^{*\wedge R}$. Conversely, if $x^{\wedge *} = x^{\wedge R}$ then $x^{\wedge *} = x^{*\wedge R} = x^{\wedge R}$, and thus $x^{*\wedge RR\vee} = x^{\wedge RR\vee}$ which reduces to $x^* = x$.

- The operator pairs $(*, R), (\wedge, R)$, and $(\vee, R)$ commute but $(\vee, *)$ and $(\wedge, *)$ do not. Thus $x^{*R} = x^{R*}, x^{R\wedge} = x^{\wedge R}$, and $f^{R\vee} = f^{\vee R}$.

  Exercise 2.7: Show that $x^{\wedge}(-s) = \int_0^p x(t)e^{2\pi ist}\,dt = \int_0^p x(-t)e^{-2\pi ist}\,dt = (x(-t))^{\wedge}(s)$.

  Exercise 2.8: Is the complex-conjugate operator $*$ a linear operator? Hint: it depends on what the field of scalars is taken to be for our Hilbert space of periodic functions.

- Note if $x^{\wedge}$ is Hermitian then $\mathrm{Re}(x^{\wedge})$ is even and $\mathrm{Im}(x^{\wedge})$ is odd, where a function, y, is *even* if $y(t) = y(-t)$, and *odd* if $y(t) = -y(-t)$. Every function x can be decomposed into its even part, $\mathrm{even}(x)(t) = (x(t)+x(-t))/2$, and its odd part, $\mathrm{odd}(x)(t) = (x(t)-x(-t))/2$, such that $\mathrm{even}(x)$ is even, $\mathrm{odd}(x)$ is odd, and $x = \mathrm{even}(x) + \mathrm{odd}(x)$.

- x is odd (*i.e.*, $x(t) = -x(-t)$,) if and only if $x^{\wedge}$ is odd. x is even (*i.e.*, $x(t) = x(-t)$,) if and only if $x^{\wedge}$ is even. Thus, if x is real and even, or imaginary and odd, $x^{\wedge}$ must be real.

- The product of two even functions is even, and so is the product of two odd functions, but the product of an even function and an odd function is odd. Using this fact, together with Euler's relation $e^{i\theta} = cos(\theta) + i\,sin(\theta)$, and the fact that $cos(\theta)$ is even and $sin(\theta)$ is odd, we can state the following.

  If $x \in L^2(Q)$ is even, $x^{\wedge}$ is the *cosine* transform of x with

$$x^{\wedge}(h/p) = (1/p)\int_{-p/2}^{p/2} x(t)\,cos(2\pi(h/p)t)\,dt \quad \text{and}$$
$$x(t) = x^{\wedge}(0) + \sum_{h\geq 1} 2x^{\wedge}(h/p)\,cos(2\pi(h/p)t).$$

  If $x \in L^2(Q)$ is odd, $x^{\wedge}$ is the *sine* transform of x with

$$x^{\wedge}(h/p) = (1/p)\int_{-p/2}^{p/2} x(t)\,sin(-2\pi(h/p)t)\,dt \quad \text{and}$$
$$x(t) = x^{\wedge}(0) + \sum_{h\geq 1} 2x^{\wedge}(h/p)\,sin(2\pi(h/p)t).$$

  The development of a trigonometric Fourier series for a real-valued function, x, can be done without resorting to complex numbers via the cosine transform of the even part of x and the sine transform of the odd part of x.

- Note that when x is real, $x^\wedge(s) = (M(s)/(2-\delta_{s,0}))e^{i\phi(s)}$, and moreover M is an even function and ϕ is an odd function.

- $(x(at))^{\wedge(p/|a|)}(s) = x^{\wedge(p)}(s/a)$ for $s \in \{\ldots, -\frac{|a|}{p}, 0, \frac{|a|}{p}, \ldots\}$. Note $x(at)$ has period $p/|a|$ when $x(t)$ has period p.

  Exercise 2.9: Show that $(x(at))^{\wedge(p/|a|)}(s) = x^{\wedge(p)}(s/a)$ for $s \in \{\ldots, -\frac{|a|}{p}, 0, \frac{|a|}{p}, \ldots\}$.

  Solution 2.9: Let $y(t) = x(at)$. Since x has period p, y has period $p/|a|$.
  Now $y^{\wedge(p/|a|)}(s) = \frac{|a|}{p}\int_0^{\frac{p}{|a|}} x(at)e^{-2\pi ist}\,dt$ for $s \in \{\ldots, -\frac{|a|}{p}, 0, \frac{|a|}{p}, \ldots\}$. Let $r = at$, then $y^{\wedge(p/|a|)}(s) = \frac{|a|}{p}\int_0^{sign(a)p} \frac{1}{a}x(r)e^{-2\pi isr/a}\,dr = \frac{sign(s)^2}{p}\int_0^p x(r)e^{-2\pi i(s/a)r}\,dr = x^{\wedge(p)}(s/a)$.

- $(x(t+b))^\wedge(s) = e^{2\pi ibs}x^\wedge(s)$. Also, for k an integer, $(e^{-2\pi i(k/p)t}x(t))^\wedge(s) = x^\wedge(s+k/p)$.

- For x of period p and for k a positive integer, $x^{\wedge(kp)} = kx^{\wedge(p)}$.

- $(x')^\wedge(s) = 2\pi is \cdot x^\wedge(s)$ for $s \in \{\ldots, -1/p, 0, 1/p, \ldots\}$, where x' denotes the derivative of x.

  Exercise 2.10: Show that $(x')^\wedge(k/p) = 2\pi i(k/p)\cdot x^\wedge(k/p)$ for $k \in \mathcal{Z}$. Hint: write the Fourier series for $x(t)$, differentiate term-by-term with respect to t, and use the orthogonality relation for complex exponentials to extract $(x')^\wedge(k/p)$.

  Solution 2.10: Integrating "by parts", we have $(x')^\wedge(k/p) = \int_{-p/2}^{p/2} x'(t)e^{-2\pi itk/p}dt = x(t)\left[e^{-2\pi itk/p}\right]_{t=-p/2}^{t=p/2} - \int_{-p/2}^{p/2} x(t)\left[e^{-2\pi itk/p}\right]' dt = -\int_{-p/2}^{p/2} x(t)(-2\pi ik/p)e^{-2\pi itk/p}dt = 2\pi i(k/p)\cdot x^\wedge(k/p)$ for $k \in \mathcal{Z}$.

- Let B be a basis for $\mathcal{C}^n$ ($\mathcal{C}$ is the set of complex numbers and $\mathcal{C}^n$ is the n-dimensional vector space of n-tuples of complex numbers over the field of scalars $\mathcal{C}$.) Suppose the $n \times n$ $\langle B, B\rangle$-matrix D is diagonal. Recall that the application of the linear transformation given by the diagonal matrix D to a vector x just multiplies each component of the vector x by a constant, namely $x_j \to D_{jj}x_j$ for $j = 1, 2, \ldots, n$.

  Now note that expressing an $L^2(Q)$-function x as its Fourier series is equivalent to representing x with respect to the countably-infinite Fourier basis $\langle\ldots, e^{2\pi i(-1/p)t}/\sqrt{p}, 1/\sqrt{p}, e^{2\pi i(1/p)t}/\sqrt{p}, \ldots\rangle$. The j-th component of x in the Fourier basis is just $x^\wedge(j/p)$.

  By analogy with the finite-dimensional case, we can say that, in the Fourier basis, the differentiation linear transformation is "diagonal"; the j-th component of x' in the Fourier basis is just the j-th component of x multiplied by the constant $2\pi i(j/p)$.

2.5 Circular Convolution

Define the circular p-convolution function $x \circledast y$ for complex-valued functions x and y by

$$(x \circledast y)(r) := (1/p)\int_0^p x(t)y(r-t)\,dt.$$

The circular p-convolution $x \circledast y$ is a periodic period-p function when x and y are period-p functions in $L^2(Q)$. For functions in $L^2(Q)$, circular p-convolution is commutative, associative, and distributive: $x \circledast y = y \circledast x$, $x \circledast (y \circledast z) = (x \circledast y) \circledast z$, and $x \circledast (y+z) = x \circledast y + x \circledast z$. Also $x \circledast y = x^R \circledast y^R$, and we have the fundamental identities:

$$(x \circledast y)^\wedge = x^\wedge y^\wedge \quad \text{and} \quad (x^\wedge y^\wedge)^\vee = x \circledast y,$$

whenever all the integrals involved exist.

The term "circular" emphasizes the periodicity of x and y. We may just say "circular convolution", leaving-out the period parameter p which is then understood to exist without explicit mention.

Exercise 2.11: Show that $(x \circledast y) = (y \circledast x)$. Hint: hold r constant and change variables: $t \to r - s$.

Solution 2.11: $(x \circledast y)(r) = \frac{1}{p}\int_0^p x(t)y(r-t)\,dt = \frac{1}{p}\int_r^{r-p} x(r-s)y(s)(-1)\,ds = \frac{1}{p}\int_{r-p}^{r} x(r-s)y(s)\,ds = \frac{1}{p}\int_0^p x(r-s)y(s)\,ds = (y \circledast x)(r)$.

Exercise 2.12: Show that $(x \circledast y)^\wedge = x^\wedge y^\wedge$ for $x, y \in L^2(Q)$.

Solution 2.12:

$$\begin{aligned}
(x \circledast y)^\wedge(h/p) &= \frac{1}{p}\int_{r\in Q}\left[\frac{1}{p}\int_{t\in Q} x(t)y(r-t)\right] e^{-2\pi i(h/p)r} \\
&= \frac{1}{p}\int_{r\in Q}\left[\frac{1}{p}\int_{t\in Q} x(t)y(r-t)\right] e^{-2\pi i(h/p)(r-t)}e^{-2\pi i(h/p)t} \\
&= \frac{1}{p}\int_{t\in Q}\left[x(t)e^{-2\pi i(h/p)t}\frac{1}{p}\int_{r\in Q} y(r-t)e^{-2\pi i(h/p)(r-t)}\right] \\
&= \frac{1}{p}\int_{t\in Q}\left[x(t)e^{-2\pi i(h/p)t}\frac{1}{p}\int_{s\in Q} y(s)e^{-2\pi i(h/p)s}\right] \\
&= \left[\frac{1}{p}\int_{t\in Q} x(t)e^{-2\pi i(h/p)t}\right]\left[\frac{1}{p}\int_{s\in Q} y(s)e^{-2\pi i(h/p)s}\right] \\
&= x^\wedge(h/p)y^\wedge(h/p)
\end{aligned}$$

Also for $r \in \mathcal{Z}$, we define the convolution of the transforms $x^\wedge$ and $y^\wedge$ in $L^2(\mathcal{Z}/p)$, corresponding to period-p functions x and y in $L^2(Q)$, by:

$$(x^\wedge \circledast y^\wedge)(r/p) := \sum_{h=-\infty}^{\infty} x^\wedge(h/p)y^\wedge(r/p - h/p).$$

Then $x^\wedge \circledast y^\wedge = y^\wedge \circledast x^\wedge$, $x^\wedge \circledast y^\wedge = x^{\wedge R} \circledast y^{\wedge R}$, and also $(xy)^\wedge = x^\wedge \circledast y^\wedge$ and $(x^\wedge \circledast y^\wedge)^\vee = xy$.

Exercise 2.13: Show that $(xy)^\wedge = x^\wedge \circledast y^\wedge$.

Solution 2.13:

$$\begin{aligned} x(s)y(s) &= \left[\sum_k x^\wedge(k/p)e^{2\pi i(k/p)s}\right]\left[\sum_h y^\wedge(h/p)e^{2\pi i(h/p)s}\right] \\ &= \sum_k x^\wedge(k/p)e^{2\pi i(k/p)s}\sum_h y^\wedge(h/p-k/p)e^{2\pi i(h/p-k/p)s} \\ &= \sum_h\sum_k x^\wedge(k/p)y^\wedge(h/p-k/p)e^{2\pi i(h/p)s} \\ &= (x^\wedge \circledast y^\wedge)^\vee, \end{aligned}$$

so $(xy)^\wedge = x^\wedge \circledast y^\wedge$.

The operation of circular p-convolution of a function $x \in L^2(Q)$ by a fixed function g such that $x \circledast g$ exists is a linear transformation on $L^2(Q)$ (*i.e.*, $(\alpha x + \beta y) \circledast g = (\alpha(x \circledast g) + \beta(y \circledast g)$.) The relation $(x \circledast g)^\wedge(j/p) = g^\wedge(j/p)x^\wedge(j/p)$ shows that circular p-convolution by g is a "diagonal" linear transformation when applied to functions expressed in the Fourier basis in the same way as we described the differentiation linear transformation above.

Also, the operation of circular p-convolution by a fixed function g commutes with translations. Define $(T_a x)(t) = x(t+a)$. Then $(T_a)x \circledast g = T_a(x \circledast g)$. It is an interesting fact that any linear operator mapping $L^2(Q)$ to $L^2(Q)$ is expressible as circular p-convolution by some function (or functional,) this is precisely the reason that convolution is important and strongly connected to integral equations.

For suitably-smooth functions, the operation of circular p-convolution by a fixed function g also commutes with differentiation: $(x' \circledast g) + (x \circledast g)'$. (This means $(D_t^1 x(t)) \circledast g = D_r^1[(x \circledast g)(r)]$ where D_z^1 denotes differentiation with respect to z.) This further confirms that, just like differentiation, the operation of circular p-convolution by g is a "diagonal" transformation in the Fourier basis, since the "product" of diagonal linear transformations is commutative.

Exercise 2.14: Show that $(x' \circledast g) + (x \circledast g)'$.

Solution 2.14: $(x \circledast g)' = D_r^1\left[\frac{1}{p}\int_0^p x(t)g(r-t)\,dt\right] = \frac{1}{p}\int_0^p x(t)g'(r-t)\,dt = (x \circledast g')$. But then we also have $(g \circledast x)' = (g \circledast x')$, and $(x \circledast g)' = (g \circledast x)'$ and $(g \circledast x') = (x' \circledast g)$, so $(x \circledast g)' = (x' \circledast g)$.

Note this is to be expected because the product of diagonal linear transformations is again diagonal.

It is appropriate at this point to clarify operator notation. We may have *prefix* operators, say A and B, where we write $B(A(f)) = BA(f)$, and we can write $X = BA$ to define the prefix operator X as being the application of the operator A, followed by the application of the operator B. We may also have *postfix* operators, say C and D, where we write $f^{CD} = (f^C)^D$, and we can write $Y = CD$ to define the postfix operator Y as being the application of the operator C, followed by the application of the operator D.

Clearly, you need to be alert for whether an operator symbol is used as a postfix operator or a prefix operator. Mixing the two notations, which is commonly done, requires even more vigilance. Also note that many operators, like differentiation, have multiple ways of being denoted, sometimes using postfix symbols, sometimes using prefix symbols, and sometimes, if two arguments are involved, using infix symbols. And such operator symbols can be written as superscripts, as subscripts, or in "in-line" position. (Can you make a list of all the various differentiation notations in common use?)

Also for $x, y \in L^2(Q)$, we define the cross-correlation kernel function $x \otimes y$ by

$$(x \otimes y)(r) := (1/p) \int_0^p x(t)y(r+t)\,dt.$$

Then $x \otimes y = y \otimes x$, $(x \otimes y)^R = x^R \otimes y^R$, $x \otimes y = x^R \circledast y$, and $(x \otimes y)^\wedge = x^\wedge y^{\wedge R} = x^{\wedge R} y^\wedge$. When x and y are real, $(x \otimes y)^\wedge = x^\wedge y^{\wedge R} = x^\wedge y^{\wedge *}$ and $(x \otimes x)^\wedge = x^\wedge x^{\wedge *} = |x^\wedge|^2$.

2.6 The Dirichlet Kernel

Let x be a complex-valued periodic function, in $L^2(Q)$, so that x is of period p. The k-th partial sum $S_k(t) := \sum_{-k \le j \le k} x^\wedge(j/p) e^{2\pi i (j/p) t}$ of the Fourier series of $x(t)$ can be summed to yield

$$S_k(t) = \int_0^p D_k(y) x(t-y)\,dy = (D_k \circledast x)(t),$$

where, for $v \in [0, p)$,

$$D_k(v) = \begin{cases} 2k+1 & \text{if } v = 0 \\ \dfrac{sin(\pi(2k+1)v/p)}{sin(\pi v/p)} & \text{otherwise.} \end{cases}$$

Note the function $S_k(t)$ is just the *band-limited* (filtered) form of $x(t)$ having just the oscillatory terms with frequencies in $\{-k/p, \ldots, -1/p, 0, 1/p, \ldots, k/p\}$.

The function D_k is called the Dirichlet kernel; D_k is extended to be periodic with period p by defining $D_k(v + mp) = D_k(v)$ for $v \in [0, p)$ and $m \in \mathcal{Z}$.

Exercise 2.15: Show that

$$\sum_{-k \le j \le k} e^{2\pi i (j/p) v} = \begin{cases} 2k+1 & \text{if } v \bmod p = 0 \\ \dfrac{sin(\pi(2k+1)v/p)}{sin(\pi v/p)} & \text{otherwise.} \end{cases}$$

Hint: use the formula for the sum of a geometric series.

Solution 2.15:

If $v \in (0, p)$, we have:

$$\begin{aligned}
\sum_{-k\le j\le k} e^{2\pi i(j/p)v} &= \sum_{0\le j\le 2k} e^{2\pi i(j-k)v/p} \\
&= e^{-2\pi ikv/p} \sum_{0\le j\le 2k} e^{2\pi ijv/p} \\
&= e^{-2\pi ikv/p}[e^{2\pi i(2k+1)v/p} - 1]/[e^{2\pi iv/p} - 1] \\
&= e^{-2\pi ikv/p} e^{2\pi ikv/p}[e^{2\pi i(k+1)v/p} - e^{-2\pi ikv/p}]/[e^{2\pi iv/p} - 1] \\
&= [e^{2\pi i(k+1)v/p} - e^{-2\pi ikv/p}]/[e^{2\pi iv/p} - 1] \\
&= e^{\pi iv/p}[e^{\pi i(2k+1)v/p} - e^{-\pi i(2k+1)v/p}]/[e^{2\pi iv/p} - 1] \\
&= e^{\pi iv/p}[2i\, sin(\pi(2k+1)v/p)]/[(e^{\pi iv/p} - e^{-\pi iv/p})e^{\pi iv/p}] \\
&= 2i\, sin(\pi(2k+1)v/p)/[2i\, sin(\pi v/p)] \\
&= sin(\pi(2k+1)v/p)/\, sin(\pi v/p),
\end{aligned}$$

since $e^{i\theta} = cos(\theta) + i\, sin(\theta)$, $e^{-i\theta} = cos(\theta) - i\, sin(\theta)$, and taking the difference yields $e^{i\theta} - e^{-i\theta} = 2i\, sin(\theta)$. (Note since $v \in (0, p)$ by assumption, $sin(\pi v/p) \neq 0$.) Moreover, $\sum_{-k\le j\le k} e^{2\pi i(j/p)0} = 2k+1$. Thus the period-$p$ function $\sum_{-k\le j\le k} e^{2\pi i(j/p)v} = D_k(v)$.

Note $D_k(v)$ is the inverse Fourier transform of the discrete function b, where

$$b(s) = \begin{cases} 1 & \text{for } s = -k/p, \ldots, -1/p, 0, 1/p, \ldots, k/p, \\ 0 & \text{otherwise.} \end{cases}$$

The function b is a discrete function with stepsize $1/p$.

We have $D_k(v) = \sum_{-k\le j\le k} e^{2\pi i(j/p)v}$, and this exhibits the Fourier series for D_k, so we see that

$$D_k^\wedge(s/p) = b(s/p) = \begin{cases} 1 & \text{for } s = -k, \ldots, -1, 0, 1, \ldots, k \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $D_k^\wedge(s/p)x^\wedge(s/p) = \begin{cases} x^\wedge(s/p) & \text{for } s = -k, \ldots, -1, 0, 1, \ldots, k \\ 0 & \text{otherwise} \end{cases}$. And $D_k^\wedge x^\wedge = (D_k \circledast x)^\wedge$, so $D_k \circledast x = (D_k^\wedge x^\wedge)^\vee = \sum_{-k\le j\le k} 1 \cdot x^\wedge(v/p)e^{2\pi i(j/p)t} = S_k(t)$. And thus $S_k(t) = \int_0^p D_k(y)x(t-y)\, dy$.

But the function S_k is just the projection of x into the subspace B_k of $L^2(Q)$ spanned by $\{\, e^{2\pi i(h/p)t} \mid |h| \le k \,\}$ (!). Thus convolution with D_k is the projection operator into B_k.

2.7 The Spectra of Extensions of a Function

If x is given only on $[0, p)$, then $x^\wedge$ is, in fact, the Fourier transform of the strict period-p periodic extension of x. For x defined on $[0, q)$ with $q \ge p$, the strict period-p periodic extension of x is defined as $x_{[p]}(t) = x(t \bmod p)$ where $t \bmod p$ is that value $v \in [0, p)$ such that $t = v + kp$ with

$k \in \mathcal{Z}$. In general, if x is defined on $[a, a+q)$ with $q \geq p$, then the strict period-p periodic extension of x is the function $x_{[p]}(t) := x(a + ((t-a) \bmod p))$ for $t \in \mathcal{R}$. In order for the strict period-p periodic extension of x to exist, we must have x defined on an interval of length no less than p.

For x defined on $[0,p]$, we also have the even period-$2p$ periodic extension of x,

$$x_{ep}(t) = (x_e)_{[2p]}(t), \quad \text{where} \quad x_e(t) = \begin{cases} x(-t) & \text{if } -p \leq t \leq 0, \\ x(t) & \text{if } 0 \leq t < p, \\ 0 & \text{otherwise.} \end{cases}$$

and the odd period-$2p$ periodic extension of x,

$$x_{op}(t) = (x_o)_{[2p]}(t), \quad \text{where} \quad x_o(t) = \begin{cases} -x(-t) & \text{if } -p \leq t < 0, \\ x(t) & \text{if } 0 \leq t < p, \\ 0 & \text{otherwise.} \end{cases}$$

Now $(x_{ep})^{\wedge(2p)}(s) = ((x_{[p]})^{\wedge(p)}(-s) + (x_{[p]})^{\wedge(p)}(s))/2$ and $(x_{op})^{\wedge(2p)}(s) = ((x_{[p]})^{\wedge(p)}(s) - (x_{[p]})^{\wedge(p)}(-s))/2$, where $s \in \{\ldots, -1/(2p), 0, 1/(2p), \ldots\}$.

(In these equations, and those that follow, we may ignore the stricture that $y^{\wedge(a)}(s)$ is only computed for $s \in \{\ldots, -1/a, 0, 1/a, \ldots\}$, and instead, we may allow s to range over any desired discrete range $\ldots, -1/b, 0, 1/b, \ldots\}$. We will denote this by writing $y^{\wedge(a)}(s)_b$.

Also, when x coincides on $[0,p]$ with a function y of period $q \geq p$, we define the period-q y-extension of x, $x_{\langle y \rangle}(t) = y(t)$.

Let $0 < p \leq q$ and let $y(t)$ be defined on $[0,q]$ with $y(t) = x(t)$ for $t \in [0,p]$. Take $s \in \{\ldots, -1/q, 0, 1/q, \ldots\}$. Now $(x_{\langle y \rangle})^{\wedge(q)}(s) = \dfrac{1}{q}\left[\int_0^p x(t)e^{-2\pi ist}\,dt + \int_p^q y(t)e^{-2\pi ist}\,dt\right]$ and $\int_p^q y(t)e^{-2\pi ist}\,dt = \int_0^{q-p} y(t+p)e^{-2\pi ist}e^{-2\pi isp}\,dt = e^{-2\pi isp}\int_0^{q-p} y(t+p)e^{-2\pi ist}\,dt$. Note when s is an integer multiple of $1/p$, we have $e^{-2\pi isp} = 1$. Let $z(t) := y(t+p)$ for $0 < t \leq q-p$. Then $q \cdot (x_{\langle y \rangle})^{\wedge(q)}(s) = p \cdot (x_{[p]})^{\wedge(p)}(s)_q + (q-p) \cdot e^{-2\pi isp}(z_{[q-p]})^{\wedge(q-p)}(s)_q$.

When $y(t+p) = x(t)$ for $0 < t \leq q-p$, so that we have used the period-p function x on $[0,q]$ instead of on $[0,p]$ to define y, then we have $y(t) = x_{[q]}(t)$ and $q \cdot (x_{[q]})^{\wedge(q)}(s) = \lfloor q/p \rfloor \cdot p \cdot (x_{[p]})^{\wedge(p)}(s)_q + r \cdot (x_{[r]})^{\wedge(r)}(s)_q$ where $r = q \bmod p$.

Exercise 2.16: Show that $q \cdot (x_{[q]})^{\wedge(q)}(s) = \lceil q/p \rceil \cdot r \cdot (x_{[r]})^{\wedge(r)}(s)_q + \lfloor q/p \rfloor \cdot a \cdot (z_{[a]})^{\wedge(a)}(s)_q$ where $z(t) = x(t+r)$ for $0 \leq t \leq p-r$ and $r = q \bmod p$ and $a = p - r$.

If $0 < q \leq p$, then $q \cdot (x_{[q]})^{\wedge(q)}(s) = p \cdot (x_{[p]})^{\wedge(p)}(s)_q - b \cdot (z_{[b]})^{\wedge(b)}(s)_q$ where $z(t) = x(t+q)$ for $0 \leq t \leq p-q$ and $b = p - q$.

When a function x given on $[0,p]$ is to be extended, the function x_{ep} is often used, since, if x has no discontinuities in $[0,p]$, x_{ep} is continous everywhere, and hence no artificial high-frequency components arise in the spectrum of x_{ep} due to discontinuities.

2.8 Spectral Power Density Function

Let $x(t)$ be the voltage across a 1 Ohm resistance at time t. By Ohm's law, the current through the resistance at time t is $x(t)/1$ Amperes. Thus, the power being used at time t to heat the resistance is $x(t) \cdot x(t)/1$ Joules/second, $(1/p)\int_0^p x(t)^2\,dt$ is the average power used, measured in Joules per second, averaged over p seconds, and $\int_0^p x(t)^2\,dt$ Joules is the total amount of energy converted into heat in p seconds. Note power is the rate at which energy is converted, so to say the average power used over p seconds is y Joules/second is the same as saying that the average rate at which energy is used (converted) over p seconds is y Joules/second. The square root of $(1/p)\int_0^p x(t)^2\,dt$ is called the root-mean-square (RMS) value of x over p seconds, measured in the same units as x.

Now suppose $x(t) \in L^2(Q)$ is a complex-valued periodic function of period p. By Plancherel's identity, $\|x\|^2_{L^2(Q)} = \|x^\wedge\|^2_{L^2(\mathcal{Z}/p)}$, so the average power used over p seconds is $\|x\|^2_{L^2(Q)}/p = \|x^\wedge\|^2_{L^2(\mathcal{Z}/p)}/p = \sum_{h=-\infty}^{\infty} |x^\wedge(h/p)|^2$, and the total amount of energy transformed into heat over one period is $\|x\|^2_{L^2(Q)} = p\sum_{h=-\infty}^{\infty} |x^\wedge(h/p)|^2 = \|x^\wedge\|_{L^2(\mathcal{Z}/p)}$.

The function $|x^\wedge(s)|^2$ is called the spectral power density function of x. The average power used over one period due to the complex spectral components of x in the frequency band $[a, b]$ is $\sum_{h=ap}^{bp} |x^\wedge(h/p)|^2$.

When x is real-valued, $x^\wedge$ is Hermitian, so $|x^\wedge|^2 = (x \circledast x^R)^\wedge$, and the spectral power density function is even. Thus, folding into the positive frequencies results in the energy due to the real spectral components in the positive frequency band $[a, b]$, with $0 \le a \le b$, being $\sum_{h=ap}^{bp}(2-\delta_{h,0}) \cdot |x^\wedge(h/p)|^2$. Note $|x^\wedge(h/p)|^2 = x^\wedge(h/p)x^\wedge(h/p)^* = M(h/p)^2/(2-\delta_{h,0})^2$ when x is real, so

$$(2-\delta_{h,0})|x^\wedge(h/p)|^2 = M(h/p)^2.$$

3 The Discrete Fourier Transform

Let $x(t)$ be a discrete complex-valued periodic function of period p with stepsize T defined at $t = \ldots$, $-2T$, $-T$, 0, T, $2T$, $\ldots$ with $p = nT$, where $n \in \mathcal{Z}^+$ ($\mathcal{Z}^+ := \{j \in \mathcal{Z} \mid j > 0\}$). This means that $x(kT) = x((k+n)T)$ for $k \in \mathcal{Z}$. We shall only consider *regular* discrete functions in this section. Thus either of the discrete sequences $0, T, \ldots, (n-1)T$ or $-\lfloor n/2\rfloor T, \ldots, -T, 0, T, \ldots, (\lceil n/2\rceil - 1)T$, among others, constitutes the domain of x confined to one period. Of course, x may in fact be defined on the whole real line.

The discrete Fourier transform of x is

$$x^\wedge(s) = (T/p)\sum_{h=-\lfloor n/2\rfloor}^{\lceil n/2\rceil-1} x(hT)e^{-2\pi ishT},$$

where $x^\wedge(s)$ is defined for $s = \ldots, -2/p, -1/p, 0, 1/p, 2/p, \ldots$, with $p = nT$. Note if $T \in \mathcal{Z}^+$ then $p \in \mathcal{Z}^+$, and if $p = 1$ and $T \in \mathcal{Z}^+$ then $n = T = 1$ and x is the (degenerate) period-1 discrete function of stepsize 1. In general, if x is defined with an integral stepsize, $x^\wedge$ will have a non-integral stepsize unless x is degenerate. The transform $x^\wedge$ is a discrete periodic function of period n/p with stepsize $1/p$. The ratio of the period and the stepsize is the same value n for both x and $x^\wedge$.

This sum is just a rectangular Riemann sum approximation of the integral form of the Fourier transform of an integrable periodic function x, computed on a regular mesh of points, each of which is T units apart from the next. (The value hT serves in the role of t and the factor T serves in the role of dt in the discretization.)

Unlike the transform of a periodic function defined on the entire real line, $x^\wedge$ is also a (discrete) periodic function, and hence the inverse operator, $\vee$, acts on the same type of functions as the direct operator, $\wedge$, but, in general, with a different period and stepsize. When necessary we shall write $\wedge(p;n)$ and $\vee(p;n)$ to denote the discrete Fourier transform and the inverse discrete Fourier transform for discrete periodic functions of period p with stepsize p/n. Then $\vee(n/p;n)$ is the inverse operator of the operator $\wedge(p;n)$.

The inverse discrete Fourier transform of the discrete function y of period p with stepsize $T = p/n$ is

$$y^{\vee(p;n)}(r) = \sum_{h=-\lfloor n/2 \rfloor}^{\lceil n/2 \rceil - 1} y(hT)e^{2\pi i r hT},$$

for $r = \ldots, -2/p, -1/p, 0, 1/p, 2/p, \ldots$. Here $y^{\vee(p;n)}$ is a discrete period-n/p function with stepsize $1/p$. By convention, the period-$\frac{n}{p}$ functions $y^{\wedge(p;n)}(s)$ and $y^{\vee(p;n)}(s)$ are taken to be discrete functions with values at $\ldots, -2/p, -1/p, 0, 1/p, 2/p, \ldots$, even though the expressions at hand may make sense on an entire interval of real numbers. Just as $\vee(n/p;n)$ is the inverse operator of $\wedge(p;n)$, $\vee(p;n)$ is the inverse operator of $\wedge(n/p;n)$. When $\wedge$ is understood to be $\wedge(p;n)$, $\vee$ shall normally be understood to be $\vee(n/p;n)$. Note when the function y is of period $p = nT$ with stepsize $p/n = T$, both $y^{\wedge(p;n)}$ and $y^{\vee(p;n)}$ are periodic of period $n/p = 1/T$, and are defined on a mesh of stepsize $1/p$.

For a regular discrete periodic function, the relationship between the period p, the stepsize T, and the number of steps in one period, n, is $p = nT$. The discrete Fourier transform of a period-p, stepsize-T, number-of-steps-n periodic discrete function is a period-n/p, stepsize-$1/p$, number-of-steps-n periodic discrete function. Thus, we should, perhaps, denote $\wedge(p;n)$ by $\wedge(p;T;n)$, and denote $\vee(n/p;n)$ by $\vee(n/p;1/p;n)$.

In general, with $p = nT$, we have:

period	stepsize	#-steps		period	stepsize	#-steps
$p = nT$	$\frac{p}{n} = T$	n	$\xrightarrow{\wedge}$	$\frac{n}{p} = \frac{1}{T}$	$\frac{1}{p} = \frac{1}{nT}$	n

And with $T \to 0$ and $n \to \infty$ such that $nT = p$, we have

period	stepsize	#-steps		period	stepsize	#-steps
p	0	∞	$\xrightarrow{\wedge}$	∞	$\frac{1}{p}$	∞

And now, when $p \to \infty$, we have

$$\begin{array}{ccc} \underline{\text{period}} & \underline{\text{stepsize}} & \underline{\#\text{-steps}} \\ \infty & 0 & \infty \end{array} \xrightarrow{\wedge} \begin{array}{ccc} \underline{\text{period}} & \underline{\text{stepsize}} & \underline{\#\text{-steps}} \\ \infty & 0 & \infty. \end{array}$$

We see that the 'period' of the Fourier transform of a function is the reciprocal of its 'stepsize' and its stepsize is the reciprocal of its period. Our notation for the periodic Fourier transform operator on the period-p functions in $L^2(Q)$ is $\wedge(p)$, and the corresponding inverse Fourier transform operator on the stepsize-$1/p$ functions in $L^2(\mathcal{Z}/p])$ is written as $\vee(p)$; whereas it would be more consistent to write $\wedge(p; 0; \infty)$ and $\vee(\infty; 1/p; \infty)$, but, as with the discrete Fourier transform $\wedge(p; n)$ and its inverse $\vee(n/p; n)$, we opt for brevity with mnemonic content instead.

Note that when x is a periodic period-1 function in $L^2([0, 1])$, the Fourier transform $x^\wedge$ is a discrete stepsize-1 function; in this case, the scale-factor $\frac{1}{p}$ "vanishes" in the Fourier transform integral; because of this, we often see the Fourier transform introduced for period-1 functions to simplify the expressions involved.

For x of period $p = nT$ with stepsize T, the inverse discrete Fourier transform of the stepsize-$1/p$, period-$\frac{n}{p}$ function $x^\wedge$ is

$$x^{\wedge(p;n)\vee(n/p;n)}(t) = \sum_{h=-\lfloor n/2\rfloor}^{\lceil n/2\rceil-1} x^{\wedge(p;n)}(h/p)e^{2\pi i(h/p)t},$$

where $x^{\wedge\vee}$ is of period p defined on a mesh of stepsize $T = p/n$, and $x^{\wedge\vee}(t) = x(t)$ for $t = \ldots, -2T, -T, 0, T, 2T, \ldots$. This is the Fourier series of the discrete function x; note it is a finite sum.

For $-\lfloor n/2\rfloor \le h \le \lceil n/2\rceil - 1$, $x^\wedge(h/p)$ is the complex amplitude of the complex oscillation $e^{2\pi i(h/p)t}$ of frequency h/p cycles per t-unit in the finite Fourier series $x^{\wedge\vee}$, and $x^{\wedge\vee}$ is a sum of n complex oscillations of frequencies $-\lfloor n/2\rfloor/p, \ldots, 0, \ldots, (\lceil n/2\rceil - 1)/p$. Thus $x^{\wedge\vee}$ is *band-limited*; that is, the finite n-term Fourier series $x^{\wedge\vee}$ has no terms for frequencies outside the finite interval or band $[-\lfloor n/2\rfloor/p, (\lceil n/2\rceil - 1)/p]$.

Observe that by allowing t to be any real value, $x^{\wedge(p;n)\vee(n/p;n)}(t)$ is defined for all t; it is a continuous periodic function of period p which coincides with x at $t = \ldots, -2T, -T, 0, T, 2T, \ldots$. Indeed the function $x^{\wedge(p;n)\vee(n/p;n)}(t)$ defined on $\mathcal{R}$ is the unique period-p periodic function in $L^2(Q)$ with this property which is band-limited with $x^{\wedge(p;n)\vee(n/p;n)\wedge(p;n)}(s) = 0$ for s outside the band $[-\lfloor n/2\rfloor/p, (\lceil n/2\rceil - 1)/p]$. If x is real, then when n is odd, $x^{\wedge\vee}$ is real, but when n is even, $x^{\wedge\vee}$ is complex in general, even though $x^{\wedge\vee}(t)$ is real when t is an integral multiple of T.

Note $x^{\wedge(p;n)\vee(n/p;n)}(t)$ is an *interpolating* function for the points $(kT, x(kT))$ for $k = -\lfloor n/2\rfloor, \ldots, \lceil n/2\rceil - 1$. Since $x^{\wedge(p;n)\vee(n/p;n)}(t)$ is periodic, it is, in fact, an interpolation function defined over the entire real line that interpolates the points $(kT, x(kT))$ for $k = \ldots, -2, -1, 0, 1, 2, \ldots$.

Exercise 3.1: Show that for a discrete periodic period-p function x, with $p \in \mathcal{Z}^+$, when the number of steps $n = p^2$, the stepsize of x is $\frac{1}{p}$, and in this case, the discrete Fourier transform $x^{\wedge(p)}$ is also a discrete periodic period-$-p$ function with stepsize-$\frac{1}{p}$ and the number of steps in each period is n.

Another useful form of the discrete Fourier Inversion theorem is

$$x^{\wedge(p;n)\vee(n/p;n)}(kT) = x(kT) = \sum_{h=-\lfloor n/2\rfloor}^{\lceil n/2\rceil-1} x^{\wedge(p;n)}(h/p)e^{2\pi i(h/n)k},$$

where x is a discrete period-p function with $p = nT$.

For $nT = p$, the functions $x(hT)$ and $e^{-2\pi ishT}$ with s an integral multiple of $1/p$ are both periodic functions of hT with period p and are simultaneously periodic functions of h with period n, and hence the discrete Fourier transform of x can be obtained by summing over any contiguous index sequence of length n, so that $x^\wedge(s) = (T/p)\sum_{h=a}^{n-1+a} x(hT)e^{-2\pi ishT}$ for $s = \ldots, -1/p, 0, 1/p, \ldots$. Similarly, $x^\wedge(h/p)$ and $e^{2\pi i(h/p)t}$ with t a multiple of T are both periodic functions of h/p with period n/p and periodic functions of h with period n, so $x^{\wedge\vee}(t) = \sum_{h=a}^{n-1+a} x^\wedge(h/p)e^{2\pi i(h/p)t}$, for $t = \ldots, -2T, -T, 0, T, 2T, \ldots$. Note that the periodicity of x, where $x(-kT) = x((n-k)T)$ for $k \in \mathcal{Z}$, implies that $x^\wedge(-k/p) = x^\wedge((n-k)/p)$ for $k \in \mathcal{Z}$.

Exercise 3.2: Show that for $p = n$ with $n \in \mathcal{Z}^+$,

$$y(s) := x^{\wedge(n;n)}(\frac{s}{n}) = \frac{1}{n}\sum_{h=0}^{n-1} x(h)e^{2\pi ish/n} \text{ for } s = 0, 1, 2, \ldots, n-1.$$

Thus when $p = n$, x and y are both period-n, stepsize-1 discrete functions defined on $\mathcal{Z}$, and $x^{\wedge(n;n)}(s) = y(ns)$.

Indeed, when summing $\frac{T}{p}x(hT)e^{2\pi ishT}$ over $h = a, a+1, \ldots, n-1+a$, the periodicity of x insures the *same* values are being summed, regardless of the value of the integer a, so the Fourier series denoted by $x^{\wedge\vee}$ is a unique sum of complex oscillations, which, when expressed in the particular form where $a = -\lfloor n/2\rfloor$, allows us to easily combine the positive and negative frequency terms (treating $x^\wedge(\lceil n/2\rceil/p)$ as 0 in these combinations when n is even) and shows the spectral decomposition of $x^{\wedge\vee}$ to be

$$\tilde{x}(t) = \sum_{h=0}^{\lfloor n/2\rfloor} M(h/p)\, cos(2\pi(h/p)t + \phi(h/p)),$$

where M and ϕ are defined as for the spectral decomposition of a period-p function in $L^2(Q)$ and $\tilde{x}$ uniformly approximates x. When x is real and n is even, we may introduce a special redefinition of $M(n/(2p))$ and $\phi(n/(2p))$ which makes $\tilde{x}$ real and satisfies the identity $\tilde{x}(kT) = x(kT)$. (This redefinition is based on the fact that $x^\wedge(-k/p) = x^\wedge(k/p)^*$ for $k \in \mathcal{Z}$ when x is real.) Take

$$M(h/p) = \left[\sqrt{(x^\wedge(h/p) + x^\wedge(-h/p))^2 - (x^\wedge(h/p) - x^\wedge(-h/p))^2}\right]/(1 + \delta_{h,0})$$

for $0 \le h \le \lceil n/2 \rceil - 1$ and, when n is even, let $M(n/(2p)) = |x^\wedge(-n/(2p))|$, and, by definition, $M(h/p) = 0$ for $h > \lfloor n/2 \rfloor$. Also take

$$\phi(h/p) = \text{atan2}(-i(x^\wedge(h/p) - x^\wedge(-h/p)), x^\wedge(h/p) + x^\wedge(-h/p)).$$

for $0 \le h \le \lceil n/2 \rceil - 1$, and, when n is even, define $\phi(n/(2p)) = \text{atan2}(0, x^\wedge(-n/(2p)))$, and $\phi(h/p) = 0$ for $h > \lfloor n/2 \rfloor$.

Exercise 3.3: Show that $x^\wedge(h/(2p))$ is real when x is real and n is even.

Exercise 3.4: Show that $\sqrt{(x^\wedge(h/p) + x^\wedge(-h/p))^2 - (x^\wedge(h/p) - x^\wedge(-h/p))^2} = 2|x^\wedge(h/p)|^2$ for $h \in \mathcal{Z}$ when x is real.

Exercise 3.5: Let $w = e^{-2\pi i/n}$ with $n \in \mathcal{Z}^+$. Show that w is a *primitive* n-th *root of unity*, *i.e.*, $w^h \ne 1$ for $h \in \{1, 2, \ldots, n-1\}$ and $w^n = 1$. Also show that, for $k \in \mathcal{Z}^+$, w^k is a primitive $(n/\gcd(k,n))$-th root of unity. How many distinct primitive n-th roots of unity are there?

Let $x(t)$ be a discrete complex-valued periodic function of period p with stepsize $T = p/n$ and let $q(z)$ be the polynomial $q_0 + q_1 z + q_2 z^2 + \cdots + q_{n-1} z^{n-1}$, where $q_h = \dfrac{1}{n} x(hT)$ for $h \in \{0, 1, 2, \ldots, n-1\}$ Show that $x^\wedge(k/p) = q(w^k)$ for $k \in \mathcal{Z}$. Thus the value $x^\wedge(k/p)$ is computed by obtaining the value of the polynomial q at w^k.

Finally, show that the periodicity of x and $e^{-2\pi ikh/n}$ yields $x^\wedge(k/p) = q(r^k)$ where r is *any* primitive n-th root of unity.

3.1 Geometrical Interpretation

Let $\mathcal{Z}$ denote the set of integers $\{\ldots, -1, 0, 1, \ldots\}$ and let $T\mathcal{Z}$ denote the set of values $\{\ldots, -T, 0, T, \ldots\}$. Let $d_n(T\mathcal{Z})$ denote the set of complex-valued discrete periodic functions defined on $T\mathcal{Z}$ of period $p = nT$.

Note $T/p = 1/n$ and introduce the inner product $(x, y)_{d_n(T\mathcal{Z})} = T \sum_{h=0}^{n-1} x(hT) y(hT)^*$ and the norm $\|x\|_{d_n(T\mathcal{Z})} = (x, x)^{1/2}_{d_n(T\mathcal{Z})}$. Then $d_n(T\mathcal{Z})$ is a finite-dimensional Hilbert space of dimension n, and the sequence of functions $\langle e_0, e_1, \ldots, e_{n-1} \rangle$ is an orthogonal basis for $d_n(T\mathcal{Z})$, where $e_k(hT) = e_k(hp/n) := e^{2\pi i(k/p)(hp/n)} = e^{2\pi ik(h/n)}$ for $h \in \mathcal{Z}$.

Exercise 3.6: Show that $(e_j, e_k)_{d_n(T\mathcal{Z})} = \delta_{j,k} p$, and $\|e_j\| = p^{1/2}$.

Solution 3.6: $(e_j, e_k)_{d_n(T\mathcal{Z})} = T \sum_{0 \le h \le n-1} (e^{2\pi ijh/n})(e^{2\pi ikh/n})^* = T \sum_{0 \le h \le n-1} e^{2\pi i(j-k)h/n}$. Let $w = e^{2\pi i(j-k)/n}$. Then, $(e_j, e_k)_{d_n(T\mathcal{Z})} = T[1 + w + w^2 + \cdots + w^{n-1}]$. Then if $j = k$, $w = 1$ and $(e_j, e_k)_{d_n(T\mathcal{Z})} = Tn = p$. If $j \ne k$, $(e_j, e_k)_{d_n(T\mathcal{Z})} = T[(w^n - 1)/(w - 1)]$, and $w^n = 1$, so $(e_j, e_k)_{d_n(T\mathcal{Z})} = 0$.

The discrete Fourier transform $\wedge(p; n)$ maps $d_n(T\mathcal{Z})$ onto $d_n(\mathcal{Z}/p)$, which is also an n-dimensional Hilbert space. Thus, with respect to the basis $\langle e_0, e_1, \ldots, e_{n-1} \rangle$, $\wedge(p; n)$ is explicitly representable by an $n \times n$ non-singular matrix, F_n, where $(F_n)_{j,k} = (1/n) e^{-2\pi i(j-1)(k-1)/n}$ for $1 \le j \le n$ and

$1 \le k \le n$, and $x^\wedge = xF_n$ where x and $x^\wedge$ are taken as the vectors $[x(0), x(T), \ldots, x((n-1)T)]$ and $[x^\wedge(0), x^\wedge(1/p), \ldots, x^\wedge((n-1)/p)]$.

Exercise 3.7: Show that the matrix F_n is symmetric, *i.e.*, $(F_n)_{j,k} = (F_n)_{k,j}$ for $1 \le j, k \le n$.

Note that the matrix F_n representing $\wedge(p;n)$ does not depend on the period p or the stepsize T, but only on their ratio n. Also note that, formally, $domain(F_n) \ne range(F_n)$, unless $n = p^2$ and $p^2 \in \mathcal{Z}$, however, both $d_n(T\mathcal{Z})$ and $d_n(\mathcal{Z}/p)$ are isomorphic to $\mathcal{C}^n$, so this nit can be resolved as follows.

Introduce the "vectorizing" isomorphism $c_{p,n} : d_n(T\mathcal{Z}) \to \mathcal{C}^n$ where $c_{p,n}(x) = (x(0), x(T), \ldots, x((n-1)T))$, with $T = p/n$. This is, in essence, just a rescaling of $T\mathcal{Z}$ by $1/T$. Note $c_{p,n}$ is a linear one-to-one map of $d_n(T\mathcal{Z})$ onto $\mathcal{C}^n$, and $c_{n/p,n}$ is a linear one-to-one map of $d_n(\mathcal{Z}/p)$ onto $\mathcal{C}^n$.

Then the "correct" way to specify the discrete Fourier transform via the matrix F_n is: $c^{-1}_{n/p,n}(c_{p,n}(x)F_n) = x^{\wedge(p;n)}$.

If we want a "nice" form of Parseval's identity to hold, we need to define the alternate inner product $\langle f, g\rangle_{d_n(\mathcal{Z}/p)}$ on $d_n(\mathcal{Z}/p)$ as $\langle f, g\rangle_{d_n(\mathcal{Z}/p)} = p\sum_{h=0}^{n-1} f(h/p)g(h/p)^*$. With this choice, Parseval's identity holds with the factor p hidden: we have $(x, y)_{d_n(T\mathcal{Z})} = \langle x^{\wedge(p;n)}, y^{\wedge(p;n)}\rangle_{d_n(\mathcal{Z}/p)}$.

We also have the "standard" Hermitian inner-product $(u, v)_{\mathcal{C}^n} := \sum_{j=0}^{n-1} u_j v_j^*$ for vectors $u, v \in \mathcal{C}^n$, and we see that the isomorphism $c_{p,n}$ "relates" the inner-product $(\cdot,\cdot)_{d_n(T\mathcal{Z})}$ on $d_n(T\mathcal{Z})$ to the Hermitian inner-product on $\mathcal{C}^n$ as $(x, y)_{d_n(T\mathcal{Z})} = T(c_{p,n}(x), c_{p,n}(y))_{\mathcal{C}^n}$. Similarly, $\langle f, g\rangle_{d_n(\mathcal{Z}/p)} = nT(c_{n/p,n}(f), c_{n/p,n}(g))_{\mathcal{C}^n}$.

Therefore Parseval's identity for a discrete period p, stepsize T, sample-size n function corresponds to $T(c_{p,n}(x), c_{p,n}(y))_{\mathcal{C}^n} = nT(c^{-1}_{n/p,n}(c_{p,n}(x)F_n), c^{-1}_{n/p,n}(c_{p,n}(y)F_n))_{\mathcal{C}^n}$.

Since $c_{p,n}$ and $c_{n/p,n}$ are linear mappings, we can use the factor $\sqrt{n}$ to "rescale" the matrix F_n to write $(c_{p,n}(x), c_{p,n}(y))_{\mathcal{C}^n} = (c^{-1}_{n/p,n}(c_{p,n}(x)\sqrt{n}F_n),\ c^{-1}_{n/p,n}(c_{p,n}(y)\sqrt{n}F_n))_{\mathcal{C}^n}$. For $\sqrt{n}F_n$ taken as an $n \times n$ matrix of a linear transformation mapping $\mathcal{C}^n$ onto C^n, this statement is just $(u, v)_{\mathcal{C}^n} = (u\sqrt{n}F_n,\ v\sqrt{n}F_n)_{\mathcal{C}^n}$ for $u, v \in \mathcal{C}^n$, *i.e.*, the linear transformation given by $\sqrt{n}F_n$ *preserves lengths*.

This identity means that, as a linear transformation mapping $\mathcal{C}^n$ onto $\mathcal{C}^n$, the matrix $\sqrt{n}F_n$ is *unitary*, *i.e.*, $(\sqrt{n}F_n)^{-1} = (\sqrt{n}F_n)^{*\mathrm{T}} = \sqrt{n}F_n^{*\mathrm{T}}$. And since F_n is symmetric, $F_n^{-1} = nF_n^*$. Thus the inverse discrete Fourier transform is represented as a matrix by nF_n^* which matches the defining inverse summation formula.

Exercise 3.8: Show that the matrix $F_n^{-1} = nF_n^*$. Hint: $(\sqrt{n}F_n)^{-1} = \dfrac{1}{\sqrt{n}}F_n^{-1}$.

Exercise 3.9: Show that $(F_n^{-1})_{j,k} = e^{2\pi i(j-1)(k-1)T/p}$ for $1 \le j \le n$ and $1 \le k \le n$.

Exercise 3.10: Show that $\|F_n \text{ row } j\|_{d_n(T\mathcal{Z})} = n^{-1/2}$.

Exercise 3.11: Show that the matrix F_n is normal, *i.e.*, $F_nF_n^H = F_n^HF_n$ where $F_n^H := F_n^{*\mathrm{T}}$. This means F_n is unitarily diagonalizable, and there is an orthonormal $\mathcal{C}^n$-basis consisting of eigenvectors of F_n.

Exercise 3.12: Let $v \in \mathcal{C}^n$. Show that $v = (v^* \sqrt{n} F_n)^* \sqrt{n} F_n$.

Exercise 3.13: Let $\bar{e}_k(h/p) = e^{2\pi i k h/n} = e^{2\pi i k (h/p) T}$ with $p = nT$. Show that $\langle \bar{e}_0, \bar{e}_1, \ldots, \bar{e}_{n-1} \rangle$ is an orthogonal basis for $d_n(\mathcal{Z}/p)$ and show that $\langle \bar{e}_0, \bar{e}_1, \ldots, \bar{e}_{n-1} \rangle = \langle e_0, e_{n-1}, e_{n-2}, \ldots, e_1 \rangle$. Hint: $\bar{e}_j = e_j^*$.

Note, for $n \geq 4$, that $F_n^4 = n^{-2} I$ as shown below, and therefore $F_n^{-1} = n^2 F_n^3$. The matrix F_n is a matrix "analog" of the scalar $\frac{1}{\sqrt{n}} \cdot r$ where r is a primitive fourth root of unity in $\mathcal{C}$.

Exercise 3.14: Show that $F_n^2 = n^{-1} S_n$ where $S_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & . & \vdots & \vdots \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}$.

Explicitly, for $1 \leq j \leq n$ and $1 \leq k \leq n$, $(S_n)_{jk} = \begin{cases} 1 & \text{if } (j+k-2) \bmod n = 0, \\ 0 & \text{otherwise.} \end{cases}$

Solution 3.14:

$$\begin{aligned} (F_n F_n)_{s,t} &= \sum_{j=1}^{n} (F_n)_{s,j} (F_n)_{j,t} \\ &= \sum_{j=1}^{n} \frac{1}{n} e^{-2\pi i (s-1)(j-1)/n} \cdot \frac{1}{n} e^{-2\pi i (j-1)(t-1)/n} \\ &= n^{-2} \sum_{j=1}^{n} e^{-2\pi i (j-1)[s-1+t-1]/n} \\ &= n^{-2} \sum_{j=0}^{n-1} e^{-2\pi i j [s+t-2]/n} \\ &= n^{-2} \sum_{j=0}^{n-1} w^j, \text{ where } w = e^{-2\pi i [s+t-2]/n}. \end{aligned}$$

If $w = 1$, we have $(F_n F_n)_{s,t} = n^{-2} n$. If $w \neq 1$, we have $(F_n F_n)_{s,t} = n^{-2}(w^n - 1)/(w - 1)$ and $w^n = 1$, so $(F_n F_n)_{s,t} = 0$. But $w = e^{-2\pi i [s+t-2]/n} = 1$ precisely when $s + t - 2$ is an integral multiple of n which occurs when $s = t = 1$ or when $s = k$ and $t = n + 2 - k$ for $k = 2, 3, \ldots, n$. Thus $(F_n^2)_{s,t} = \begin{cases} n^{-1} & \text{if } (s+t-2) \bmod n = 0, \\ 0 & \text{otherwise,} \end{cases}$ so $F_n^2 = n^{-1} S_n$.

Note $S_n^2 = I$, so $F_n^4 = n^{-2} S_n^2 = n^{-2} I$. Also $(\sqrt{n} F_n)^{-1} = \sqrt{n} F_n \sqrt{n} F_n \sqrt{n} F_n$, so $F_n^{-1} = n^2 F_n^3$.

The matrix S_n corresponds to the reversal operator: $c_{p,n}^{-1}(c_{p,n}(x) S_n) = x^R$. This means we can imagine the Fourier transform matrix F_n as representing a 90° *rotation* composed with a scale by $\sqrt{n}$, since $(\sqrt{n} F_n)^2 = S_n$. We may imagine the *reversal* of a function x to be the 180° rotation of the function x about 0. (Note the reversal of the function x is also the *reflection* of x in the complex plane through the subspace orthogonal to the real line. Thus a reflection transformation is the composition of two 90° rotations, *i.e.*, a reflection transformation has a 90° rotation as a "square-root".)

Now we see that, for $n \geq 4$, the minimal polynomial of F_n is $\lambda^4 - n^{-2}$, and thus the distinct eigenvalues of F_n are the fourth-roots of n^{-2}: $1/\sqrt{n}, i/\sqrt{n}, -1/\sqrt{n}$, and $-i/\sqrt{n}$. This follows from the fact that $F_n^4 - n^{-2}I = 0$. If $vF_n = \lambda v$ for some vector $v \in \mathcal{C}^n - 0$, then $vF_n^2 = \lambda v F_n = \lambda^2 v$, $vF_n^3 = \lambda^3 v$, and $vF_n^4 = \lambda^4 v$, *i.e.*, the eigenvalues of F_n^4 are the 4-th powers of the eigenvalues of F_n. But, $vF_n^4 = n^{-2}v$, so n^{-2} is the only eigenvalue of F_n^4, and hence the 4-th roots $(\sqrt{n})^{-1}$, $-(\sqrt{n})^{-1}$, $i(\sqrt{n})^{-1}$, and $-i(\sqrt{n})^{-1}$ are the eigenvalues of F_n.

Exercise 3.15: Show that $xF_n^4 = n^2 x^{\wedge(p;n)\wedge(n/p;n)\wedge(p;n)\wedge(n/p;n)} = x$.

Solution 3.15: $F_n^{-1} = n^2 F_n^3$, so $x^{\wedge\vee\wedge\vee} = xF_n n^2 F_n^3 F_n n^2 F_n^3 = n^4 x F_n^8 = n^4 x(n^{-2}I)(n^{-2}I) = x$.

The multiplicities of these eigenvalues have been determined [MP72]. Let $m = \lfloor n/4 \rfloor$ and let $r = n \bmod 4$. Then $\mathit{multiplicity}(1/\sqrt{n}) = m + 1$, $\mathit{multiplicity}(i/\sqrt{n}) = m - \delta_{r,0}$, $\mathit{multiplicity}(-1/\sqrt{n}) = m + \delta_{r,2} + \delta_{r,3}$, and $\mathit{multiplicity}(-i/\sqrt{n}) = m + \delta_{r,3}$. Since the matrix F_n is normal, these multiplicities are the dimensions of the four corresponding eigenspaces of F_n.

Exercise 3.16: Show that for $n \in \mathcal{Z}^+$ with $m = \lfloor m/4 \rfloor$ and $r = n \bmod 4$, $4m + 1 + \delta_{r,2} + 2\delta_{r,3} - \delta_{r,0} = n$.

For $n \geq 4$, F_n has only four eigenspaces, and since F_n is unitarily diagonalizable, F_n has n mutually orthonormal eigenvectors, and hence the four eigenspaces of F_n are mutally orthogonal and their direct-sum is equal to $\mathcal{C}^n$. For larger values of n, these eigenspaces have high dimension, and there are many choices for bases for these eigenspaces; the union of these basis sets, whatever they are chosen to be, form a complete set of n linearly-independent eigenvectors of F_n.

Exercise 3.17: Show that the eigenvalues of $\sqrt{2}F_2$ are 1 and -1, and the eigenvalues of $\sqrt{3}F_3$ are 1, -1 and $-i$.

Exercise 3.18: Show that for n odd, the characteristic polynomial of the unitary matrix $\sqrt{n}F_n$ is $(x^4 - 1)^{\lfloor n/4 \rfloor}(x - 1)$ for $n \bmod 4 = 1$, and is $(x^4 - 1)^{\lfloor n/4 \rfloor}(x - 1)(x + 1)(x + i)$ for $n \bmod 4 = 3$. Hint: Assume n is a positive odd integer and form the polynomial whose roots are 1 with multiplicity $m + 1$, -1 with multiplicity $m + \delta_{r,3}$, i with multiplicity m, and $-i$ with multiplicity $m + \delta_{r,3}$, with $m = \lfloor n/4 \rfloor$ and $r = n \bmod 4$.

3.2 Aliasing

Let $x \in L^2(Q)$ so that x is of period p. By sampling x at n equally-spaced points in each period, we may also consider x as a member of $d_n(T\mathcal{Z})$ where $T = p/n$. Then we have the *aliasing relation*:

$$x^{\wedge(p;n)}(h/p) = \sum_{m=-\infty}^{\infty} x^{\wedge(p)}((h + mn)/p).$$

Thus, if $x^{\wedge(p)}$ is 0 outside the discrete interval $[-\lfloor n/2\rfloor/p, \ldots, (\lceil n/2\rceil - 1)/p]$, then $x^{\wedge(p;n)} = x^{\wedge(p)}$, but otherwise, $x^{\wedge(p;n)}(h/p)$ is a sum of $x^{\wedge(p)}(h/p)$ and various *aliases*, which are the values of $x^{\wedge(p)}$ defining the complex amplitudes at higher frequencies: $x^{\wedge(p)}((h/p) \pm (mn))/p)$ for $\mid m \mid \geq 1$. (Remember $x^{\wedge(p)}(k/p) \to 0$ as $|k| \to \infty$.)

Exercise 3.19: Show that $x^{\wedge(p;n)}(h/p) = \displaystyle\sum_{m=-\infty}^{\infty} x^{\wedge(p)}((h+mn)/p)$.

Solution 3.19: Let $p = nT$ where n is a positive integer. The Fourier inversion theorem for the periodic function $x \in L^2(Q)$ states:

$$x(t) = \sum_{\infty \leq j \leq -\infty} x^{\wedge(p)}(j/p) e^{2\pi i (j/p)t},$$

so in particular,

$$x(kT) = \sum_{\infty \leq j \leq -\infty} x^{\wedge(p)}(j/p) e^{2\pi i (j/p)kT}.$$

Moreover,

$$x(kT) = \sum_{j} x^{\wedge(p)}(j/p) e^{2\pi i (j/p)kT} = \sum_{0 \leq h < n} \sum_{m} x^{\wedge(p)}((h+mn)/p) e^{2\pi i ((h+mn)/p)kT}.$$

We may write $e^{2\pi i ((h+mn)/p)kT} = e^{2\pi i (h/p)kT} e^{2\pi i (mn/p)kT}$, and since $p = nT$, $e^{2\pi i (mn/p)kT} = 1$, so:

$$x(kT) = \sum_{0 \leq h < n} \left[\sum_{m} x^{\wedge(p)}((h+mn)/p) \right] e^{2\pi i (h/p)kT}.$$

Now, the Fourier inversion theorem for the discrete periodic function $x \in d_n(T\mathcal{Z})$ (which is a *sampled* form of $x \in L^2(Q)$) states:

$$x(kT) = \sum_{0 \leq h < n} x^{\wedge(p;n)}(h/p) e^{2\pi i (h/p)kT},$$

and equating coefficients of $e^{2\pi i (h/p)kT}$ in these two identically-valued sums, we have $x^{\wedge(p;n)}(h/p) = \sum_{m=-\infty}^{\infty} x^{\wedge(p)}((h+mn)/p)$. (Strictly, we need to compute the inner-product $(\cdot, e^{2\pi i (j/p)kT})$ of our two sums for $j = 0, 1, \ldots, n-1$ to extract the matching coefficients, and then assert that they are identical.)

The discrete Fourier transform $x^{\wedge(p;n)}$ thus approximates the Fourier transform $x^{\wedge(p)}$. This approximation is good to the extent that $x^{\wedge(p;n)}$ is not excessively contaminated by aliasing. To be sure that $x^{\wedge(p;n)}$ is a good approximation to $x^{\wedge(p)}$ it is necessary that the sampling rate n/p has been chosen large enough so that no significant high-frequency components of x are missed; that is $x^{\wedge(p)}$ must be nearly zero outside the discrete band $[-\lfloor n/2\rfloor/p, \ldots, (\lceil n/2\rceil - 1)/p]$; we correctly deal with this band by sampling at the rate n/p.

In particular, if x is band-limited, then we can specify a bound for the error in $x^{\wedge(p;n)}$ as $|x^{\wedge(p;n)}(h/p) - x^{\wedge(p)}(h/p)| \le O(2ce^{-cn})$, where c is a constant, and, of course, when $x^{\wedge(p)}$ is 0 outside the discrete band $[-\lfloor n/2 \rfloor/p, \ldots, (\lceil n/2 \rceil - 1)/p]$, the approximation is perfect.

Note if x is the period-p periodic extension of a function, y, defined on $[0,p]$, such that $y(0) \ne y(p)$ (so that $x(0) \ne \lim_{\epsilon\downarrow 0} y(p-\epsilon) = x(p)$,) then artificial discontinuities have been introduced and x is certainly not band-limited, so some aliasing error must occur in $x^{\wedge(p;n)}$. We may avoid the error due to an introduced discontinuity by computing $x^{\wedge}_{ep}$ instead of $x^{\wedge}$, but of course, this also introduces aliasing error.

3.3 Structural Relations

Let $\wedge$ denote the discrete Fourier transform $\wedge(p;n)$, let $\vee$ denote the inverse discrete Fourier transform $\vee(p;n)$, and let x be periodic with period $p = nT$. Note that here $\vee$ is not the inverse operator of $\wedge$ (unless the integer $n = p^2$.)

- If x is real, then $x^{\wedge}$ is Hermitian, *i.e.*, $x^{\wedge}(-k/p) = x^{\wedge}(k/p)^*$ for $k \in \mathcal{Z}$. This is summarized by writing $x^{\wedge R} = x^{\wedge *}$ where $y^R(s) = y(-s)$. Note if x is real, $x^{\wedge}(0)$ is real; also, if n is even, $x^{\wedge}(-n/(2p))$ is real.
- If x is imaginary, then $x^{\wedge}$ is anti-Hermitian, *i.e.*, $x^{\wedge R} = -x^{\wedge *}$
- x is even, if and only if $x^{\wedge}$ is even, and x is odd, if and only if $x^{\wedge}$ is odd. Thus, if x is real and even, or imaginary and odd, $x^{\wedge}$ must be real.
- The operator pairs $(*, R)$, $(\wedge, R)$, $(\vee, R)$, and $(\wedge(p;n), \vee(n/p;n))$ commute, but $(\vee, *)$ and $(\wedge, *)$ do not. Thus $x^{*R} = x^{R*}$, $x^{R\wedge} = x^{\wedge R}$, and $x^{R\vee} = x^{\vee R}$. However $x^{\wedge(p;n)*} = (T/p)x^{*\vee(p;n)} = x^{*\wedge(p;n)R}$, $x^{*\wedge(p;n)} = (T/p)x^{\vee(p;n)*} = x^{\wedge(p;n)*R}$, $x^{\vee(p;n)*} = (p/T)x^{*\wedge(p;n)} = x^{*\vee(p;n)R}$, and $x^{*\vee(p;n)} = (p/T)x^{\wedge(p;n)*} = x^{\vee(p;n)*R}$.

Exercise 3.20: Show that $x^{\wedge(p;n)R}(s) = x^{R\wedge(p;n)}(s)$ for $s \in \{\ldots, -1/p, 0, 1/p, \ldots\}$.

Solution 3.20: For $s \in \{\ldots, -1/p, 0, 1/p, \ldots\}$, we have:

$$\begin{aligned} x^{\wedge(p;n)R}(s) &= \left[\frac{T}{p}\sum_{0\le h\le n-1} x(hT)e^{-2\pi i s hT}\right]^R \\ &= \frac{T}{p}\sum_{0\le h\le n-1} x(hT)e^{-2\pi i(-s)hT} \\ &= \frac{T}{p}\sum_{1\le h\le n} x((n-h)T)e^{-2\pi i(-s)(n-h)T}. \end{aligned}$$

And $x(-hT) = x((n-h)T)$ and $e^{2\pi i s nT} = 1$, for s an integer multiple of $1/(nT)$, so:

$$x^{\wedge(p;n)R}(s) = \frac{T}{p}\sum_{1\le h\le n} x(-hT)e^{-2\pi i s hT} = x^{R\wedge(p;n)}(s).$$

- We have $x^{\wedge(p;n)} = (T/p)x^{\vee(p;n)R}$, so $x^{\vee(p;n)} = (p/T)x^{\wedge(p;n)R}$, and $x^{\wedge(p;n)\wedge(n/p;n)} = (T/p)x^R$, and thus $(p^2/T^2)x^{\wedge(p;n)\wedge(n/p;n)\wedge(p;n)\wedge(n/p;n)} = x$.

- $x(at)^{\wedge(p/|a|;n)}(as) = x(t)^{\wedge(p;n)}(s)$ where s is an integral multiple of $1/p$.

  Exercise 3.21: Let $x \in d_n(T\mathcal{Z})$ where $T = p/n$. Show that $(x(at))^{\wedge(p/|a|;n)}(ka/p) = x^{\wedge(p;n)}(k/p)$ for $k \in \mathcal{Z}$.

  Solution 3.21: Let $y(t) = x(at)$, y is a discrete periodic function with period $p/|a|$ and stepsize $T/|a|$. Thus $(y(t))^{\wedge(p/|a|;n)}(kr) = \dfrac{1}{n}\sum_{h=0}^{n-1} y(hv)e^{-2\pi ikrhv}$ where $v = T/|a|$ and $r = |a|/p$.

  Thus

$$\begin{aligned}
(y(t))^{\wedge(p/|a|;n)}(k|a|/p) &= \frac{1}{n}\sum_{h=0}^{n-1} x(ahT/|a|)e^{-2\pi ikh(T/p)} \\
&= \frac{1}{n}\sum_{h=0}^{n-1} x(sign(a)hT)e^{-2\pi i(k/p)hT} \\
&= (x(sign(a)t))^{\wedge(p;n)}(k/p).
\end{aligned}$$

  If $a > 0$, $y^{\wedge(p/|a|;n)}(ka/p) = (x(at))^{\wedge(p/|a|;n)}(ka/p) = x^{\wedge(p;n)}(k/p)$.

  If $a < 0$, then

$$\begin{aligned}
y^{\wedge(p/|a|;n)}(k(-a)/p) &= (x(at))^{\wedge(p/|a|;n)}(-ka/p) \\
&= (x(-t))^{\wedge(p;n)}(k/p) \\
&= \frac{1}{n}\sum_{h=0}^{n-1} x((n-h)T)e^{-2\pi i(k/p)hT} \\
&= \frac{1}{n}\sum_{h=1}^{n} x(hT)e^{-2\pi i(k/p)(n-h)T} \\
&= \frac{1}{n}\sum_{h=1}^{n} x(hT)e^{-2\pi ik+2\pi i(k/p)hT} \\
&= \frac{1}{n}\sum_{h=1}^{n} x(hT)e^{-2\pi i(-k/p)hT} \\
&= x^{\wedge(p;n)}(-k/p) \text{ for } k \in \mathcal{Z}.
\end{aligned}$$

  Thus, $(x(at))^{\wedge(p/|a|;n)}(ka/p) = x^{\wedge(p;n)}(k/p)$ for $k \in \mathcal{Z}$.

  This is consistent with the identity $(x(at))^{\wedge(p/|a|)}(s) = x^{\wedge(p)}(s/a)$ for $s \in \{\ldots, -\dfrac{|a|}{p}, 0, \dfrac{|a|}{p}, \ldots\}$ in the situation where x is defined on $[0, p)$ and periodically extended to $\mathcal{R}$.

- $x(t + t_0)^{\wedge}(s) = e^{2\pi it_0 s}x^{\wedge}(s)$, where t_0 is a multiple of $p/n = T$.

- $(e^{-2\pi i s_0 t}x(t))^\wedge(s) = x^\wedge(s+s_0)$, where s_0 is a multiple of $1/p$.
- $x^\wedge(0) = \sum_{h=0}^{n-1} x(hT)/n$, and $x(0) = \sum_{h=-\lfloor n/2\rfloor}^{\lceil n/2\rceil-1} x^\wedge(h/p)$.

3.4 Discrete Circular Convolution

Let x and y be discrete periodic functions with period p and stepsize $T = p/n$. By analogy with the circular convolution of two periodic period-p functions on the real line, we define the discrete circular convolution of the two discrete periodic period-p functions, x and y, defined on $\ldots$, $-2T$, $-T$, 0, T, $2T$, $\ldots$, as:

$$(x \circledast y)(rT) := (1/n)\sum_{h=0}^{n-1} x(hT)y(rT-hT).$$

This is just a Riemann sum approximation of the circular convolution integral for period-p functions in $L^2(Q)$. When there is danger of confusion, we shall write $\circledast_d$ to denote the discrete circular convolution operator.

Also, we define the discrete cross-correlation kernel function as:

$$(x \otimes y)(rT) := (1/n)\sum_{h=0}^{n-1} x(hT)y(rT+hT).$$

The following relations hold.

$$\begin{aligned}
(x \circledast y)^\wedge &= x^\wedge y^\wedge, \quad \text{so } x \circledast y = (x^{\wedge(p;n)}y^{\wedge(p;n)})^{\vee(p/n;n)}\\
(xy)^\wedge &= x^\wedge \circledast y^\wedge\\
(xy)^{\vee(p;n)} &= \frac{1}{n}x^{\wedge(p;n)} \circledast y^{\wedge(p;n)}\\
(x \otimes y)^\wedge &= x^\wedge y^{\wedge R}
\end{aligned}$$

Note when x is real, $(x \otimes x)^\wedge = |x^\wedge|^2$.

Exercise 3.22: Show that $(x \circledast y)^\wedge(s) = x^\wedge(s)y^\wedge(s)$, where x and y are discrete periodic functions with period $p = nT$ and stepsize T, and where $x^\wedge(s) = \sum_{0\le h\le n} x(hT)e^{-2\pi i s hT}$ and $y^\wedge(s) = \sum_{0\le h\le n} y(hT)e^{-2\pi i s hT}$ with $s = \ldots, -1/p, 0, 1/p, \ldots$ are discrete periodic functions with period $n/p = 1/T$ and stepsize $1/p = 1/(nT)$.

Solution 3.22:

$$(x \circledast y)^\wedge(s) \quad = \quad \frac{T}{p}\sum_{0\le h<n}\left[\frac{T}{p}\sum_{0\le k<n} x(kT)y(hT-kT)\right]e^{-2\pi i s hT}$$

$$
\begin{aligned}
&= \frac{T}{p}\sum_{0\le h<n}\left[\frac{T}{p}\sum_{0\le k<n} x(kT)y(hT-kT)\right] e^{-2\pi is(h-k)T}e^{-2\pi iskT} \\
&= \left[\frac{T}{p}\sum_{0\le k<n} x(kT)e^{-2\pi iskT}\frac{T}{p}\sum_{0\le h<n} y(hT-kT)\right] e^{-2\pi is(h-k)T} \\
&= \left[\frac{T}{p}\sum_{0\le k<n} x(kT)e^{-2\pi iskT}\right]\left[\frac{T}{p}\sum_{0\le m<n} y(mT)e^{-2\pi ismT}\right] \\
&= x^\wedge(s)y^\wedge(s).
\end{aligned}
$$

Exercise 3.23: Show that $(xy)^\wedge = x^\wedge \circledast y^\wedge$.

Solution 3.23:

$$
\begin{aligned}
x(s)y(s) &= \left[\sum_{0\le k<n} x^\wedge(k/p)e^{2\pi i(k/p)s}\right]\left[\sum_{0\le h<n} y^\wedge(h/p)e^{2\pi i(h/p)s}\right] \\
&= \sum_{0\le k<n} x^\wedge(k/p)e^{2\pi i(k/p)s}\sum_{0\le h<n} y^\wedge(h/p-k/p)e^{2\pi i(h/p-k/p)s} \\
&= \sum_{0\le h<n}\sum_{0\le k<n} x^\wedge(k/p)y^\wedge(h/p-k/p)e^{2\pi i(h/p)s} \\
&= (x^\wedge \circledast y^\wedge)^\vee,
\end{aligned}
$$

so $(xy)^\wedge = x^\wedge \circledast y^\wedge$.

Exercise 3.24: Let V be the $n\times n$ matrix such that $V_{jk} = \begin{cases} 1, & \text{if } j+k=n+1; \\ 0, & \text{otherwise,}\end{cases}$ and let $C_\rightarrow$ be the $n\times n$ matrix such that $(C_\rightarrow)_{jk} = \begin{cases} 1, & \text{if } k = 1+(j \bmod n); \\ 0, & \text{otherwise.}\end{cases}$

Now let $N_r = VC_\rightarrow^{r+1}$ and let $x, y \in d_n(T\mathcal{Z})$ be period-p, stepsize-T, discrete functions with $p = nT$.

Show that $(x \circledast y)(rT) = c_{p,n}(x)N_r c_{p,n}(y)^{\mathrm{T}}$ for $r \in \{0, 1, \ldots, n-1\}$, where $c_{p,n}(x) = (x(0), x(T), \ldots, x((n-1)T)) \in \mathcal{C}^n$. Hint: $C_\rightarrow$ is the $n\times n$ permutation matrix such that for any $n\times n$ matrix M, $MC_\rightarrow = M$ col $(n, 1, 2, \ldots, n-1)$, and $C_\rightarrow^n = I$.

Note that $N_r = N_r^{\mathrm{T}}$ implies that $x \circledast y = y \circledast x$.

Also show that $c_{p,n}(x \circledast y) = c_{p,n}(x)Y$, where Y is the $n\times n$ *Toeplitz matrix* defined by $Y_{jk} = y(jT-kT)$ for $1 \le j,k \le n$. Hint: remember that $y(-kT) = y((n-k)T)$.

Discrete circular convolution provides a fast way to multiply polynomials. Given two polynomials $A(z) = a_0 + a_1 z + \ldots + a_{n-1}z^{n-1}$ and $B(z) = b_0 + b_1 z + \ldots + b_{n-1}z^{n-1}$ of degree at most $n-1$, the product $A(z)B(z) = c_0 + c_1 z + \ldots + c_{2n-2} z^{2n-2}$ has the coefficients

$$c_j = \sum_{k=0}^{\min(j,n-1)} a_k b_{j-k}, \quad \text{where} \quad a_t = b_t = 0 \quad \text{for} \quad t \geq n.$$

Thus if we define the vectors $a = [a_0, \ldots, a_{n-1}, 0, \ldots, 0]$, $b = [b_0, \ldots, b_{n-1}, 0, \ldots, 0]$, and $c = [c_0, \ldots, c_{2n-1}]$ with $2n$ components, then $c = (aF_{2n} \odot bF_{2n})F_{2n}^{-1}$, where F_{2n} is the $2n \times 2n$ discrete Fourier transformation matrix, and where $\odot$ denotes the operation of element-by-element multiplication of two vectors, producing another vector. Equivalently, $c = (a^{\wedge} \odot b^{\wedge})^{\vee}$. Note $c_{2n-1} = 0$. This computation of c is fast because there is an algorithm called 'the fast Fourier transform' algorithm (discussed below) that computes aF in $O(2n \log(2n))$ steps when n is a power of 2; when n is not a power of 2, we can take the coefficient sequences a, b, and c as $2 \cdot 2^k$-vectors where $k = \lceil \log_2 n \rceil$, *i.e.*, we extend our length-n coefficient sequences to the least power of 2 greater than or equal to $2n$.

3.5 Interpolation

Let x be a complex-valued periodic function in $L^2(Q)$, so that x is of period p. Recall that the k-th partial sum $S_k(t) := \sum_{-k \leq j \leq k} x^{\wedge(p)}(j/p) e^{2\pi i (j/p) t}$ of the Fourier series of $x(t)$ can be summed to yield

$$S_k(t) = \int_0^p D_k(y) x(t-y)\, dy = (D_k \circledast x)(t)$$

where, for $v \in [0, p)$,

$$D_k(v) = \begin{cases} 2k+1 & \text{if } v = 0 \\ \dfrac{sin(\pi(2k+1)v/p)}{sin(\pi v/p)} & \text{otherwise.} \end{cases}$$

The Dirichlet kernel, D_k, is extended to be periodic with period p by defining $D_k(v + mp) = D_k(v)$ for $v \in [0, p)$ and $m \in \mathcal{Z}$.

Although $D_k(t)$ is defined on the entire real line, as a discrete function, D_k is restricted to the discrete domain $\{-\lfloor n/2 \rfloor T, \ldots, -T, 0, T, \ldots, (\lceil n/2 \rceil - 1)T\}$ where $T = p/n$, and extended periodically to the domain $\{\ldots, -T, 0, T, \ldots\}$ by taking $D_k(jT) = D_k((j \mod n)T)$. We will generally want to use the discrete sampled form of D_k with $n = 2k+1$, so that $\{-kT, \ldots, -T, 0, T, \ldots, kT\}$ spans one period of the discrete form of D_k.

If $x \in L^2(Q)$ is band-limited, so that $x^{\wedge(p)}(s) = 0$ for $|s| > k/p$, with k a fixed non-negative integer, then $x = S_k$. (Note it could be that $x^{\wedge(p)}(s) = 0$ for $|s| > m/p$ with $m < k$, in which case our choice of k is not the least possible.) Moreover if we define the discrete periodic function X with period p and stepsize $T = p/(2k+1)$ so that $X(hT) = x(hT)$ for $h = \cdots, -1, 0, 1, \cdots$, then $S_k = X^{\wedge(p;2k+1)\vee((2k+1)/p;2k+1)}$ where $p = (2k+1)T$, since in this case, the aliasing relation states that $x^{\wedge(p)} = X^{\wedge(p;2k+1)}$ and, since $x^{\wedge(p)}(s) = 0$ for $|s| > k/p$, $x^{\wedge(p)\vee(p)} = X^{\wedge(p;2k+1)\vee((2k+1)/p;2k+1)}$. The discrete function X is just the stepsize-T sampled form of the periodic period p function x.

Thus, $x = X^{\wedge(p;2k+1)\vee((2k+1)/p;2k+1)}$, allowing the argument of X to range over the real line. This shows that a band-limited periodic function is completely determined by an odd number of sufficiently closely-spaced samples over one period. In particular, the sampling stepsize must be no greater than $T = p/(2k+1)$, where k is chosen to be the least non-negative integer such that x has no spectral components outside the frequency band $[-k/p, k/p]$. The required rate of sampling is thus at least $(2k+1)/p$ samples per unit of t, so the required rate of sampling must be greater than the sampling frequency $2k/p$ samples per unit of t, called the *Nyquist sampling frequency*, at which aliasing error can begin to appear. When such samples $X(0)$, $X(T)$, ..., $X(2vT)$ are known, with the integer $v \geq k$ and the value T redefined as $T := p/(2v+1)$, the unique band-limited interpolating function x that satisfies $x(hT) = X(hT)$ can be reconstituted via the discrete Fourier transform as $X^{\wedge(p;2v+1)\vee((2v+1)/p;2v+1)}$, or directly as

$$x(t) = \frac{1}{n} \sum_{0 \leq j < n} X(jT)\, sin(\pi(t/T - j))/\, sin(\pi(t/T - j)/n),$$

where $n = 2v+1, T = p/n$, and $sin(0)/0$ is taken as n. Note the above sum can instead be taken over $-v \leq j \leq v$. This identity is known as Whittaker's interpolation formula.

Exercise 3.25: Let $v \in \mathcal{Z}$ with $v \geq 0$. Show that, if the period-p function $x \in L^2(Q)$ is band-limited with $x^{\wedge(p)}(s) = 0$ for $|s| > v/p$, then

$$x(t) = \frac{1}{n} \sum_{0 \leq j < n} x(jT)\, sin(\pi(t/T - j))/\, sin(\pi(t/T - j)/n),$$

with $n = 2v+1, T = p/n$, and $sin(0)/0$ taken as n.

Solution 3.25:

$$\begin{aligned} x(t) &= x^{\wedge(p;n)\vee(n/p;n)}(t) \\ &= \sum_{0 \leq h < n} \left[\frac{T}{p} \sum_{0 \leq j < n} x(jT) e^{-2\pi i h j T/p} \right] e^{2\pi i (h/p) t} \\ &= \frac{1}{n} \sum_{0 \leq j < n} x(jT) \sum_{0 \leq h < n} e^{-2\pi i h (t - jT)/p}. \end{aligned}$$

And $\sum_{0 \leq h < n} e^{-2\pi i h u/p} = sin(\pi n u/p)/\, sin(\pi u/p)$ with $sin(0)/0$ taken as n; this sum is just the Dirichlet kernel $D_{\lfloor n/2 \rfloor}(u)$.

Thus,

$$\begin{aligned} x(t) &= \frac{1}{n} \sum_{0 \leq j < n} x(jT)\, sin(\pi n(t - jT)/p)/\, sin(\pi(t - jT)/p) \\ &= \frac{1}{n} \sum_{0 \leq j < n} x(jT)\, sin(\pi(t/T - j))/\, sin(\pi(t/T - j)/n) \\ &= (x \circledast_d D_{\lfloor n/2 \rfloor})(t), \end{aligned}$$

where $n/p = 1/T$ and $n = 2v + 1$.

Let x be a discrete complex-valued period-p stepsize-T function with $n = p/T \in \mathcal{Z}^+$. Consider the polynomial $V(z) = x(0) + x(T)z + \ldots + x((n-1)T)z^{n-1}$ whose coefficient vector is $x = [x(0), \ldots, x((n-1)T)]$. The discrete inverse Fourier transform $\vee(p;n)$ of x, tabulated in a vector, is $[x^\vee(0), x^\vee(1/p), \ldots, x^\vee((n-1)/p)] = xF_n^{-1}$, where F_n is the $n \times n$ discrete Fourier transform matrix. The vector xF_n^{-1} is just $[V(1), V(w), V(w^2), \ldots, V(w^{n-1})]^{\mathrm{T}}$ where $w = e^{2\pi i/n}$. Thus, given values of $V(z)$ for $z = 1, w, w^2, \ldots, w^{n-1}$, as the components of a vector y, $yF_n = x$ is just the vector of the coefficients of the polynomial V of degree $n-1$ which interpolates the n given points $(w^k, V(w^k))$ for $k = 0, 1, \cdots, n-1$. This polynomial is, of course, the Lagrange interpolating polynomial of degree $n-1$ for the n given data points equally-spaced on the unit circle in the complex plane.

3.6 The Fast Fourier Transform

The fast Fourier transform algorithm is a method to numerically compute $nx^\wedge$, given values of the discrete periodic function x on $0, 1, \ldots, n-1$, where x is of period n and stepsize 1; thus here the discrete period-n function x is represented with n "steps" and $\wedge$ denotes $\wedge(n;n)$. The result of the fast Fourier transform algorithm applied to the sequence of values $\langle x(0), \ldots, x(n-1)\rangle$ is the sequence of values $nx^\wedge(0)$, $nx^\wedge(1/n)$, $\ldots$, $nx^\wedge((n-1)/n)$. Since $x^\wedge$ is just a discrete periodic function with period 1 and stepsize $1/n$, this is just the sequence $nx^\wedge(0)$, $nx^\wedge(1/n)$, $\ldots$, $nx^\wedge((\lceil n/2\rceil - 1)/n)$, $nx^\wedge(-\lfloor n/2\rfloor/n)$, $\ldots$, $nx^\wedge(-1/n)$, which can be reordered into its corresponding natural order by "swapping" the first and last halves.

The fast Fourier transform algorithm is "fast" only when n is highly composite. It is particularly convenient to choose n to be a power of 2. We can develop the formula that characterizes the fast Fourier transform algorithm as follows. Given $x(0), x(1), \ldots, x(n-1)$, where x is of period n, we have $nx^\wedge(\frac{j}{n}) = \sum_{0 \le k \le n-1} x(k)w_n^{jk}$ for $j \in \{0, 1, \ldots, n-1\}$, where $w_n = e^{-2\pi i/n}$, a primitive nth root of unity. (Note the function w_n^j as a function of $j \in \mathcal{Z}$ is a discrete period-n function.) Let $n = uv$, where u and v are positive integers. Then

$$nx^\wedge(\frac{j}{n}) = \sum_{h=0}^{v-1}\sum_{g=0}^{u-1} x(hu+g)w_n^{j(hu+g)} = \sum_{g=0}^{u-1}\sum_{h=0}^{v-1} x(hu+g)w_n^{jhu}w_n^{jg}.$$

Note that $w_n^{jhu} = w_v^{jh}$. Thus:

$$nx^\wedge(\frac{j}{n}) = \sum_{g=0}^{u-1} w_n^{jg} \sum_{h=0}^{v-1} x(hu+g)w_v^{jh}.$$

This latter identity shows that $nx^\wedge(\frac{j}{n})$ is the sum of u discrete Fourier transforms of period-v functions defined by u evenly-spaced length-v transform sums $x(0\cdot u+g)w_v^{j0} + x(1\cdot u+g)w_v^{j1} + \cdots + x((v-1)\cdot u+g)w_v^{j(v-1)}$ for $g = 0, 1, \ldots, u-1$, weighted by the complex oscillations 1, w_n^j, w_n^{2j}, $\ldots$,

$w_n^{(u-1)j}$. If v is composite, this formula can be applied recursively to each of the u transforms that construct the final result.

In particular, for n a power of two, we can apply the fast Fourier transform to $\langle x(0), x(2), \ldots, x(n-2)\rangle$ to obtain the sequence $\langle a_0, a_1, \ldots, a_{n/2-1}\rangle$, and we can apply the fast Fourier transform to $\langle x(1), x(3), \ldots, x(n-1)\rangle$ to obtain $\langle b_0, b_1, \ldots, b_{n/2-1}\rangle$. And, given the result sequences $\langle a_0,\ a_1,\ \ldots,$ $a_{(n/2)-1}\rangle$ and $\langle b_0, b_1, \ldots, b_{(n/2)-1}\rangle$, we have $nx^\wedge(j/n) = a_j + b_j e^{-2\pi i j/n}$. This follows by specializing the fast Fourier transform formula above to obtain:

$$nx^\wedge(j/n) = \sum_{k=0}^{(n/2)-1} [x(2k)w_{n/2}^{jk} + (x(2k+1)w_{n/2}^{jk})w_n^j] = a_j + b_j w_n^j,$$

for $j \in \{0, 1, \ldots, n-1\}$, where n is a power of 2. The sequences a and b represent period-$\dfrac{n}{2}$ discrete functions, so for $j \geq n/2$, the sequences a and b are extended by $a_j = a_{(j \bmod (n/2))}$ and $b_j = b_{(j \bmod (n/2))}$.

If we have n equally-spaced "samples" $x(0)$, $x(1)$, …, $x(n-1)$ over one period, where n is not a power of 2, we can interpolate this data, perhaps with a cubic spline interpolation function, and re-sample to increase n to be a power of 2; alternately we can extend the sequence x to be a power-of-2 length as discussed on page 21.

Exercise 3.26: Show that $nx^\wedge(j/n) = \sum_{k=0}^{(n/2)-1} [x(2k)w_{n/2}^{jk} + (x(2k+1)w_{n/2}^{jk})w_n^j] = a_j + b_j w_n^j$, for $0 \leq j < n$, where n is a power of 2 and the sequences a and b are the discrete Fourier transform sequences defined above.

Solution 3.26: Assume $n = uv$ where $u = 2$. For $j \in \{0, 1, \ldots, n-2\}$ we have

$$\begin{aligned}
nx^\wedge(\frac{j}{n}) &= \sum_{g=0}^{1} w_n^{jg} \sum_{h=0}^{(n/2)-1} x(2h+g)w_{n/2}^{jh} \\
&= \sum_{h=0}^{(n/2)-1} w_{n/2}^{jh} \sum_{g=0}^{1} w_n^{jg} x(2h+g) \\
&= \sum_{h=0}^{(n/2)-1} w_{n/2}^{jh} \left[w_n^{0j} x(2h) + w_n^{1j} x(2h+1)\right] \\
&= \left[\sum_{k=0}^{(n/2)-1} x(2k)w_{n/2}^{jk}\right] + \left[\sum_{k=0}^{(n/2)-1} (x(2k+1)w_{n/2}^{jk}\right] w_n^j \\
&= a_j + b_j w_n^j.
\end{aligned}$$

In order to compute the sequences $\langle a_0, a_1, \ldots, a_{n/2-1}\rangle$ and $\langle b_0, b_1, \ldots, b_{n/2-1}\rangle$, we can use the fast Fourier transform recursively. By recursively using the FFT on every sequence of more than 2 values, we obtain the full fast Fourier transform algorithm for the case where n is a power of two.

Here is a program for this recursive form of the FFT alogrithm for n a power of 2.

```
complex array address FFT(complex array x[0:n-1]; integer n; integer s):
  "The basic discrete direct n-scaled or inverse Fourier transform
   of the data in x is computed and the address of the result complex
   array is returned.  If s=1, the direct n-scaled transform is returned;
   if s=-1, the inverse transform is returned."
{static integer j, k; static real f; complex w, u;
 complex array address a,b;

 allocate-space for complex array xh[0:n-1];
 if (n < 1) goto exit;
 if (n = 1) {xh[0]←x[0]; goto exit;}

 allocate-space for complex array xa[0:n/2-1];
 allocate-space for complex array xb[0:n/2-1];

 w←1; f←2π/n; u←cos(f)-i*s*sin(f); "u = exp(-s2πi/n)"

 for j←0:n-1:2 do {xa[j]←x[j]; xb[j]←x[j+1];};
 if (n = 2) {a←address(xa); b←address(xb); goto finish;}

 a←FFT(xa, n/2, s);
 b←FFT(xb, n/2, s);

 free-space xa,xb;
 finish:
 for j←0:n-1 do {k←j mod (n/2); xh[j]←a[k]+b[k]*w; w←w*u;};
 free-space a,b;
 exit:
 return(address(xh));
}
```

The notation "`for j←0:n-1:2`" represents "`for j ← 0,2, ..., ⌊(n-1)/2⌋`", *i.e.*, "from 0 to n-1 in steps of size 2."

Exercise 3.27: What is the memory space requirement of this recursive FFT algorithm? Hint: count the maximum number of complex-numbers that may exist simultaneously during execution.

Exercise 3.28: Revise this recursive FFT algorithm to use the arguments: `complex array x[0:n-1]; integer starti, stepi, leni, s`, where `fft(x, starti, stepi, leni, s)` computes the scaled transform (or inverse transform) of the sequence `x[starti], x[starti+stepi], x[starti+2*stepi], ..., x[leni-starti+1]`. This permits you to avoid the use of the arrays `xa` and `xb`.

This recursive algorithm can be "unwrapped" into an iterative form known as the power-of-two

Cooley-Tukey algorithm which can leave its output in the input array `x` of $2n$ real values when this is advantageous, and uses a total of $(n-3)\log_2 n$ complex multiplications and $n\log_2 n$ complex additions, plus $(\log_2 n - 1)$ sine and cosine evaluations. Here n must be a power of 2.

In order to understand the "unwrapping" that leads to the Cooley-Tukey algorithm, let us look at an example with $n = 8$. We have the input sequence $\langle x(0), x(1), x(2), x(3), x(4), x(5), x(6), x(7)\rangle$. The recursive algorithm first computes the scaled discrete transforms of $\langle x(0), x(2), x(4), x(6)\rangle$ and $\langle x(1), x(3), x(5), x(7)\rangle$ and combines them. To do this, the scaled transforms of $\langle x(0), x(4)\rangle$ and $\langle x(2), x(6)\rangle$ and $\langle x(1), x(5)\rangle$ and $\langle x(3), x(7)\rangle$ are computed and combined in pairs. Finally, computing these four scaled transforms is done, conceptually, by computing the trivial single-element transforms of $\langle x(0)\rangle$, $\langle x(4)\rangle$, $\langle x(2)\rangle$, $\langle x(6)\rangle$, $\langle x(1)\rangle$, $\langle x(5)\rangle$, $\langle x(3)\rangle$, and $\langle x(7)\rangle$ and combining them in pairs. (The single-element discrete Fourier transform is just the identity operator.)

In tabular form we have:

level 0	$\langle x(0)\rangle^\wedge$	$\langle x(4)\rangle^\wedge$	$\langle x(2)\rangle^\wedge$	$\langle x(6)\rangle^\wedge$	$\langle x(1)\rangle^\wedge$	$\langle x(5)\rangle^\wedge$	$\langle x(3)\rangle^\wedge$	$\langle x(7)\rangle^\wedge$
level 1	$\langle x(0)$	$x(4)\rangle^\wedge$	$\langle x(2)$	$x(6)\rangle^\wedge$	$\langle x(1)$	$x(5)\rangle^\wedge$	$\langle x(3)$	$x(7)\rangle^\wedge$
level 2	$\langle x(0)$	$x(2)$	$x(4)$	$x(6)\rangle^\wedge$	$\langle x(1)$	$x(3)$	$x(5)$	$x(7)\rangle^\wedge$
level 3	$\langle x(0)$	$x(1)$	$x(2)$	$x(3)$	$x(4)$	$x(5)$	$x(6)$	$x(7)\rangle^\wedge$.

We combine each pair of adjacent length-2^k transform sequences at level k to produce the length-2^{k+1} transform sequences at level $k+1$. At level $\log_2 n$, we have the final length-n transform sequence which is our result.

It is instructive to rewrite our table with each index value given in binary form.

level 0	$\langle x(000)\rangle^\wedge$	$\langle x(100)\rangle^\wedge$	$\langle x(010)\rangle^\wedge$	$\langle x(110)\rangle^\wedge$	$\langle x(001)\rangle^\wedge$	$\langle x(101)\rangle^\wedge$	$\langle x(011)\rangle^\wedge$	$\langle x(111)\rangle^\wedge$
level 1	$\langle x(000)$	$x(100)\rangle^\wedge$	$\langle x(010)$	$x(110)\rangle^\wedge$	$\langle x(001)$	$x(101)\rangle^\wedge$	$\langle x(011)$	$x(111)\rangle^\wedge$
level 2	$\langle x(000)$	$x(010)$	$x(100)$	$x(110)\rangle^\wedge$	$\langle x(001)$	$x(011)$	$x(101)$	$x(111)\rangle^\wedge$
level 3	$\langle x(000)$	$x(001)$	$x(010)$	$x(011)$	$x(100)$	$x(101)$	$x(110)$	$x(111)\rangle^\wedge$.

Now, note that the index sequence at level 0 is just the sequence $rev_n(0)$, $rev_n(1)$, $rev_n(2)$, $rev_n(3)$, $rev_n(4)$, $rev_n(5)$, $rev_n(6)$, $rev_n(7)$ where, for $v \in \{0, 1, \ldots, n-1\}$, $rev_n(v)$ is the integer whose $\log_2 n$-bit binary form is the *reverse* of the $\log_2 n$-bit binary form of the integer v, *e.g.* $rev_8(3) = 6$ because 011 reversed is 110. You can see why this happens – at each recursive application of the FFT procedure, we segregate the even-index (low-order 0-bit) and odd-index (low-order 1-bit) elements of the input sequence into two sequences for further processing.

This same pattern applies for n an arbitrary power of 2. Thus, if we permute the elements of the input sequence $\langle x(0), x(1), \ldots, x(n-1)\rangle$, to obtain the sequence $\langle x(rev_n(0)), x(rev_n(1)), \ldots, x(rev_n(n-1))\rangle$, then starting at level 0 where we have n length-1 transform sequences, we can just combine adjacent pairs of length-2^k transform sequences in level k and replace each such pair by the resulting length-2^{k+1}-transform sequence. When we fill-in the values of the length-n transform sequence in level $\log_2 n$, we are done.

The Cooley-Tukey form of the fast Fourier algorithm for n a power of two is given as follows.

```
complex array address FFT(complex array x[0:n-1]; integer n; integer s):
  "The basic discrete direct n-scaled or inverse Fourier transform
   of the data in x is computed and the address of the result complex
   array is returned.  If s=1, the n-scaled direct transform is returned;
   if s=-1, the inverse transform is returned."
{integer a, b, j, k, h; complex t, w, u; real f;

 allocate complex array z[0:n-1];

 j←0;
 for k←0:n-1 do
 {z[k]←x[j];    "z = shuffled form of x array, j = bit reversal of k"
  h←n/2; while (h<=j) {j←j-h; h←h/2};
  j←j+h};

 a←1;
 while(a<n)
 {b←a+a; f←π/a; u←cos(f)-i*s*sin(f); w←1; "u = exp(-s2πi/n)"
  "compute the FFT of the n/b different length-b sequences.
   each is computed in 'a' steps (j=0, 1, ..., a-1), where each step
   computes two of the complex result values:  z[h] and z[k], that
   involve u^j."
  for j←0:a-1 do       "here w=u^j"
  {for h←j:n-1:b do {k←h+a; t←z[k]; z[k]←z[h]-w*t; z[h]←z[h]+w*t};
   w←w*u};
  a←b};
 return(z)
}
```

Note that $w_n^{j+n/2} = -w_n^j$, so only $(n-1)/2$ complex multiplications are involved in computing $a_j + b_j w_n^j$ for $0 \le j < n$, not counting obtaining the powers $w_n^1, w_n^2, \ldots, w_n^{n/2}$.

If we are given the sequence $\langle x(0), \ldots, x(n-1)\rangle = \langle y(t_0), y(t_0+T), \ldots, y(t_0+(n-1)T)\rangle$, where the discrete function y is of period nT, t_0 is an integral multiple of T, and n is a power of two, we can use the power-of-two fast Fourier transform algorithm given above to compute the sequence $\langle nx^\wedge(0), \ldots, nx^\wedge((n-1)/n)\rangle$, and then obtain $y^\wedge(h/(nT)) = e^{2\pi i t_0 h/n} x^\wedge(((h+n) \bmod n)/n)$ for $h = -\lfloor n/2\rfloor, \ldots, -1, 0, 1, \ldots, \lceil n/2\rceil - 1$.

Note the inverse discrete Fourier transform of $\langle x(0), \ldots, x(n-1)\rangle$ can be computed using the fast Fourier transform algorithm with the argument `s` $= -1$, or with `s` $= 1$ and the identity $x^{\vee(n;n)} = nx^{*\wedge(n;n)*}$, or alternately, the identities $x^{\vee(n;n)} = nx^{R\wedge(n;n)} = nx^{\wedge(n;n)R}$.

When x is a real discrete periodic function with period n and stepsize 1, $x^\wedge$ is a discrete periodic Hermitian function with period 1 and stepsize $1/n$. In this case, the relation $x^{\wedge R} = x^{\wedge *}$ stated elementwise is $x^{\wedge R}(k/n) = x^\wedge((n-k)/n) = x^\wedge(k/n)^*$ for $k \in \mathcal{Z}$, and similarly, if x is imaginary, we have $x^{\wedge R} = -x^{\wedge *}$, and stated elementwise, this is just $x^{\wedge R}((n-k)/n) = -x^\wedge(k/n)^*$ for $k \in \mathcal{Z}$.

Now let us consider the complex discrete period-n stepsize-1 function $z(j) = x(j) + iy(j)$ where x and y are real discrete period-n stepsize-1 functions. Let $u(j) = iy(j)$. Then $z^\wedge = x^\wedge + u^\wedge$, and $x^{\wedge R} = x^{\wedge *}$ and $u^{\wedge R} = -u^{\wedge *}$. Thus, $\frac{1}{2}(z^\wedge + z^{\wedge * R}) = x^\wedge$ and $\frac{1}{2}(z^\wedge - z^{\wedge * R}) = u^\wedge$ and $-iu^\wedge = y^\wedge$.

If we have two length-n real sequences, x and y, each consisting of n equally-spaced samples over one period of the associated real periodic function, then the above relations allow us to "pack" the two sequences together as $x + iy$, compute the n-scaled transform of $x + iy$, and recover the sequences $x^\wedge$ and $y^\wedge$.

Exercise 3.29: Let z be the complex discrete period-n stepsize-1 function defined by $z(j) = x(j) + iy(j)$ where x and y are real discrete period-n stepsize-1 functions. Show that $\frac{1}{2}(z^\wedge + z^{\wedge * R}) = x^\wedge$ and $\frac{1}{2}(z^\wedge - z^{\wedge * R}) = u^\wedge$ and $-iu^\wedge = y^\wedge$.

Solution 3.29: $\frac{1}{2}(z^\wedge(j/n) + z^{\wedge * R}(j/n)) = \frac{1}{2}(z^\wedge(j/n) + z^\wedge((n-j)/n)^*) = \frac{1}{2}[x^\wedge(j/n) + u^\wedge(j/n) + x^\wedge((n-j)/n)^* + u^\wedge((n-j)/n)^*]$.

But, $x^\wedge((n-j)/n)^* = x^\wedge(j/n)$ and $u^\wedge((n-j)/n)^* = -u^\wedge(j/n)$, since x is real and u is imaginary, so $\frac{1}{2}(z^\wedge(j/n) + z^{\wedge * R}(j/n)) = \frac{1}{2}[x^\wedge(j/n) + u^\wedge(j/n) + x^\wedge(j/n) - u^\wedge(j/n)] = x^\wedge(j/n)$.

Similarly, $\frac{1}{2}(z^\wedge - z^{\wedge * R}) = \frac{1}{2}[x^\wedge + u^\wedge - x^\wedge + u^\wedge] = u^\wedge$. And $u^\wedge = iy^\wedge$ so $y^\wedge = -iu^\wedge$.

Exercise 3.30: Show that if x and y are real discrete period-n stepsize-1 functions and we form $v = x^{\wedge(n;n)} + iy^{\wedge(n;n)}$, then $x = \mathrm{Re}(v^{\vee(1;n)})$ and $y = \mathrm{Im}(v^{\vee(1;n)})$.

If $x^\wedge$ is computed with the fast Fourier transform algorithm using floating-point arithmetic with b-bit precision, and $n = 2^k$, then the Euclidian norm of the error in the sequence $\langle x^\wedge(0), x^\wedge(1/n), \ldots, x^\wedge((n-1)/n)\rangle$ is bounded as follows: Let the resulting b-bit precision floating-point sequence be denoted by $x^{\wedge(p;n;b)}$. Then $\|x^{\wedge(p;n;b)} - x^{\wedge(p;n)}\| < 1.06 n^{1/2} \cdot 2^{3-b} k \cdot \|x^{\wedge(p;n)}\|$ [BP94].

4 The Fourier Integral Transform

Let $C^\infty_\downarrow(\mathcal{R})$ be the set of infinitely-differentiable rapidly-decreasing complex-valued functions on $\mathcal{R}$; $x \in C^\infty_\downarrow(\mathcal{R})$ means that the n-th derivative of x, $x^{(n)}$, exists for all $n \geq 0$, and that $t^h x^{(n)}(t) \to 0$ as $|t| \to \infty$ for every pair of non-negative integral values for h and n, *i.e.*, $x^{(n)}(t) = O(|t|^{-h})$ for every non-negative integral value n and every non-negative integral value h; this is what we mean by *rapidly-decreasing*. In other words, a rapidly-decreasing function $x(t)$ approaches 0 as $t \to \infty$ *and* as $t \to -\infty$ faster than any polynomial function increases, so that, for any polynomial function p and any rapidly-decreasing function x, $p(t)x(t) \to 0$ as $|t| \to \infty$ fast enough that $\int_{-\infty}^{\infty} p(t)x(t)dt$ exists. An example of a rapidly-decreasing function is $x(t) = e^{-t^2}$.

The set of functions $C^\infty_\downarrow(\mathcal{R})$ is called the *Schwartz* space; it is closed under differentiation, addition, multiplication, complex-conjugation, reversal, convolution, and Fourier transformation.

Introduce the inner product $(x, y)_{L^2(\mathcal{R})} = \int_{\mathcal{R}} xy^*$ and the norm $\|x\|_{L^2(\mathcal{R})} = (x, x)^{1/2}_{L^2(\mathcal{R})}$. (Recall that $\displaystyle\int_{\mathcal{R}} f = \int_{-\infty}^{\infty} f(t)dt := \lim_{\substack{a \to -\infty \\ b \to \infty}} \int_a^b f(t)dt$.)

Let the set of functions $L^2(\mathcal{R})$ be the completion of $C_\downarrow^\infty(\mathcal{R})$ with respect to the metric $\|x-y\|_{L^2(\mathcal{R})}$; $L^2(\mathcal{R})$ is obtained by taking all functions which are limits of Cauchy sequences of functions in $C_\downarrow^\infty(\mathcal{R})$ relative to the metric induced by the just-introduced $L^2(\mathcal{R})$-norm. The space $L^2(\mathcal{R})$ consists of the set of complex-valued measurable functions, x, defined on $[-\infty,\infty]$, such that $\int_{-\infty}^{\infty}|x(t)|^2\,dt<\infty$.

The Fourier integral transform of $x\in L^2(\mathcal{R})$ is defined by

$$x^\wedge(s)=\lim_{n\to\infty}\int_{-\infty}^{\infty}x_n(t)e^{-2\pi i s t}\,dt,$$

where $\langle x_0,x_1,\ldots\rangle$ is a Cauchy sequence of functions in $C_\downarrow^\infty(\mathcal{R})$ which converges to x in the sense that $\|x_n-x\|_{L^2(\mathcal{R})}\to 0$ as $n\to\infty$; the inverse Fourier integral transform is defined by:

$$x^{\wedge\vee}(t)=\lim_{n\to\infty}\int_{-\infty}^{\infty}x_n^\wedge(s)e^{2\pi i s t}\,ds \text{ where } x^\wedge(s)=\int_{-\infty}^{\infty}x_n(t)e^{-2\pi i s t}\,dt.$$

We may also write merely

$$x^\wedge(s)=\int_{-\infty}^{\infty}x(t)e^{-2\pi i s t}\,dt \quad\text{and}\quad x^{\wedge\vee}(t)=\int_{-\infty}^{\infty}x^\wedge(s)e^{2\pi i s t}\,ds,$$

where we interpret $\int_{-\infty}^{\infty}f(r,v)dr$ as the function $g(v)$ such that

$$\lim_{\substack{a\to-\infty\\ b\to\infty}}\|g(v)-\int_a^b f(r,v)dr\|_{L^2(\mathcal{R})}=0.$$

We may write $\wedge(\infty)$ and $\vee(\infty)$ to distinguish the Fourier integral transform and the Fourier integral inverse transform from the transforms $\wedge(p)$ and $\vee(p)$ defined on $L^2(Q)$.

Consider a function $x\in C_\downarrow^\infty(\mathcal{R})$, and consider the periodic extension $x_{[p]}$ where $x_{[p]}(t)=x(t)$ on $[-p/2,p/2)$. We can write the Fourier series for $x_{[p]}$ as

$$x_{[p]}(t)=\sum_h\left[\frac{1}{p}\int_{-p/2}^{p/2}x_{[p]}(r)e^{-2\pi i(h/p)r}\,dr\right]e^{2\pi i(h/p)t}.$$

Then

$$\begin{aligned}
x(t) &= \lim_{p\to\infty}x_{[p]}(t)\\
&= \lim_{p\to\infty}x_{[p]}^{\wedge(p)\vee(p)}(t)\\
&= \lim_{p\to\infty}\sum_h p\,(x_{[p]})^{\wedge(p)}(h/p)\,e^{2\pi i(h/p)t}\frac{1}{p}\\
&= \lim_{p\to\infty}\sum_h\left[\frac{p}{p}\int_{-p/2}^{p/2}x_{[p]}(r)e^{-2\pi i(h/p)r}\,dr\right]e^{2\pi i(h/p)t}\frac{1}{p}
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{p\to\infty} \sum_h \int_{-p/2}^{p/2} x_{[p]}(r) e^{-2\pi i (h/p)(r-t)}\, dr\, \frac{1}{p} \\
&= \lim_{p\to\infty} \int_{-\infty}^{\infty} \int_{-p/2}^{p/2} x_{[p]}(r) e^{-2\pi i s(r-t)}\, dr\, ds \\
&= \int_{-\infty}^{\infty} \left[\lim_{p\to\infty} \int_{-p/2}^{p/2} x_{[p]}(r) e^{-2\pi i s r}\, dr \right] e^{2\pi i s t}\, ds \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x(r) e^{-2\pi i s r}\, dr \right] e^{2\pi i s t}\, ds \\
&= \int_{-\infty}^{\infty} x^{\wedge(\infty)}(s) e^{2\pi i s t}\, ds \\
&= x^{\wedge(\infty)\vee(\infty)}(t), \text{ where we take } s = \frac{h}{p} \text{ as a function of } h, \text{ so } \frac{1}{p} \to ds \text{ as } p \to \infty.
\end{aligned}
$$

(Note how we compute the limit for $p \to \infty$ in two steps. First we note that $\lim_{p\to\infty} \sum_h \left[\int_{-p/2}^{p/2} x_{[p]}(r) e^{-2\pi i (h/p)(r-t)}\, dr\, \frac{1}{p} \right]$ is a Riemann sum over the mesh $\cdots, -\frac{2}{p}, -\frac{1}{p}, 0, \frac{1}{p}, \frac{2}{p}, \cdots$ with the stepsize $\frac{1}{p}$. We substitute $s = h/p$ and take the part-limit for $p \to \infty$ so that $\sum_h \to \int_{-\infty}^{\infty}$ together with $\frac{1}{p} \to ds$. Second, we take the remaining limit for $p \to \infty$ so that $\int_{-p/2}^{p/2} x_{[p]}(r) e^{-2\pi i s r}\, dr \to x^{\wedge(\infty)}(s)$.)

We have heuristically shown that the Fourier inversion theorem: $x(t) = x^{\wedge(\infty)\vee(\infty)}(s)$ a.e. holds for $x \in C_{\downarrow}^{\infty}(\mathcal{R})$. The same result holds for limits of sequences of functions in $C_{\downarrow}^{\infty}(\mathcal{R})$; this completion of $C_{\downarrow}^{\infty}(\mathcal{R})$ yields either the function space $L^2(\mathcal{R})$ or $L^1(\mathcal{R})$, depending on which norm is used in defining convergence, (with some exceptions in $L^1(\mathcal{R})$). (In these situations we have to introduce the limit of a convergent sequence of $C_{\downarrow}^{\infty}(\mathcal{R})$-functions, restricted to a finite support interval and periodically extended, in place of $x_{[p]}$.)

The gaps in this heuristic "proof" of the Fourier inversion theorem for the Fourier integral transform can be filled-in to establish a firm foundation for the Fourier integral transform and its inverse [DM72] [Rud98]. Basically, we need to justify exchanging limits and proper integration operations (*i.e.*, integration of functions of compact support,) which is generally done by appealing to a dominated convergence theorem.

When $x \in L^2(\mathcal{R})$ is real, $x^{\wedge R} = x^{\wedge *}$, and we have:

$$x^{\wedge\vee}(t) = \int_0^{\infty} M(s)\, cos(2\pi s t + \phi(s))\, ds,$$

with $x^{\wedge}(-s)e^{-2\pi i s t} + x^{\wedge}(s)e^{2\pi i s t} = M(s)\, cos(2\pi s t + \phi(s))$, where M and ϕ are functions of s defined as follows. Let $A(s) = x^{\wedge}(s) + x^{\wedge}(-s)$ and $B(s) = i(x^{\wedge}(s) - x^{\wedge}(-s))$. Define $M(s) = \left[A(s)^2 + B(s)^2\right]^{1/2} / (1 + \delta_{s,0})$ and $\phi(s) = \text{atan2}(-B(s), A(s))$. When x is real, M and ϕ are real and $M(s) \geq 0$.

Note that for $x \in L^2(\mathcal{R}), x^\vee(s) = [x(-t)]^\wedge$, *i.e.*, $x^\vee = x^{R\wedge}$. Thus, if x is even, $x^\wedge = x^\vee$.

We may also define the Fourier integral transform on the class $L^1(\mathcal{R})$, of complex-valued functions, x, such that $\int_{-\infty}^{\infty} |x(t)|\,dt < \infty$. For $x \in L^1(\mathcal{R})$, $x^\wedge(s) = \int_{-\infty}^{\infty} x(t)e^{-2\pi ist}\,dt$. However $x^\wedge$ may not, itself, be in $L^1(\mathcal{R})$, and the inverse Fourier transform integral, $\displaystyle\int_{-\infty}^{\infty} x^\wedge(s)e^{2\pi ist}ds$, does not exist for every such function, $x^\wedge$; however x can be recovered from $x^\wedge$ by a special smoothing device. Introduce the norm $\|x\|_{L^1(\mathcal{R})} = \int_{\mathcal{R}} |x|$. Then $x^{\wedge\vee}(t) = \lim_{a\downarrow 0}^{L^1(\mathcal{R})} \int_{-\infty}^{\infty} e^{-2\pi^2 s^2 a} x^\wedge(s) e^{2\pi ist}\,dt$, where $\lim_{a\downarrow 0}^{L^1(\mathcal{R})} f(a)$ denotes the function, g, such that $\lim_{a\downarrow 0} \|f(a) - g\|_{L^1(\mathcal{R})} = 0$.

The relationships between the Fourier transform on the spaces $L^2(Q)$ and $L^2(\mathcal{Z}/p)$ are: (periodic function)$^\wedge$ = discrete function and (discrete function)$^\vee$ = periodic function. As the period p tends to ∞ so that compliant periodic functions are extended to converge to rapidly-decreasing functions, we have periodic function$\rightarrow C_\downarrow^\infty(\mathcal{R})$-function, and discrete function $\rightarrow C_\downarrow^\infty(\mathcal{R})$-function. The space $L^2(\mathcal{R})$ is the completion of this common limit space obtained as $p \rightarrow \infty$, and the Fourier transform operators $\wedge(p)$ and $\vee(p)$ become the Fourier integral transform operators $\wedge(\infty)$ and $\vee(\infty)$ on $L^2(\mathcal{R})$.

Note the Fourier integral transform and inverse transform can be "re-parmeterized" in many forms. The general relations that satisfy $x^{\wedge\vee} = x$ are:

$$x^\wedge(s) = \left[\frac{|b|}{(2\pi)^{1-a}}\right]^{1/2} \int_{-\infty}^{\infty} x(t)e^{ibst}\,dt \quad \text{and} \quad x^{\wedge\vee}(t) = \left[\frac{|b|}{(2\pi)^{1+a}}\right]^{1/2} \int_{-\infty}^{\infty} x^\wedge(s)e^{-ibts}\,ds,$$

where $b \neq 0$. We use $a = 0$ and $b = -2\pi$ here; this is the usual choice for signal processing applications. Other common choices are: $a = 1$ and $b = -1$ (mathematics), $a = 1$ and $b = 1$ (probability), $a = 0$ and $b = 1$ (modern physics), and $a = -1$ and $b = 1$ (classical physics),

The common choice $a = 1$ and $b = -1$ results in the argument, w, of the Fourier transform function, $x^\wedge(w)$, having the natural unit of radians per t-unit (angular frequency) rather than cycles per t-unit (linear frequency). Note if s is a value representing a number of cycles per t-unit, then $w = 2\pi s$ is the correspondimg number of radians per t-unit.

Define $\wedge[a, b]$ and $\vee[a, b]$ as the operators such that $x^{\wedge[a,b]}(s) = \left[\frac{|b|}{(2\pi)^{1-a}}\right]^{1/2} \displaystyle\int_{-\infty}^{\infty} x(t)e^{ibst}\,dt$ and $y^{\vee[a,b]}(t) = \left[\frac{|b|}{(2\pi)^{1+a}}\right]^{1/2} \displaystyle\int_{-\infty}^{\infty} x^\wedge(s)e^{-ibts}\,ds$ where $b \neq 0$. Then we have $x^{\wedge[0,-2\pi]}(s) = x^{\wedge[1,-1]}(2\pi s)$ and $y^{\vee[0,-2\pi]}(t) = 2\pi y^{\vee[1,-1]}(2\pi t)$.

Thus, if $x^\wedge(w) = f(w)$ with $a = 1$ and $b = -1$, then $x^\wedge(s) = f(2\pi s)$ with $a = 0$ and $b = -2\pi$, and if $y^\vee(r) = g(r)$ with $a = 1$ and $b = -1$, then $y^\vee(t) = 2\pi g(2\pi t)$ with $a = 0$ and $b = -2\pi$ – with the proviso that if a Fourier transform, $v^\wedge(w)$, occurs in the expression defining $f(w)$ or $g(r)$, then it is to be replaced by $v^\wedge(s)$, and if an inverse Fourier transform, $u^\vee(r)$, occurs in the expression defining $f(w)$ or $g(r)$, then it is to be replaced by $u^\vee(t)$ (these rules follow from the preceding paragraph.) We need these formulas to convert among the most common differing choices found in different books.

In the case $b = -1$ with a other than 1, the conversion from $x^{\wedge[a,-1]}$ to $x^{\wedge[0,-2\pi]}$ is complicated by the need to multiply $x^{\wedge[a,-1]}$ by $\left[(2\pi)^{1-a}\right]^{1/2}$ and multiply $y^{\vee[a,-1]}$ by $2\pi\left[(2\pi)^{1+a}\right]^{1/2}$.

4.1 Geometrical Interpretation

With the inner product $(x,y)_{L^2(\mathcal{R})} := \int xy^*$, and the norm $\|x\|_{L^2(\mathcal{R})} := (x,x)^{1/2}_{L^2(\mathcal{R})}$, $L^2(\mathcal{R})$ is a Hilbert space, but unlike the situation for periodic functions, the functions $e^{2\pi ist}$ do not belong to $L^2(\mathcal{R})$. Nevertheless, the Fourier integral transform, $\wedge$, is a one-to-one distance-preserving map of $L^2(\mathcal{R})$ onto $L^2(\mathcal{R})$, and we have Parseval's identity: $(x,y)_{L^2(\mathcal{R})} = (x^\wedge, y^\wedge)_{L^2(\mathcal{R})}$ and Plancherel's identity: $\|x\|_{L^2(\mathcal{R})} = \|x^\wedge\|_{L^2(\mathcal{R})}$. (The reason $x^\wedge$ exists for $x \in L^2(\mathcal{R})$, even though $e^{2\pi ist} \notin L^2(\mathcal{R})$, is because $|x|$ decreases rapidly enough to ensure that $\left|\int_{-\infty}^{\infty} x(t)e^{2\pi ist}dt\right| < \infty$.)

The Fourier integral transform is a unitary linear transformation on $L^2(\mathcal{R})$, and since $x^{\wedge\wedge\wedge\wedge} = x$, $\wedge$ is an infinite-dimensional analog of a "90-degree" rotation mixture. The inverse Fourier integral transform, $\vee$, is also a one-to-one distance-preserving map of $L^2(\mathcal{R})$ onto $L^2(\mathcal{R})$.

The Fourier integral transform is defined on $L^1(\mathcal{R})$ and for $x \in L^1(\mathcal{R})$, $x^\wedge \in C(\mathcal{R})$ (The set $C(\mathcal{R})$ is the set of continuous complex functions defined on $\mathcal{R}$,) however, similar to the situation in $L^1(Q)$, $x^\wedge$ may be unbounded, so that $x^{\wedge\vee}$ does not exist.

$C^\infty_\downarrow(\mathcal{R})$ is dense in $L^2(\mathcal{R})$ (and in $L^1(\mathcal{R})$), and $\wedge$ maps $C^\infty_\downarrow(\mathcal{R})$ onto itself. Unlike the class of square-integrable finite-period periodic functions, $L^2(\mathcal{R}) \not\subseteq L^1(\mathcal{R})$ (and $L^1(\mathcal{R}) \not\subseteq L^2(\mathcal{R})$.)

Exercise 4.1: If $C^\infty_\downarrow(\mathcal{R})$ is dense in $L^2(\mathcal{R})$ and also in $L^1(\mathcal{R})$, why doesn't $L^2(\mathcal{R}) = L^1(\mathcal{R})$?

Solution 4.1: The completion process that forms $L^2(\mathcal{R})$ and similarly $L^1(\mathcal{R})$ from $C^\infty_\downarrow(\mathcal{R})$ is done with respect to distinct norms.

4.2 Band-Limited Functions

A band-limited function, $x \in L^2(\mathcal{R})$, which has no spectral component whose frequency lies outside the band $[-b,b]$, is determined by an infinite sequence of discrete samples $\ldots$, $x(-2/(2b))$, $x(-1/(2b))$, $x(0)$, $x(1/(2b))$, $x(2/(2b))$, $\ldots$, as:

$$x(t) = \frac{1}{2b}\sum_{-\infty<h<\infty} x(h/(2b))\frac{sin\,(2\pi b(t-h/(2b)))}{\pi(t-h/(2b))}.$$

This series is known as the cardinal series expansion of x.

Let $x \in L^2(\mathcal{R})$ be band-limited with $x^\wedge(s) = 0$ for $|s| > b$.

Then we can write $x(t) = \int_{-b}^{b} x^\wedge(s)e^{2\pi ist}\,ds$. (Here $\wedge$ denotes the Fourier integral transform $\wedge(\infty)$.)

And since $x^\wedge(s) = 0$ outside $[-b, b]$, we can write $x^\wedge(s)$ as a Fourier series valid for $s \in (-b, b)$:

$$x^\wedge(s) = \sum_{-\infty<h<\infty} x^{\wedge\wedge(p)}(h/(2b))e^{2\pi ihs/(2b)}.$$

And we have $x^{\wedge\wedge(p)}(v) = \frac{1}{2b}\int_{-b}^{b} x^\wedge(r)e^{-2\pi irv}\,dr$ for $v = \ldots, -1/(2b), 0, 1/(2b), \ldots$, so,

$$x^\wedge(s) = \sum_{-\infty<h<\infty} \left[\frac{1}{2b}\int_{-b}^{b} x^\wedge(r)e^{-2\pi irh/(2b)}\,dr\right] e^{2\pi ish/(2b)},$$

and, inverting the order of summation by writing $-h$ for h, we have:

$$\begin{aligned} x^\wedge(s) &= \sum_{\infty>h>-\infty} \left[\frac{1}{2b}\int_{-b}^{b} x^\wedge(r)e^{2\pi irh/(2b)}\,dr\right] e^{-2\pi ish/(2b)} \\ &= \sum_{\infty>h>-\infty} \left[\frac{1}{2b}x(h/(2b))\right] e^{-2\pi ihs/(2b)}. \end{aligned}$$

But now,

$$\begin{aligned} x(t) &= \int_{-b}^{b} x^\wedge(s)e^{2\pi ist}\,ds \\ &= \int_{-b}^{b} \left[\sum_h \frac{1}{2b}x(h/(2b))e^{-2\pi ish/(2b)}\right] e^{2\pi ist}\,ds \\ &= \sum_h \frac{1}{2b}x(h/(2b))\int_{-b}^{b} e^{2\pi i(t-h/(2b))s}\,ds \\ &= \sum_h \frac{1}{2b}x(h/(2b))\left[\frac{e^{2\pi i(t-h/(2b))s}}{2\pi i(t-h/(2b))}\right]_{s=-b}^{s=b} \\ &= \sum_h \frac{1}{2b}x(h/(2b))\left[\frac{e^{2\pi i(t-h/(2b))b} - e^{-2\pi i(t-h/(2b))b}}{2\pi i(t-h/(2b))}\right] \\ &= \frac{1}{2b}\sum_h x(h/(2b))\frac{sin\,(2\pi b(t-h/(2b)))}{\pi(t-h/(2b))}, \end{aligned}$$

where we compute 0/0 as 1.

The identity $x(t) = \frac{1}{2b}\sum_{h=-\infty}^{\infty} x(h/(2b))\frac{sin\,(2\pi b(t-h/(2b)))}{\pi(t-h/(2b))}$ is called the Shannon sampling theorem; it is a generalization of Whittaker's interpolation formula to band-limited functions in $L^2(\mathcal{R})$.

In general, for $x \in L^2(\mathcal{R})$, not necessarily band-limited, we have both $x(t) \to 0$ as $|t| \to \infty$ and $x^\wedge(s) \to 0$ as $|s| \to \infty$. Therefore the truncated sum $\frac{1}{2b} \sum_{-m \le h \le m} x(h/(2b)) \frac{sin\,(2\pi b(t - h/(2b)))}{\pi(t - h/(2b))}$ will be a good approximation to $x(t)$ when m is large enough.

Exercise 4.2: Explain what the Shannon sampling theorem says about the value of $x(t)$ for $t \in \mathcal{Z}/(2b)$.

Exercise 4.3: Let b be a positive real value and let $x(t) \in L^2(\mathcal{R})$ be an even function with $x(t) = 0$ for $|t| \ge b$ such that $x(t) + x(b-t)$ is constant for $0 \le t \le b$. We shall call such a function a *bi-constant* function.

An example of an even function, x, that satisfies $x(t) = 0$ for $|t| \ge b$ and $x(t) + x(b-t) = c$ for $0 \le t \le b$ is $x(t) = \begin{cases} c(1 - t/b), & \text{if } |t| < b \\ 0, & \text{otherwise.} \end{cases}$

In general, an even function, x, that satisfies $x(t) = 0$ for $|t| \ge b$ and $x(t) + x(b-t) = c$ for $0 \le t \le b$ also satisfies $x'(t) - x'(b-t) = 0$ for $0 \le t \le b$.

Show that, for $b = 1$ and $0 < t < 1$, $x'(t) = e^{-1/(t-t^2)}$ satisfies $x'(t) - x'(b-t) = 0$ for $0 \le t \le 1$. We can thus take $x(t) = \int_t^1 e^{-1/(r-r^2)} dr$ for $0 \le t \le 1$, and extend $x(t)$ to be a continuous even bi-constant function on $\mathcal{R}$ with support in $(-1, 1)$. What is this function?

Also show that any bi-constant function x with the support set $(-b, b)$ satisfies $x^\wedge(t) = x^\vee(s) = 0$ for $s = k/b$ with $k \in \mathcal{Z} - \{0\}$

Solution 4.3: Let x be a bi-constant function. Recall that x is an even function and let $c = x(0) = x(t) + x(b-t)$ for $0 \le t \le b$. Note $x(t) + x(b-t) = c$ for $0 \le t \le b$ is equivalent to $x(\alpha b) + x((1-\alpha)(-b))$ for $\alpha \in [0, 1]$.

Now, $x^{R\wedge} = x^\vee$ and since x is even, $x^\wedge = x^\vee$. Also, for $k \in \mathcal{Z}$, we have

$$\begin{aligned} x^\wedge(k/b) &= \int_{-\infty}^{\infty} x(t) e^{-2\pi i (k/b) t} dt \\ &= \int_{-b}^{b} x(t) e^{-2\pi i (k/b) t} dt \\ &= \sum_{j=-1}^{0} \int_{jb}^{(j+1)b} x(t) e^{-2\pi i (k/b) t} dt \\ &= \sum_{j=-1}^{0} \int_{0}^{b} x(t + jb) e^{-2\pi i (k/b) t} dt, \end{aligned}$$

since $e^{-2\pi i (k/b) t}$ is periodic with period b, so that $e^{-2\pi i (k/b)(t+jb)} = e^{-2\pi i (k/b) t} \cdot e^{-2\pi i k j}$ and $e^{-2\pi i k j} = 1$.

But then,

$$x^\wedge(k/b) = \int_0^b [x(t-b) + x(t)] e^{-2\pi i (k/b) t} dt$$

$$\begin{aligned} &= \int_0^b [x(t) + x(b-t)] e^{-2\pi i (k/b) t} dt \\ &= \int_0^b c e^{-2\pi i (k/b) t} dt. \end{aligned}$$

Thus, for $k \neq 0$,

$$\begin{aligned} x^\wedge(k/b) &= \int_0^b c e^{-2\pi i (k/b) t} dt \\ &= c e^{-2\pi i (k/b) t} / (-2\pi i (k/b)) \Big|_{t=0}^{t=b} \\ &= -\frac{cb}{2\pi i k} \left[e^{-2\pi i k} - 1 \right] \\ &= -\frac{cb}{2\pi i k} [1 - 1] \\ &= 0, \end{aligned}$$

and, for $k = 0$,

$$\begin{aligned} x^\wedge(k/b) = x^\wedge(0) &= \int_0^b c e^{-2\pi i (k/b) t} dt \\ &= \int_0^b c dt \\ &= bc. \end{aligned}$$

Beware. We have shown that the integral Fourier transform of a bi-constant $L^2(\mathcal{R})$-function x is 0 at $\pm 1/b, \pm 2/b, \ldots$, but this does not imply that $x^\wedge(s)$ is 0 everywhere away from $s = 0$(!) What can you say about the Fourier series of the period $2b$-function $x_{[2b]}(t)$ constructed from the bi-constant function x given on $[-b, b]$?

Paley and Wiener have shown that any band-limited function, $x \in L^2(\mathcal{R})$, with $x^\wedge(s) = 0$ for $|s| > b$, can be extended to a unique function, y, of a complex variable, z, such that $y(z)$ is an entire function, $y(z) = x(z)$ for $z \in \mathcal{R}$ and $|y(z)| \leq c e^{a\pi|z|}$ for some constants c and a. In fact, this function, $y(z)$, is just the inverse Fourier transform of the function $x^\wedge$, so

$$y(z) = \int_{-b}^{b} x^\wedge(s) e^{2\pi i s z} \, ds.$$

Thus, a band-limited function x extends to a particular function y of a complex variable, where the Fourier transform of x determines y on the entire complex plane, by means of the Fourier inversion formula. The converse is also true: if $y(z)$ is an entire function with $|y(z)| \leq c e^{a\pi|z|}$ for some constants c and a, then y is band-limited.

The cardinal series expansion of x also determines y as

$$y(z) = \frac{1}{2b} \sum_{h=-\infty}^{\infty} x(h/(2b)) \frac{sin\,(2\pi b(z - h/(2b)))}{\pi(z - h/(2b))}.$$

Note that a non-zero band-limited function $x \in L^2(\mathcal{R})$ is never a domain-limited function; *i.e.*, there is no value $a > 0$ such that $x(t) = 0$ for $|t| > a$. The converse is also true; a non-zero domain-limited function is not a band-limited function.

In general, the more "spread-out" the function x is, the more "concentrated" $x^\wedge$ is, and vice-versa; Dym and McKean [DM72] report a proof of a descriptive result relating "spread" and "concentration". Consider $x \in L^2(\mathcal{R})$ normalized to have total "power" $\|x\|^2_{L^2(\mathcal{R})} = 1 = \|x^\wedge\|^2_{L^2(\mathcal{R})}$.

Let $\alpha_x^2 \leq 1$ be the "power" of x in the interval $[-a, a]$, so that $\alpha_x^2 = \int_{-a}^{a} |x(t)|^2 dt$. Let $\beta_x^2 \leq 1$ be the "power" of $x^\wedge$ in the interval $[-b, b]$, so that $\beta_x^2 = \int_{-a}^{a} |x^\wedge(s)|^2 ds$. Note α_x is a function of a and β_x is a function of b.

Fix the values a and b, with $a > 0$ and $b > 0$. No matter what values of a and b we choose, it is not possible to choose a function x so that $\alpha_x^2 = 1$ and $\beta_x^2 = 1$ simultaneously. However, for any pair $(\bar{\alpha}, \bar{\beta}) \in \{(\alpha, \beta) \mid \beta \leq \alpha\gamma_{ab}^{1/2} + (1-\alpha^2)^{1/2}(1-\gamma_{ab})^{1/2}\} - \{(0,1),(1,0)\}$, where γ_{ab} is a certain non-negative monotonically-increasing function with γ_{ab} rapidly approaching $.916\ldots$ as ab approaches ∞, there is a function y that achieves $\bar{\alpha}^2 = \alpha_y^2 = \int_{-a}^{a} |y(t)|^2 dt$ and $\bar{\beta}^2 = \beta_y^2 = \int_{-a}^{a} |y^\wedge(s)|^2 ds$ for the prior-chosen fixed positive values of a and b. ($\gamma_{ab} \approx .916\ldots \cdot (1 - e^{-4\pi ab})$.)

Another famous example of the "spread" vs. "concentration" relation between a function x and its Fourier transform $x^\wedge$ is the Heisenberg inequality. The classical form of the Heisenberg inequality for $x \in L^2(\mathcal{R})$ with $\|x\|_{L^2(\mathcal{R})} = 1$ is:

$$\left[\int_{-\infty}^{\infty} (t - E(A))^2 |x(t)|^2 dt\right] \cdot \left[\int_{-\infty}^{\infty} (s - E(B))^2 |x^\wedge(s)|^2 ds\right] \geq 1/(16\pi^2),$$

where A is a real random variable with the probability density function $[P(A \leq t)]' = |x(t)|^2$ and B is a real random variable with the probability density function $[P(B \leq s)]' = |x^\wedge(s)|^2$. The Heisenberg inequality asserts that $Var(A) \cdot Var(B) \geq 1/(16\pi^2)$.

The unnormalized form of the Heisenberg inequality for $x \in L^2(\mathcal{R})$ is:

$$\left[\int_{-\infty}^{\infty} t^2 |x(t)|^2 dt\right] \cdot \left[\int_{-\infty}^{\infty} s^2 |x^\wedge(s)|^2 ds\right] \geq \|x\|^4_{L^2(\mathcal{R})}/(16\pi^2);$$

this states that the "spread-weighted" power of x and the "spread-weighted" power of $x^\wedge$ are (roughly) inversely related. (Recall that Parseval's identity says that the total unweighted powers of x and $x^\wedge$ are equal.)

4.3 The Poisson Summation Formula

Take $p > 0$. Given the rapidly-decreasing function $f(t)$ defined on $\mathcal{R}$, define

$$g_p(t) := \sum_{-\infty < k < \infty} f(t + kp),$$

where the summation index $k \in \mathcal{Z}$. The function g_p is periodic of period p(!) Moreover because $f(r) \to 0$ rapidly as $|r| \to \infty$, we have $\lim_{p \to \infty} g_p(t) = f(t)$.

Exercise 4.4: Explain graphically what the function $g_p(t)$ is as the 'sum' of the graphs of the functions $f(t+kp)$ for $k = \ldots, -1, 0, 1, \ldots$.

We have the Fourier series:

$$g_p(s) = \sum_{-\infty<n<\infty} \left[\frac{1}{p}\int_{-p/2}^{p/2} g_p(t)e^{-2\pi it(n/p)}\,dt\right] e^{2\pi i(n/p)s},$$

and,

$$\begin{aligned}
\frac{1}{p}\int_{-p/2}^{p/2} g_p(t)e^{-2\pi i(n/p)t}\,dt &= \frac{1}{p}\int_{-p/2}^{p/2}\left[\sum_{-\infty<k<\infty} f(t+kp)\right]e^{-2\pi i(n/p)t}\,dt \\
&= \frac{1}{p}\sum_{-\infty<k<\infty}\int_{-p/2}^{p/2} f(t+kp)e^{-2\pi i(n/p)t}\,dt \\
&= \frac{1}{p}\sum_{-\infty<k<\infty}\int_{-p/2}^{p/2} f(t+kp)e^{-2\pi i(n/p)(t+kp)}\,dt \\
&= \frac{1}{p}\sum_{-\infty<k<\infty}\int_{(k-\frac{1}{2})p}^{(k+\frac{1}{2})p} f(r)e^{-2\pi i(n/p)r}\,dr \\
&= \frac{1}{p}\int_{-\infty}^{\infty} f(r)e^{-2\pi i(n/p)r}\,dr \\
&= \frac{1}{p}f^{\wedge(\infty)}(n/p).
\end{aligned}$$

So,

$$g_p(s) = \sum_{-\infty<n<\infty}\left[\frac{1}{p}f^{\wedge(\infty)}(n/p)\right]e^{-2\pi i(n/p)s}.$$

Note for $s = 0$, we have $\sum_{-\infty<k<\infty} f(kp) = \sum_{-\infty<n<\infty}\frac{1}{p}f^{\wedge(\infty)}(n/p)$, and in addition, $\sum_k f(k) = \sum_n f^{\wedge(\infty)}(n)$ when we take $p = 1$. This identity is the Poisson summation formula; like the Plancherel identity, it shows the "equivalence" of the "sizes" of the rapidly-decreasing function f and its Fourier integral transform.

Also, we obtain the Fourier inversion theorem:

$$\lim_{p\to\infty} g_p(s) = f(s) = \lim_{p\to\infty}\sum_n f^{\wedge(\infty)}(n/p)e^{-2\pi isn/p}\frac{1}{p} = \int_{-\infty}^{\infty} f^{\wedge(\infty)}(t)e^{-2\pi ist}\,dt = f^{\wedge(\infty)\vee(\infty)}(s).$$

4.4 Structural Relations

- Recall that any function, x, can be written $x = \mathrm{even}(x) + \mathrm{odd}(x)$, where $\mathrm{even}(x)(t) = (x(t) + x(-t))/2$ and $\mathrm{odd}(x) = (x(t) - x(-t))/2$. If $x(t) = 0$ for $t < 0$, then $\mathrm{even}(x)(t) = x(|t|)/2$ and $\mathrm{odd}(x)(t) = \mathit{sign}(t) \cdot x(|t|)/2$. We have $x^\wedge = \mathrm{even}(x)^\wedge + \mathrm{odd}(x)^\wedge$, and $\mathrm{even}(x)^\wedge = \mathrm{Re}(x^\wedge)$, $\mathrm{odd}(x)^\wedge = i\,\mathrm{Im}(x^\wedge)$.

- $x^{\wedge *} = x^{*\vee}$ and if x is real, $x^\wedge$ is Hermitian, *i.e.*, $x^{\wedge R} = x^{\wedge *}$, also if x is imaginary then $x^{\wedge *} = -x^{\wedge R}$. Also $x^{R\wedge} = x^{*\wedge *} = x^{**\vee} = x^\vee$, and $x^{*\wedge} = x^{R\wedge *}$.

  Exercise 4.5: Show that $* \vee * = \wedge$.

- $x^{\wedge\wedge} = x^R$, and so $x^{\wedge\wedge\wedge\wedge} = x$ and $x^{\wedge R \wedge} = x$, *i.e.*, $R = \wedge\wedge$, $R\wedge = \wedge R$, $\vee = \wedge R = R\wedge$, $\wedge = R\vee$, and $\wedge \wedge \wedge\wedge = I$.

  Exercise 4.6: Show that $\wedge\wedge = R$.

  Solution 4.6: $x^{\wedge\wedge}(s) = \int_{-\infty}^{\infty} x^\wedge(t) e^{-2\pi i s t}\, dt$, so $x^{\wedge\wedge}(-s) = \int_{-\infty}^{\infty} x^\wedge(t) e^{2\pi i s t}\, dt = x^{\wedge\vee}(s) = x(s)$. Thus, $x^{\wedge\wedge}(s) = x(-s)$, so $\wedge\wedge = R$.

  Exercise 4.7: Show that $R\vee = \vee R$.

  Exercise 4.8: Show that $\wedge * \wedge = R * R = *$.

  Exercise 4.9: Use Parseval's identity: $(x, y)_{L^2(\mathcal{R})} = (x^\wedge, y^\wedge)_{L^2(\mathcal{R})}$ and the operator identities above to show that $(x, y^\wedge)_{L^2(\mathcal{R})} = (x^\wedge, y^R)_{L^2(\mathcal{R})}$ and $(x^\wedge, y^{\wedge *})_{L^2(\mathcal{R})} = (x^R, y^*)_{L^2(\mathcal{R})}$ or equivalently, $(x^\wedge, y)_{L^2(\mathcal{R})} = (x, y^{R\wedge})_{L^2(\mathcal{R})}$.

  Exercise 4.10: Show that $(x, y^{\wedge *})_{L^2(\mathcal{R})} = (x^\wedge, y^*)_{L^2(\mathcal{R})}$. Hint: remember $\wedge\wedge = R$.

  Solution 4.10: Parseval's identity implies that $(x^\wedge, y^*)_{L^2(\mathcal{R})} = (x^{\wedge\wedge}, y^{*\wedge})_{L^2(\mathcal{R})}$ and $(x^{\wedge\wedge}, y^{*\wedge})_{L^2(\mathcal{R})} = (x^R, y^{R\wedge *})_{L^2(\mathcal{R})} = (x^R, y^{\wedge R *})_{L^2(\mathcal{R})}$. But

$$\begin{aligned} (x^R, y^{\wedge R *})_{L^2(\mathcal{R})} &= \int_{-\infty}^{\infty} x(-r) y^\wedge(-r)^{**} dr \\ &= \int_{-\infty}^{\infty} x(-r) y^\wedge(-r) dr \\ &= \int_{\infty}^{-\infty} x(r) y^\wedge(r)(-1) dr \\ &= \int_{-\infty}^{\infty} x(r) y^\wedge(r) dr \\ &= (x, y^{\wedge *})_{L^2(\mathcal{R})}. \end{aligned}$$

  Exercise 4.11: Show that $(x, y^{\vee *})_{L^2(\mathcal{R})} = (x^\vee, y^*)_{L^2(\mathcal{R})}$.

- $(x')^\wedge(s) = 2\pi i s \cdot x^\wedge(s)$.

  Exercise 4.12: Prove that $(x')^\wedge(s) = 2\pi i s \cdot x^\wedge(s)$ for $x \in L^2(\mathcal{R})$. Hint: use integration by parts, plus the fact that $x(t) \to 0$ as $t \to \infty$ or $t \to -\infty$.

- Let $v(t)$ be a polynomial function. In general, using operator notation, $\left[v(D_t^1)(x(t))\right]^\wedge(s) = v(2\pi i s) x^\wedge(s)$, where D_t^1 denotes the prefix differentiation operator with respect to t.

- $(x^\wedge)'(s) = (-i2\pi t \cdot x(t))^\wedge(s)$.
- In general, using operator notation, $[(2\pi is)^p D_s^q](x^\wedge(s)) = [[D_t^p(-2\pi it)^q](x(t))]^\wedge(s)$ for non-negative integers p and q. (The differentiation operator D_s^q denotes q-fold differentiation with respect to s.)
- The operators $[D_r^1 - (2\pi r)^2]$ and $\wedge$ commute, where "r" denotes the argument of the function to which the operators are applied.
- $(e^{i2\pi vt}x(t))^\wedge(s) = x^\wedge(s-v)$.
- $(x(at+b))^\wedge(s) = (1/|a|)e^{2\pi is(b/a)}x(t)^\wedge(s/a)$. Taking $b=0$ and $a=-1$ shows that $R\wedge = \wedge R$.
- $x(t)\cdot cos(wt)$ is the carrier oscillation $cos(wt)$ of frequency $w/(2\pi)$ amplitude-modulated by the signal $x(t)$: we have $(x(t)\cdot cos(wt))^\wedge(s) = (1/2)x^\wedge(s - w/(2\pi)) + (1/2)x^\wedge(s + w/(2\pi))$; these are the two "side-band" terms of x.
- if $x(t) = e^{-t^2/(2\sigma^2)}$, then $x^\wedge(s) = (2\pi)^{1/2}\sigma e^{-2\pi^2\sigma^2 s^2}$. In particular, $(e^{-\pi t^2})^\wedge(s) = e^{-\pi s^2}$.
- The eigenvalues of $\wedge$ are $-i$, i, -1, and 1, and the eigenfunctions are the Hermite functions: $h_n(r) := (-1)^n e^{\pi r^2}(e^{-2\pi r^2})^{(n)}/n!$ for $n \geq 0$, where the superscript (n) denotes the n-th derivative. Specifically we have $h_n^\wedge = (-i)^n \cdot h_n$ [DM72].
- $x^\wedge$ is continuous for $x \in L^1(\mathcal{R})$.

4.5 Convolution

For $x, y \in L^2(\mathcal{R})$, we define the convolution

$$(x \circledast y)(t) = \int_{-\infty}^{\infty} x(s)y(t-s)\,ds.$$

Note if x and y are both 0 outside the interval $[-p/2, p/2]$ then $x \circledast y$ is 0 outside the interval $[-p, p]$.

The following identities hold.

- $(x \circledast y)^\wedge = x^\wedge y^\wedge$.
- $x \circledast y = y \circledast x$, $x \circledast (y \circledast z) = (x \circledast y) \circledast z$, and $x \circledast (y+z) = x \circledast y + x \circledast z$.
- $(x \circledast y)' = x' \circledast y = x \circledast y'$.
- $\int_{-\infty}^{\infty}(x \circledast y)(t)\,dt = (\int_{-\infty}^{\infty} x(t)\,dt)\cdot(\int_{-\infty}^{\infty} y(t)\,dt)$.
- $(xy)^\wedge = x^\wedge \circledast y^\wedge$.

- If $x \in L^2(\mathcal{R})$ is band-limited with $x^\wedge(s) = 0$ for $|s| > b$, then

$$x = (x^\wedge)^\vee = (x^\wedge B_{2b})^\vee = x^{\wedge\vee} \circledast B_{2b}^\vee = x \circledast sinc_{2b},$$

where $sinc_p(t) := sin(\pi pt)/(\pi pt)$, and $B_p(t) := \begin{cases} 1, & \text{if } |t| < p/2, \\ 1/2, & \text{if } |t| = p/2, \\ 0, & \text{otherwise.} \end{cases}$

Also, we have the cross-correlation kernel function

$$(x \otimes y)(r) \quad = \quad \int_{-\infty}^{\infty} x(t)y(t+r)\,dt.$$

Thus $x \otimes y = x^R \circledast y$, and when x is real, $(x \otimes x)^\wedge = x^\wedge x^{\wedge *} = |x^\wedge|^2$.

Let $\langle y(t)\rangle_x = (\int_{-\infty}^{\infty} y(t)x(t)\,dt)/(\int_{-\infty}^{\infty} x(t)\,dt)$. Note if x is a probability density function, and if t is a random variable such that $y(t)$ has the probability density function x, then $\langle y(t)\rangle_x = E(y(t))$.

Thus,

$$\begin{aligned} \langle t\rangle_{x\circledast y} &= \langle t\rangle_x + \langle t\rangle_y, \quad \text{and} \\ \langle t^2\rangle_{x\circledast y} &= \langle t^2\rangle_x + 2\langle t\rangle_x\langle t\rangle_y + \langle t^2\rangle_y. \end{aligned}$$

4.6 Eigenvalues and Eigenfunctions

Dym and McKean [DM72] present Norbert Weiner's derivation of the eigenvalues and eigenfunctions of the Fourier integral linear operator $\wedge$.

If $f^\wedge = \lambda f$, *i.e*, if f is an eigenfunction and λ is a corresponding eigenvalue of the linear operator $\wedge$, then $f^{\wedge\wedge} = \lambda f^\wedge$, $f^{\wedge\wedge\wedge} = \lambda f^{\wedge\wedge}$, and $f^{\wedge\wedge\wedge\wedge} = \lambda f^{\wedge\wedge\wedge}$. But then, $f^{\wedge\wedge\wedge\wedge} = \lambda^4 f$, and, since $f = f^{\wedge\wedge\wedge\wedge}$, we have $f = \lambda^4 f$, *i.e.*, $f(t) = \lambda^4 f(t)$, and fixing t to be any value a such that $f(a) \neq 0$ and dividing by $f(a)$, we have $1 = \lambda^4$. Thus the eigenvalues of $\wedge$ are the fourth roots of unity: $1, i, -1, -i$.

Now, if we recall that the Fourier integral transform of $x(t) = e^{-\pi t^2}$ is $x^\wedge(s) = e^{-\pi s^2}$, then we might search for eigenfunctions of $\wedge$ related to Gaussians. And since $(f'(t))^\wedge(s) = 2\pi is \cdot f^\wedge(s)$, we might also look at functions of the form $u(t)e^{-\pi t^2}$ where $u(t)$ is a polynomial. It turns-out that the eigenfunctions of $\wedge$ are the *Hermite* functions:

$$h_n(t) = \frac{(-1)^n}{n!} e^{\pi t^2} D_t^n \left[e^{-2\pi t^2}\right] \text{ for } n \in \{0, 1, \ldots\},$$

(Here D_t^n denotes the operation of n-fold differentiation with respect to t.)

The eigenvalue associated with the eigenfunction h_n is $(-i)^n$, *i.e.*, $h_n^\wedge = (-i)^n h_n$.

Exercise 4.13: Show that $h_n(t) = e^{-\pi t^2} u_n(t)$ where $u_n(t)$ is a polynomial of exact degree n with real coefficients. Hint: the polynomials $H_n(t) := (-1)^n e^{-t^2} D_t^n \left[e^{-t^2} \right]$ for $n \in \{0, 1, \ldots\}$ (called *Hermite polynomials*) satisfy the recursion relation $H_{n+1}(t) = 2tH_n(t) - 2nH_{n-1}(t)$ with $H_0(t) = 1$, $H_1(t) = 2t$, $H_2(t) = 4t^2 - 2$, etc. Consider $n!\, u_n(t/\sqrt{2\pi}) = H_n(t)$.

Let $e_n = h_n / \|h_n\|_{L^2(\mathcal{R})}$. Then, just as happens with unitary linear transformations on $\mathcal{C}^n$, it turns-out that $\langle e_0, e_1, \ldots \rangle$ is an orthonormal approximating basis for $L^2(\mathcal{R})$. Therefore every function $x \in L^2(\mathcal{R})$ has a Hermite function expansion: $x = \sum_{0 \le n \le \infty} (x, e_n) e_n$ based on the normalized eigen-functions of $\wedge$ (!) (By $x = \sum_{0 \le n \le \infty} (x, e_n) e_n$, we mean $\lim_{k \to \infty} \|x - \sum_{0 \le n \le \infty} (x, e_n) e_n\|_{L^2(\mathcal{R})} = 0$.) The value $\|h_n\|_{L^2(\mathcal{R})} = \left[\frac{(4\pi)^n}{\sqrt{2} n!} \right]^{1/2}$.

Now for $x \in L^2(\mathcal{R})$, we have Weiner's formula:

$$x^\wedge = \sum_{0 \le n \le \infty} (x, e_n)(-i)^n e_n.$$

Exercise 4.14: Why isn't Weiner's formula $x^\wedge = \sum_{0 \le n \le \infty} (x, e_n)(-i)^n \|h_n\|_{L^2(\mathcal{R})} e_n$?

Define $E_j = \{x \in L^2(\mathcal{R}) \mid x = \sum_{0 \le n \le \infty} (x, e_{4n+j}) e_{4n+j}\}$ for $j \in \{0, 1, 2, 3\}$. The subspaces E_0, E_1, E_2, and E_3 are the eigenspaces in $L^2(\mathcal{R})$ with respect to the linear operator $\wedge$, and, since $\wedge$ has a complete set of eigenfunctions, $L^2(\mathcal{R}) = E_0 \oplus E_1 \oplus E_2 \oplus E_3$.

Note $\wedge$ on E_j is just multiplication by $(-i)^j$, *i.e.*, $f^\wedge = (-i)^j f$ for $f \in E_j$. And multiplication of a complex number by $(-i)^j$ just "rotates" that number in the complex plane by $-j\pi/2$ radians.

Exercise 4.15: Show that $h_0, h_1, \ldots,$ are eigenfunctions of $\vee$, the inverse Fourier integral transform with $h_n^\vee = (-i)^n h_n$. Hint: $\wedge \wedge \wedge = \vee$.

5 Fourier Integral Transforms of Linear Functionals

We can extend the Fourier transform operator to certain functions, x, such that $\int_{-\infty}^{\infty} |x(t)| dt = \infty$, for example: $x(t) = 1$ or $x(t) = sin(t)$. We cannot use the integral formula $x^\wedge(s) = \int_{-\infty}^{\infty} x(t) e^{-2\pi ist} dt$ to define the Fourier transform in such cases, however, without going in a round-about path and introducing the famous Dirac-δ functional (although Paul Dirac was not the earliest discoverer.)

The basic approach comes from the observation, derived from Parseval's identity, that $(x^\wedge, y^*)_{L^2(\mathcal{R})} = (x, y^{\wedge *})_{L^2(\mathcal{R})}$, together with the observation that $(x, y^*)_{L^2(\mathcal{R})} = \int_{-\infty}^{\infty} x(t) y(t) dt$ may exist for functions x that do not decrease, *i.e.*, do not satisfy $x(t) \to 0$ as $|t| \to \infty$, as long as y decreases rapidly enough to compensate. Therefore we can try to *define* the Fourier transform $x^\wedge$ for certain slowly-decreasing or non-decreasing functions x. The idea is that, given x, we can "solve" for $x^\wedge$ in the infinite set of functional equations

$$\int_{-\infty}^{\infty} x^\wedge(s)y(s)ds = \int_{-\infty}^{\infty} x(t)y^\wedge(t)dt,$$

obtained as y ranges over a suitable set of decreasing functions such as $C_\downarrow^\infty(\mathcal{R})$ with respect to which we re-define the Fourier transform; these functions are often called *test* functions. (Actually, the more restrictive the class of test functions that y ranges over, the more "expansive" the class to which x belongs can be.)

In carrying out this program, we soon find that if we want to solve for the Fourier transform of a polynomial, or even just of $x(t) = 1$, we must be willing to take the Fourier transform $x^\wedge$ which we are trying to construct to be a so-called *linear functional* (the terms "distribution" or "generalized function" are also used.) This is the case essentially because the relationship between the "spread" of a function x and the "concentration" of its Fourier transform $x^\wedge$ continues to hold as we take limits and forces us to consider objects such as the δ functional.

Recall that, in general, a *linear functional* on a vector space V with the field of scalars $\mathcal{C}$ is a function F mapping elements of V to complex values such that $F(\alpha x + \beta y) = \alpha F(x) + \beta F(y)$ for $x, y \in V$ and $\alpha, \beta \in \mathcal{C}$. The linear functionals on V are just the simplest class of linear transformations on V, those of rank 0 or 1 whose range is just C. Let the set of linear functionals on V be denoted by $V^\top$; $V^\top$ is also a vector space called the *dual space* of V, although, unlike the situation in finite-dimensional spaces, V and $V^\top$ are not necessarily isomorphic. Various linear functionals are often defined with respect to an inner-product operation $(\cdot,\cdot)$ defined on V, thus we may have $F(x) = (x, f)$ where $f \in V$ is a fixed element of V corresponding to the linear functional F; it can even be the case that (x, f) is defined when $f \notin V$.

We are specifically interested in the situation where V is the space of functions $L^2(\mathcal{R})$; in this case a linear functional is a function that maps functions in $L^2(\mathcal{R})$ to complex numbers. (The word 'functional' is used in an attempt to disambiguate what kind of function we are discussing.)

It is convenient to define the pseudo-inner-product $\langle f, g\rangle := (f, g^*)_{L^2(\mathcal{R})}$. Note $\langle f, g\rangle = \langle g, f\rangle$. The integral $\langle f, g\rangle$ is the continuous analog of the finite-dimensional Euclidean inner-product which we know represents the application of a linear functional defined by a covector to a vector in $\mathcal{R}^n$. (Also, note that $\langle f, g\rangle = (f, g)$ when f and g are real-valued functions.) Thus in the same way, we may take $\langle f, g\rangle$ as representing the application of a linear functional, F, defined by $F(g) = \int_{-\infty}^{\infty} f(t)g(t)dt$ to obtain a complex number, or symmetrically, as representing the application of a linear functional, G, defined by $G(f) = \int_{-\infty}^{\infty} f(t)g(t)dt$. Note here, the linear functional F is associated with a corresponding function f, (and the linear functional G is associated with a corresponding function g).

Exercise 5.1: Show that $\langle x, y^R\rangle = \langle x^\wedge, y^\wedge\rangle$.

Solution 5.1: We have $(x, y^{\wedge *}) = (x^\wedge, y^*)$, so $(x, y^{\wedge\wedge *}) = (x^\wedge, y^{\wedge *})$, and $y^{\wedge\wedge *} = y^{R*}$, so $\langle x, y^R\rangle = \langle x^\wedge, y^\wedge\rangle$, or equivalently, $\langle y^R, x\rangle = \langle y^\wedge, x^\wedge\rangle$.

For the function space $L^2(\mathcal{R})$, we can define continuous linear functionals as follows. Let F be a linear functional on $L^2(\mathcal{R})$. If, for every sequence of functions $x_1, x_2, \ldots \in L^2(\mathcal{R})$ that converges to a function $x \in L^2(\mathcal{R})$ in the sense that $\|x_n - x\|_{L^2(\mathcal{R})} \to 0$ as $n \to \infty$, the corresponding sequence of complex numbers $F(x_1), F(x_2), \ldots$ converges to the value $F(x)$, then the linear functional $F \in L^2(\mathcal{R})^\top$ is *continuous*. We can also define bounded linear functionals: the linear functional $F \in L^2(\mathcal{R})^\top$ is *bounded* exactly when there exists a real constant b such that $|F(x)| \le b\|x\|_{L^2(\mathcal{R})}$ for all $x \in L^2(\mathcal{R})$; in other words, $F(x)$ does not produce a result greater than $O(\|x\|_{L^2(\mathcal{R})})$. Any continuous linear functional is bounded and vice versa.

An important class of linear functionals in $L^2(\mathcal{R})^\top$ are those linear functionals, F, that correspond to functions, f, in $L^2(\mathcal{R})$ itself. We write $[f]_L = F$ to show this correspondence. For each function $f \in L^2(\mathcal{R})$, the corresponding linear functional, $[f]_L$, is computed on $x \in L^2(\mathcal{R})$ as $[f]_L(x) = \langle x, f\rangle = \langle f, x\rangle = \int_{-\infty}^{\infty} x(t)f(t)dt$, which is guaranteed to exist because the product $|f||x|$ approaches 0 sufficiently quickly.

The dual space $L^2(\mathcal{R})^\top$ also contains linear functionals corresponding to other admissible functions, f, defined on $\mathcal{R}$, such as $f(t) = t$, or $f(t) = cos(t)$, which are not found in $L^2(\mathcal{R})$. These functions have the property that the integral $\langle x, f\rangle$ exists for $x \in L^2(\mathcal{R})$. (Such admissible functions grow no faster than polynomials, as you might suppose from the definition of the Schwartz space, $C_\downarrow^\infty(\mathcal{R})$.) We shall denote the class of non-$L^2(\mathcal{R})$ admissible functions as $NL^2(\mathcal{R})$, and for $f \in NL^2(\mathcal{R})$, we shall write $[f]_{NL}$ to denote the corresponding linear functional. The linear functionals based on either $L^2(\mathcal{R})$-functions or on proper admissible $NL^2(\mathcal{R})$-functions are called *regular* linear functionals. For $f \in L^2(\mathcal{R}) \cup NL^2(\mathcal{R})$, we shall just write $[f]$ to denote the corresponding linear functional when we are indifferent as to whether $f \in L^2(\mathcal{R})$ or $f \in NL^2(\mathcal{R})$. Finally, it is convenient to define $RL^2(\mathcal{R}) := L^2(\mathcal{R}) \cup NL^2(\mathcal{R})$. [Are *all* the admissible functions in $NL^2(\mathcal{R})$, ordinary pointwise limits of $C_\downarrow^\infty(\mathcal{R})$-functions (or even stronger, limits of functions with compact support, where general $C_\downarrow^\infty(\mathcal{R})$-functions are obtained by letting the support sets of the sequence functions grow larger and larger?]

There are linear functionals in $L^2(\mathcal{R})^\top$ which are not regular. The most famous such linear functional is the Dirac-δ functional. The non-regular linear functionals in $L^2(\mathcal{R})^\top$ are limits of certain sequences of regular linear functionals in $L^2(\mathcal{R})^\top$ (they may also be the limits of many other arbitrary sequences in $L^2(\mathcal{R})^\top$.) Recall that in order for a sequence of functions $x_1, x_2, \ldots \in L^2(\mathcal{R})$ to converge (in norm) to a function $x \in L^2(\mathcal{R})$, we must have $\|x_n - x\|_{L^2(\mathcal{R})} \to 0$ as $n \to \infty$. In contrast, to have a sequence of regular linear functionals $[x_1], [x_2], \ldots \in L^2(\mathcal{R})^\top$ converge to a linear functional $X \in L^2(\mathcal{R})^\top$, we require that $|[x_n](y) - X(y)| \to 0$ as $n \to \infty$ for all test functions $y \in C_\downarrow^\infty(\mathcal{R})$. In this case, we say that $[x_1], [x_2], \ldots$ *functionally-converges* to $X \in L^2(\mathcal{R})^\top$. By extension, we shall define the linear functional $X \in L^2(\mathcal{R})^\top$ to be the *functional limit* of the sequence of (not-necessarily regular) linear functionals $X_1, X_2, \ldots$ in $L^2(\mathcal{R})^\top$ when $|X_n(y) - X(y)| \to 0$ as $n \to \infty$ for all $y \in C_\downarrow^\infty(\mathcal{R})$.

A non-regular linear functional, G, applied to a function $x \in L^2(\mathcal{R})$ is computed as $G(x) = \langle G, x\rangle = \int_{-\infty}^{\infty} \underline{G}(t)x(t)dt = \int_{-\infty}^{\infty}[\lim_{n\leftarrow\infty} g_n(t)]x(t)dt$ where $\underline{G}(t) = \lim_{n\leftarrow\infty} g_n(t)$ with $g_n \in RL^2(\mathcal{R})$. Note G denote a linear *functional* and $\underline{G}$ denotes the limit of a sequence of functions of a real argument. We have to reach for a name for the limit "function" $\lim_{n\to\infty} g_n(t)$ since this limit function is not a proper function defined on $\mathcal{R}$. This issue is akin to a sequence of rational numbers converginging to an irrational number outside the field of rational numbers where the sequence members reside.

We may abuse notation by writing $G(x)$ to denote the value of the linear functional G on the function x and writing $G(t)$ to denote the value of the associated limit $\underline{G}$ on the real value t; thus the *type* of the asrgument of G determines whether we mean $G(x)$ or $\underline{G}(t)$. (We generally mean $\underline{G}(t)$ within integrals.)

Recall that $L^2(\mathcal{R})$ is the completion of $C_{\downarrow}^{\infty}(\mathcal{R})$ with respect to the metric based on the norm $\|\cdot\|_{L^2(\mathcal{R})}$. The dual space $L^2(\mathcal{R})^{\top}$ contains the regular linear functionals corresponding to the elements of $C_{\downarrow}^{\infty}(\mathcal{R})$, plus the regular linear functionals corresponding to functions $f \in L^2(\mathcal{R}) - C_{\downarrow}^{\infty}(\mathcal{R})$, plus the additional regular linear functionals corresponding to functions $f \in NL^2(\mathcal{R})$ such that the integral $\langle x, f\rangle$ exists for all $x \in L^2(\mathcal{R})$, together with all the functional limits of functionally-convergent sequences of regular linear functionals. These limits include many non-regular linear functionals.

Note the correspondence between functions, f, and regular linear functionals, $[f]$, such that $[f](x) = \langle f, x\rangle = \int_{-\infty}^{\infty} f(t)x(t)dt$, is a many-to-one relationship. In general there are uncountably many functions, f_j, that are equivalent to f in the sense that $\langle f_j, x\rangle = \langle f, x\rangle$ for all test functions x; in this case, $f_j = f$ a.e. Thus, if $f = g$ a.e. for $f, g \in L^2(\mathcal{R})$, then we have $[f] = [g]$, *i.e.*, the corresponding regular linear functionals are equal. In general, however, we want a definition of equality for arbitrary linear functionals. For $F, G \in L^2(\mathcal{R})^{\top}$, we say that $F = G$ exactly when $F(x) = G(x)$ for all $x \in L^2(\mathcal{R})$.

The reason that regular linear functionals are infinitely differentiable, and in most other ways more "pliable" than $RL^2(\mathcal{R})$-functions, is that, of all the functions, f_j, that satisfy $[f_j] = [f]$, we may pick the smoothest such function to represent the linear functional $[f]$; we thus need not deal with the many wilder equivalent functions.

We will redefine the integral Fourier transform and its inverse to apply to linear functionals in $L^2(\mathcal{R})^{\top}$, *i.e.*, $\wedge : L^2(\mathcal{R})^{\top} \to L^2(\mathcal{R})^{\top}$. However, for $x \in L^2(\mathcal{R})$, the "new" Fourier transform $[x]_L^{\wedge}$ of the regular linear functional $[x]_L$ is just the regular linear functional $[x^{\wedge}]_L$ where $x^{\wedge}$ is the "old" Fourier transform of the function x. What is different is the wealth of new regular linear functionals corresponding to functions in $NL^2(\mathcal{R})$ and the new non-regular linear functionals that may now be assigned Fourier transforms. Many of these non-regular linear functionals have Fourier transforms involving the Dirac-δ functional.

5.1 The Dirac-δ Functional

The Dirac-δ functional, δ, is defined by $\delta(x) = x(0)$ for $x \in L^2(\mathcal{R})$.

Note the Dirac-δ functional is continuous on $C_{\downarrow}^{\infty}(\mathcal{R})$, but it is not continuous on $L^2(\mathcal{R})$. This is because the notion of convergence used to construct $L^2(\mathcal{R})$ from $C_{\downarrow}^{\infty}(\mathcal{R})$ is based on integration: $\|x_n - x\|_{L^2(\mathcal{R})} \to 0$ as $n \to \infty$, and this permits functions that are discontinuous at 0 to be limit points in $L^2(\mathcal{R})$ such that $x(0)$ is not the limit of $x_1(0), x_2(0), \ldots$, even though we may still claim that the sequence $x_1, x_2, \ldots$ converges in the $L^2(\mathcal{R})$-norm to x.

The Dirac-δ functional, δ, is often written as $\delta(t)$ with a parameter t appearing as a "phantom" argument so that $\delta(t)(x) = x(0)$. This notation is useful within formal integral expressions such as $\int_{-\infty}^{\infty} \delta(t)x(t)dt =: \langle x, \delta\rangle := x(0)$. In this context, the translated δ functional, $\delta(t-a)$ arises, where

$\delta(t-a)(x) = \delta(t)(y) = x(a)$ where $y(t)$ is the translate $x(t+a)$ (in an integral expression we have $\int_{-\infty}^{\infty} \delta(t-a)x(t)dt = \int_{-\infty}^{\infty} \delta(t)x(t+a)dt = x(a)$.) We prefer to write δ_a in place of $\delta(t-a)$ so that the functional δ may be written δ_0, but the function-argument notation is too useful in integral expressions to be abandoned.

> **Exercise 5.2:** Let T_a denote the operation of translation by a constant, a, so that $T_a(x) = y$, where $y(t) = x(t+a)$. Show that T_a is a linear tranformation on $L^2(\mathcal{R})$, and show that δ_a is the composition $\delta_0 T_a$.

We can construct the Dirac-δ functional as follows. Suppose f is continuous at 0. Choose $\epsilon_1, \epsilon_2, \ldots$ to be a decreasing sequence of positive values with $\epsilon_n \downarrow 0$ as $n \to \infty$. Then, for each ϵ_n there exists a positive value a_n such that $|f(t) - f(0)| \le \epsilon_n$ when $|t| < a_n$. The values $a_1, a_2, \ldots$ can always be taken to be a sequence of positive values a_n with $a_n \downarrow 0$ as $n \to \infty$.

Now let $p_n(t) = K(a_n)e^{-t^2/(a_n^2 - t^2)}$ where $K(a_n) = \left[\int_{-a_n}^{a_n} e^{-t^2/(a_n^2-t^2)}dt\right]^{-1}$ where $a_n > 0$. Note $p_n(t) \ge 0$, $\int_{-a_n}^{a_n} p_n(t)dt = 1$, and $p_n^* = p_n$.

Now, since $\int_{-a_n}^{a_n} p_n(t)dt = 1$,

$$
\begin{aligned}
|\,[p_n](f) - f(0)\,| &= \left| \int_{-a_n}^{a_n} f(t)p_n(t)dt - f(0)\int_{-a_n}^{a_n} p_n(t)dt \right| \\
&= \left| \int_{-a_n}^{a_n} (f(t) - f(0))p_n(t)dt \right| \\
&\le \int_{-a_n}^{a_n} |f(t) - f(0)| p_n(t)dt \\
&\le \int_{-a_n}^{a_n} \epsilon_n p_n(t)dt \\
&= \epsilon_n.
\end{aligned}
$$

And $\epsilon_n \downarrow 0$ as $n \to \infty$. Thus, $[p_n] \to \delta_0$ as $n \to \infty$ where $[p_n]$ is the regular linear functional associated with p_n, *i.e.*, $[p_n](f) \to \delta_0(f) = f(0)$ as $n \to \infty$.

> **Exercise 5.3:** Describe the graph of the function p_n on the interval $(-a_n, a_n)$.

> **Exercise 5.4:** Does the fact that $p_n(t) \ge 0$ have any import?

Let $m_1, m_2, \ldots$ be a sequence of positive real values with $m_n \uparrow \infty$ as $n \to \infty$ (for example: $m_n = n$.) Then in place of the sequence $p_1, p_2, \ldots$, we can use any sequence $q_1, q_2, \ldots$, with $q_n(t) = m_n q(m_n t)$ where $q(t)$ is an infinitely-smooth $C_{\downarrow}^{\infty}(\mathcal{R})$-function with support on $[-1, 1]$ such that $\int_{-\infty}^{\infty} q(t)dt = 1$. For any such sequence, we have $[q_n] \to \delta_0$ as $n \to \infty$. The transformation $q \to m_n q(m_n t)$ reduces the support set from $[-1, 1]$ to $[-1/m_n, 1/m_n]$ while increasing the "height" so as to maintain $\int_{-\infty}^{\infty} q_n(t)dt = 1$.

Note that $p_n \to 0$ a.e. as $n \to \infty$, but $\|p_n\|_{L^2(\mathcal{R})} = 1$ for all n. In other words, δ_0 is not a regular functional obtained from a member of $L^2(\mathcal{R})$ since $L^2(\mathcal{R})$ contains only those functions, f, that satisfy $\|f - f_n\|_{L^2(\mathcal{R})} \to 0$ as $|n| \to \infty$ for some sequence of functions $f_1, f_2, \ldots$ found in

$C_{\downarrow}^{\infty}(\mathcal{R})$. However, we have just seen that $\langle p_n, f\rangle \to \delta_0(f) = f(0)$ as $n \to \infty$ for all $f \in C_{\downarrow}^{\infty}(\mathcal{R})$, or equivalently, $[p_n] \to \delta_0$ as $n \to \infty$. It is this notion of limit that we use to define the limit of a sequence of linear functionals.

Exercise 5.5: Show that $p_1, p_2, \ldots$ is not a Cauchy sequence in $L^2(\mathcal{R})$ with respect to the metric based on the norm $\|\cdot\|_{L^2(\mathcal{R})}$.

Exercise 5.6: Show that $\boxed{\delta(at) = \frac{1}{|a|}\delta(t)}$ for $a \in \mathcal{R}$ with $a \neq 0$. (This is an example where use of a phantom argument with the δ functional simplifies matters.)

Solution 5.6: Recall that the δ functional applied to a test function, f, in $C_{\downarrow}^{\infty}(\mathcal{R})$ can be defined by $\lim_{n\to\infty}\int_{-\infty}^{\infty} p_n(t)f(t)\,dt$. Note p_n is an even function, so $p_n(at) = p_n(-at) = p_n(|a|t)$. Moreover, for $a \neq 0$, $\int_{-\infty}^{\infty} |a|p_n(|a|t)\,dt = 1$, so the sequence $|a|p_1(|a|t), |a|p_2(|a|t), \ldots$ is another suitable sequence for defining the δ functional, *i.e.*, $[|a|p_n(|a|t)] \to \delta(t)$ as $n \to \infty$.

Thus,

$$\begin{aligned}
\delta(at)(f) &= \lim_{n\to\infty}\int_{-\infty}^{\infty} p_n(at)f(t)\,dt \\
&= \lim_{n\to\infty}\int_{-\infty}^{\infty} p_n(|a|t)f(t)\,dt \\
&= \lim_{n\to\infty}\int_{-\infty}^{\infty} p_n(|a|t)f(t)\,dt \\
&= \lim_{n\to\infty}\int_{-\infty}^{\infty} \frac{1}{|a|}|a|p_n(|a|t)f(t)\,dt \\
&= \frac{1}{|a|}\lim_{n\to\infty}\int_{-\infty}^{\infty} |a|p_n(|a|t)f(t)\,dt \\
&= \frac{1}{|a|}\delta(t)(f),
\end{aligned}$$

and hence $\delta(at) = \frac{1}{|a|}\delta(t)$.

Exercise 5.7: Show that $\delta(a(t-r)) = \frac{1}{|a|}\delta(t-r)$ for $a \in \mathcal{R}$ with $a \neq 0$.

The above construction computes $\delta_0(f)$ as the limit of a sequence of integrals: $\langle f, p_n\rangle = \int_{-\infty}^{\infty} f(t)p_n(t)dt \to f(0)$ as $n \to \infty$. We generally abuse notation and indicate this by writing $f(0) = \int_{-\infty}^{\infty} f(t)\delta(t)dt$. Also, since $p_n(t)^* = p_n(t)$, we have $\delta(t)^* = \delta(t)$, so, continuing to abuse notation, $f(0) = \int_{-\infty}^{\infty} f(t)\delta(t)^*dt =: (f, \delta_0)_{L^2(\mathcal{R})} =: \delta_0(f) = \langle\delta_0, f\rangle$.

If $g \in RL^2(\mathcal{R})$, then the product $g\delta_r$ is a linear functional on $L^2(\mathcal{R})$. Specifically $(g\delta_r)(f) = \langle g(t)\delta(t-r), f(t)\rangle = \langle\delta_r, gf\rangle = g(r)f(r)$, and $\langle g(r)\delta_r, f\rangle = g(r)f(r)$, so we may write $g(t)\delta_r = g(r)\delta_r$.

In general, we define $[h](f) := \langle h, f\rangle = \int_{-\infty}^{\infty} f(t)h(t)dt$ when $[h]$ is a regular linear functional based on a function $h \in RL^2(\mathcal{R})$, and we define $G(f)$ specifically, not necessarily involving an integral expression, when G is not a regular linear functional, although we still use the integral notation $\langle G, f\rangle$ to indicate the result of applying G to f. (In this situation, $\langle G, f\rangle$ is to be taken as shorthand for $\lim_{n\to\infty}\langle g_n, f\rangle$ where $g_1, g_2, \ldots$ are $RL^2(\mathcal{R})$-functions such that the sequence of regular linear functionals $[g_1], [g_2], \ldots$ functionally-converges to G.)

We may interpret the integral $\int_{-\infty}^{\infty} f(t)h(t)dt$ as the Stieljes integral $\int_{-\infty}^{\infty} f(t)dm_h$ where m_h is a measure whose (Radon-Nikodym) derivative is the function h. Moreover, rather than compute the limit $\lim_{n\to\infty}\int_{-\infty}^{\infty} f(t)p_n(t)dt$ to get the value $f(0)$ corresponding to the symbolic result $\int_{-\infty}^{\infty} f(t)\delta(t)dt$, we can compute $\lim_{n\to\infty}\int_{-\infty}^{\infty} f(t)dm_{p_n} = \int_{-\infty}^{\infty} f(t)dm_\delta$ where m_δ is the discrete measure induced by the Dirac-δ functional: $m_\delta(s) = \begin{cases} 1, & \text{if } 0 \in s, \\ 0, & \text{otherwise.}\end{cases}$ (Note this is equivalent to $m_{p_n} \to m_\delta$ as $n \to \infty$, which is a perfectly acceptable statement about the limit of a sequence of measures.) It is a slightly lesser abuse of notation if we take $\int_{-\infty}^{\infty} f(t)\delta(t)dt$ to mean $\int_{-\infty}^{\infty} f(t)dm_\delta$, even though the Radon-Nikodym derivative of m_δ does not exist (since δ is not a proper function.)

Note that a regular linear functional can be converted to a proper function, and that function can be used in applications or derivations of results. Non-regular linear functionals, however, can really only be applied to functions in $L^2(\mathcal{R})$ to obtain complex numbers. By abusing notation as discussed above, non-regular linear functionals can be used and manipulated in certain integral expresions.

Exercise 5.8: Show that the δ functional is not bounded with respect to the norm $\|\cdot\|_{L^2(\mathcal{R})}$ (*i.e.*, there does not exist a constant b such that $|\delta_0(f)| \le b\|f\|_{L^2(\mathcal{R})}$ for all $f \in L^2(\mathcal{R})$. (Thus the Riesz representation theorem does not apply to the δ functional.) Hint: Consider a sequence of functions $p_1, p_2, \ldots$ in $L^2(\mathcal{R})$ for which $p_n(0) \to \infty$ as $n \to \infty$.

5.2 Functional Derivatives

Now we may construct the functional derivative of the δ functional. To do this we just take $\delta' = \lim_{n\to\infty}[p'_n]$. Thus,

$$\begin{aligned}
[p'_n](f) &= \int_{-a_n}^{a_n} f(t)p'_n(t)dt \\
&= f(t)p_n(t)\Big|_{t=-a_n}^{t=a_n} - \int_{-a_n}^{a_n} f'(t)p_n(t)dt \\
&= f(a_n)p_n(a_n) - f(-a_n)p_n(-a_n) - \int_{-a_n}^{a_n} f'(t)p_n(t)dt,
\end{aligned}$$

and $p_n(a_n) = p_n(-a_n) = 0$, so $[p'_n](f) = -\int_{-a_n}^{a_n} f'(t)p_n(t)dt \to -\delta_0(f')$ as $n \to \infty$.

Thus, $\delta'_0(f) = -\delta_0(f') = -f'(0)$, and in general, $\delta_0^{(k)}(f) = (-1)^k\delta_0(f^{(k)}) = (-1)^k f^{(k)}(0)$ when f is k-fold continuously-differentiable at 0. (Beware: if X is a function, then X' is the ordinary derivative of X, but if X is a linear functional, then X' is the functional derivative of X.)

Exercise 5.9: Show that $\boxed{\delta_r^{(k)}(f) = (-1)^k f^{(k)}(r)}$ for $r \in \mathcal{R}$ where f is k-fold continuously-differentiable at r.

Exercise 5.10: By definition, $\delta_r(u) = 1$ for $u(t) = 1$ (note $u \in NL^2(\mathcal{R})$.)
Show that $\int_{-\infty}^{\infty} \delta(t-r)dt := \lim_{n\to\infty} \int_{-\infty}^{\infty} p_n(t-r) \cdot 1\, dt = 1$. This is consistent with $\langle \delta_r, u \rangle = 1$.

The above development of δ_0 and $\delta_0^{(k)}$ is based on using a suitable sequence of regular $L^2(R)^\top$-functionals $[p_1], [p_2], \ldots$. There are many other sequences of functionals in $L^2(R)^\top$ that converge to δ_0, but not all of these are composed solely of regular functionals, and hence may not be ameniable to forming a sequence of integrals that converge to the value $\delta_0(f)$.

Exercise 5.11: Show that $[H]$ is a regular linear functional where $H(t)$ is the Heaviside step-function: $H(t) = \begin{cases} 0, & \text{if } t < 0, \\ \frac{1}{2}, & \text{if } t = 0, \\ 1, & \text{if } t > 0. \end{cases}$ Give a sequence of regular linear functionals $[h_1], [h_2], \ldots$ based on continuous functions $h_1, h_2, \ldots$ that converge to $[H]$. Note $H(t) = 1 - B_0(t)$. Can $h_1, h_2, \ldots$ be chosen to be $C_\downarrow^\infty(\mathcal{R})$-functions? Hint: enforce $h_n(0) = \frac{1}{2}$ for $n > 0$.

Exercise 5.12: Show that the functional d_{ba}, defined for $b > a$ by $d_{ba}(f) = (f(b)-f(a))/(b-a)$ is a non-regular linear functional. Note that the non-regular linear functional d_a defined by $d_a(f) = f'(a)$ satisfies $d_a = \lim_{b\downarrow a} d_{ba}$. Hint: $d_{ba} = (\delta_b - \delta_a)/(b-a)$.

Exercise 5.13: Consider a sequence of $C_\downarrow^\infty(\mathcal{R})$-functions $q_1, q_2, \ldots$, where q_n has support $(-a_n, a_n)$ with $a_n > 0$, such that $[q_n] \to Q \in L^2(\mathcal{R})^\top$. Show that the functional derivative, Q' defined by $\lim_{n\to\infty}[q_n']$ satisfies $Q'(f) = -Q(f')$ for all $f \in C_\downarrow^\infty(\mathcal{R})$, and hence we have $\boxed{Q^{(k)}(f) = (-1)^k Q(f^{(k)})}$ for $Q \in L^2(\mathcal{R})^\top$. (We use this relation as the definition of the functional derivative $Q^{(k)}$ of $Q \in L^2(\mathcal{R})^\top$. In prefix operator terms, we have $Q^{(k)} = (-1)^k Q D_a^k$, where a denotes the argument of whatever function D_a is applied to.)

Now let us consider the functional derivative of $[H]$. We have

$$\begin{aligned} [H]'(f) &= (-1)[H](f') \\ &= -\int_{-\infty}^{\infty} H(t) f'(t)dt \\ &= -\int_{0}^{\infty} f'(t)dt \\ &= -f(t)|_{t=0}^{t=\infty} \\ &= -(f(\infty) - f(0)) \\ &= f(0), \end{aligned}$$

Thus, $[H(t)]' = \delta(t)$.

Similarly, $[H(t-a)]' = \delta(t-a)$ and $[H(-(t-a))]' = \delta(t-a)$ as well.

Exercise 5.14: Use integration by parts: $\int_a^b pq' = p(b)q(b) - p(a)q(a) - \int_a^b p'q$ to show that $[H]' = \delta_0$.

Exercise 5.15: Define $sign(t) = 2H(t) - 1$. Show that $sign'(t) = 2H'(t) = 2\delta(t)$.

Exercise 5.16: Show that $\int_{-\infty}^{t} \delta(r)dr = H(t)$. Hint: use the sequence $p_1, p_2, \ldots$.

We can generalize the computation of $[H]'$ to see that any piecewise-continuous function, x, with a piecewise-continuous derivative can have its functional derivative expressed in terms of δ functionals.

Let $\ldots, t_{-2}, t_{-1}, t_0, t_1, t_2, \ldots$ be the separated points at which x is discontinuous, or at which x' is not defined. (In the first case, we have a jump discontinuity (or a punctured jump discontinuity) of x, and in the second case, we have a cusp of x whereat x' has a punctured jump discontinuity.) Now define $j_x(t_k) = \lim_{\epsilon\downarrow 0} x(t_k+\epsilon) - x(t_k-\epsilon)$ (this is often written $x(t_k+) - x(t_k-)$.) Then we can write:

$$x(t) = x_c(t) - \sum_{k=-\infty}^{\infty} |j_x(t_k)| \cdot H(sign(j_x(t_k))(t-t_k)),$$

where x_c is a continuous function in $RL^2(\mathcal{R})$.

In essence, when, as t increases, x jumps upward by the amount m at t_k, we *reknit* x together by subtracting the translated step-function $|m|H(t-t_k)$, and when x jumps downward by the amount m at t_k, then we reknit x together by subtracting the translated and reflected step-function $|m|H(-(t-t_k))$.

This "reknitted" form of x forms the function x_c. The function x_c will generally have a punctured discontinuity at t_k (because $H(0) = 1/2$,) but these singular points can be fixed by redefining the knitted function $x_c(t)$ at $t = t_k$ as $\lim_{t\downarrow t_k} x_c(t)$. This makes x_c a continuous function, and, although x_c may not be differentiable at the knit-points, the corresponding linear functional $[x_c]$ has a functional derivative $[x_c]'$, since this functional derivative is defined in terms of integration with respect to test functions in $C_{\downarrow}^{\infty}(\mathcal{R})$ and problems on a set of measure 0 are of no matter.

Thus, we have:

$$[x]' = [x_c]' - \sum_{k=-\infty}^{\infty} j_x(t_k) \cdot \delta(t-t_k).$$

This is because $[|m| \cdot H(sign(m)(t-t_k))]' = |m| \cdot \delta(sign(m)(t-t_k)) \cdot sign(m) = |m| \dfrac{1}{|\,sign(m)|} \cdot \delta(t-t_k) \cdot sign(m) = m\delta(t-t_k)$.

Thus, a δ functional appears in the functional derivative of a regular linear functional at each jump discontinuity, scaled by the size of the jump.

5.3 Convolution of Linear Functionals

We can extend the convolution operation to apply to linear functionals as follows. First we define the convolution of a linear functional G with a regular linear functional $[x]$ for $x \in RL^2(\mathcal{R})$ as $(G \circledast [x])(r) := \langle G, x(r-t)\rangle = \int_{-\infty}^{\infty} G(t)x(r-t)dt$. Note $(G \circledast [x])(r)$ is a linear functional determined by the convolution parameter r.

Also note the translated linear functional $G(t-s)$ is defined in terms of the linear functional G as: $\langle G(t-s), x\rangle = \int_{-\infty}^{\infty} G(t-s)x(t)dt = \int_{-\infty}^{\infty} G(t)x(t+s)dt = \langle G, x(t+s)\rangle$. Thus $(G \circledast [x])(r-s)$ has a meaning, and it makes sense to assert $\langle G, x(r-t)\rangle = \langle G(r-t), x\rangle$.

Now we may define $F \circledast G$ for $F, G \in L^2(\mathcal{R})^\top$ as $F \circledast G = E \in L^2(\mathcal{R})^\top$ where $E(x) = (F \circledast (G \circledast x^R))(0)$ for $x \in RL^2(\mathcal{R})$. (Note the similarity to the δ functional definition. This suggests there is an entire hierarchy of linear functionals, where the linear functionals at level j are applicable to the linear functionals at level $j-1$. Indeed, the Fourier transforms $\wedge$ and $\vee$ can be seen as linear functionals defined on linear functionals of $L^2(\mathcal{R})$.)

Thus,

$$
\begin{aligned}
\langle F \circledast G, x\rangle &= (F \circledast (G \circledast x^R))(0) \\
&= \left[\int_{-\infty}^{\infty} F(u)\left[\int_{-\infty}^{\infty} G(v)x^R(s-v)dv\right]_{s=t-u} du\right]_{t=0} \\
&= \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u)G(v)x(v-(t-u))dvdu\right]_{t=0} \\
&= \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u)G(v)x(v+u)dvdu\right].
\end{aligned}
$$

Note for the functional convolution $F \circledast G$ to be commutative, it is sufficient that at least one of the linear functionals F, G have bounded support. Similarly, for the functional convolution $(E \circledast F) \circledast G$ to be associative, (*i.e* $(E \circledast F) \circledast G = E \circledast (F \circledast G)$), it is sufficient that at least two of the linear functionals E, F, G have bounded support. (The support set of a linear functional $F \in L^2(\mathcal{R})^\top$ is the set $\mathcal{R} - \cup_{W \in V_F} W$, where V_F is the set of all open subsets, W, of $\mathcal{R}$ such that $F(g) = 0$ for all those functions $g \in L^2(Q)$ whose support is within W, *i.e.*, $F(g) = 0$ for $g \in \{f \in L^2(\mathcal{R}) \mid \mathit{support}(f) \subseteq W\}$ [Rud98]. Basically, the support set of F is the set of points in $\mathcal{R}$ where F "pays attention", that is, the complement of the set of points in $\mathcal{R}$ where F "pays no attention" and returns 0 for any argument function that can deviate from 0 only outside $\mathit{support}(F)$.)

Exercise 5.17: Show that $\mathit{support}(\delta_0) = \{0\}$.

Exercise 5.18: Show that $\delta_0 \circledast x = [x]$ for $x \in RL^2(\mathcal{R})$, where $(\delta_0 \circledast x)(r) = \langle \delta_0(t), x(r-t)\rangle$. Also show that $(\delta_a \circledast x)(r) = [x(r-a)]$.

Exercise 5.19: Show that $\delta_a \circledast \delta_b = \delta_{a+b}$.

Solution 5.19:

$$
\begin{aligned}
(\delta_a \circledast \delta_b)(f) &= \langle \delta_a \circledast \delta_b, f\rangle \\
&= (\delta(t-a) \circledast \delta_(t-b))(f) \\
&= \langle \int_{-\infty}^{\infty} \delta(t-a)\delta_(s-t-b))dt, f(s)\rangle \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(t-a)\delta_(s-t-b))f(s)\,dt\,ds \\
&= \int_{-\infty}^{\infty} \delta(t-a)\int_{-\infty}^{\infty} \delta_(s-t-b))f(s)\,ds\,dt \\
&= \int_{-\infty}^{\infty} \delta(t-a)f(t+b)dt
\end{aligned}
$$

$$\begin{aligned} &= f(a+b) \\ &= \langle \delta_{a+b}, f \rangle. \end{aligned}$$

Therefore $\delta_a \circledast \delta_b = \delta_{a+b}$.

In general, $(g \circledast H)(r) = \int_{-\infty}^{r} g(t)dt$ where H is the Heaviside step-function. Thus convolution by H is integration for $g \in C_{\downarrow}^{\infty}(\mathcal{R})$(!) Also, $F = \delta \circledast F$ (*i.e.*, δ is a unit in the convolution algebra $L^2(\mathcal{R})^\top$.) Thus, $F^{(k)} = (\delta \circledast F)^{(k)} = \delta^{(k)} \circledast F$ for $k \geq 0$. Also, $[H]' = \delta = \delta' \circledast H$, so $u \circledast \delta' = 0$ where $u(t) = 1$.

5.4 The Integral Fourier Transform for Linear Functionals

Let $f_1, f_2, \ldots$ be a sequence of $C_{\downarrow}^{\infty}(\mathcal{R})$-functions such that the regular linear functionals $[f_1], [f_2], \ldots$ functionally converge to the linear functional $F \in L^2(\mathcal{R})^\top$. Now we may take the integral Fourier transform of the linear functional, F, to be that linear functional that is the limit of the functionally-convergent sequence of regular linear functionals $[f_1^\wedge], [f_2^\wedge], \ldots$ *i.e.*, $F^\wedge = \lim_{n\to\infty}[f_n^\wedge]$ where $F = \lim_{n\to\infty}[f_n]$. We have $\lim_{n\to\infty}[f_n^\wedge] := \lim_{n\to\infty}\left[\int_{-\infty}^{\infty} f_n(t)e^{-2\pi i s t}dt\right]$.

Let F be a linear functional in $L^2(\mathcal{R})^\top$ where F is the functional limit of the regular functionals $[f_1], [f_2], \ldots$. Then the linear functional $F^\wedge$ satisfies

$$\begin{aligned} F^\wedge(g) = \int_{\mathcal{R}} F^\wedge g &= \int_{-\infty}^{\infty} \lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(t)e^{-2\pi i s t}dt\, g(s)\, ds \\ &= \lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(t) \int_{-\infty}^{\infty} g(s)e^{-2\pi i s t}ds\, dt \\ &= \lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(t) g^\wedge(t)dt \\ &= F(g^\wedge). \end{aligned}$$

Thus we may *define* the Fourier transform of a linear functional $F \in L^2(\mathcal{R})^\top$ to be the linear functional $F^\wedge \in L^2(\mathcal{R})^\top$ such that $\boxed{F^\wedge(g) = F(g^\wedge)}$ for all functions $g \in C_{\downarrow}^{\infty}(\mathcal{R})$. This definition is consistent with the definition $[f]^\wedge := [f^\wedge]$ for $f \in L^2(\mathcal{R})$ and extends the Fourier transform to non-regular linear functionals.

The identity $\langle f^\wedge, g \rangle = \langle f, g^\wedge \rangle$ for all $f, g \in L^2(\mathcal{R})$ is equivalent to the identity $F^\wedge(g) = F(g^\wedge)$ in the case where $F = [f]_L$ is a regular linear functional based on the function $f \in L^2(\mathcal{R})$, and we extend our notation to allow $\langle F^\wedge, g \rangle$ to represent $F^\wedge(g)$ and $\langle F, g^\wedge \rangle$ to represent $F(g^\wedge)$ in all cases.

Now, using the defining relation $F^\wedge(g) = F(g^\wedge)$ for all $g \in C_{\downarrow}^{\infty}(\mathcal{R})$, we may re-establish all the basic identities, although we have to work a bit harder to do so. For example, the identity $(F^\wedge)'(s) = [-2\pi i t \cdot F(t)]^\wedge(s)$ holds, but now it must be proven as follows. (Remember that we may abuse notation to assign the linear functional F a parameter t, and write $F(t)$ instead of F, allowing us to manipulate expressions involving the application of F in a more "suggestive" manner.)

In order to show that $\boxed{(F^\wedge)'(s) = [-2\pi it \cdot F(t)]^\wedge(s)}$ holds, we need to show that $\langle (F^\wedge)'(s), g(s)\rangle = \langle [-2\pi it \cdot F(t)]^\wedge(s), g(s)\rangle$ for every $g \in C^\infty_\downarrow(\mathcal{R})$. But, we have $\langle (F^\wedge)', g\rangle = \langle F^\wedge, -g'\rangle = \langle F, -(g')^\wedge\rangle = \langle F^\wedge, -\int_{-\infty}^{\infty} g'(t)e^{-2\pi ist}dt\rangle$, and, integrating by parts, we have

$$\begin{aligned}
\langle (F^\wedge)', g\rangle &= \langle F, -\left(\left(g(t)e^{-2\pi ist}\Big|_{t=-\infty}^{t=\infty}\right) - \int_{-\infty}^{\infty} g(t)D_t^1[e^{-2\pi ist}]dt\right)\rangle \\
&= \langle F, \int_{-\infty}^{\infty} g(t)(-2\pi is)e^{-2\pi ist}dt\rangle \\
&= \langle F, (-2\pi is)\int_{-\infty}^{\infty} g(t)e^{-2\pi ist}dt\rangle \\
&= \langle F, -2\pi is \cdot g^\wedge(s)\rangle \\
&= \langle -2\pi is \cdot F(s), g^\wedge(s)\rangle \\
&= \langle (-2\pi is \cdot F(s))^\wedge, g(s)\rangle \\
&= \langle (-2\pi it \cdot F(t))^\wedge, g(t)\rangle.
\end{aligned}$$

(D_t^k denotes k-fold differentiation with respect to t.)

Thus, $(F^\wedge)' = [-2\pi it \cdot F(t)]^\wedge$.

Note, we have $\langle F, -(g')^\wedge\rangle = \langle F(s), -2\pi is \cdot g^\wedge(s)\rangle$. (Does this imply the previously-introduced identity $(g')^\wedge(s) = 2\pi is \cdot g^\wedge(s)$?)

The following identity provides the Fourier transform of the k-th derivative of the δ functional: $\boxed{(\delta_r^{(k)})^\wedge(s) = (2\pi is)^k e^{-2\pi isr}}$.

Let $g \in L^2(\mathcal{R})$. Then

$$\begin{aligned}
\langle (\delta_r^{(k)})^\wedge, g\rangle &= \langle \delta_r^{(k)}, g^\wedge\rangle \\
&= (\delta_r^{(k)})(g^\wedge) \\
&= (-1)^k (g^\wedge)^{(k)}(r) \\
&= (-1)^k \left[D_s^k\left(\int_{-\infty}^{\infty} g(t)e^{-2\pi ist}dt\right)\right]_{s=r} \\
&= (-1)^k \left[\int_{-\infty}^{\infty} g(t)(-2\pi it)^k e^{-2\pi ist}dt\right]_{s=r} \\
&= \int_{-\infty}^{\infty} (2\pi it)^k e^{-2\pi irt} g(t)dt \\
&= \langle (2\pi it)^k e^{-2\pi irt}, g(t)\rangle \\
&= \langle (2\pi is)^k e^{-2\pi irs}, g(s)\rangle.
\end{aligned}$$

Thus $(\delta_r^{(k)})^\wedge(s) = (2\pi is)^k e^{-2\pi isr}$.

Note for $k = 0$ and $r = 0$, we have $\delta_0^\wedge(s) = \delta^\wedge = 1$. This is consistent with the direct heuristic formula $\delta_r^\wedge(s) = \int_{-\infty}^{\infty} \delta_r(t)e^{-2\pi ist}dt = \langle \delta_r(t), e^{-2\pi ist}\rangle = (\delta_r(t))(e^{-2\pi ist}) = e^{-2\pi isr}$. Also, $1^\wedge = \delta^{\wedge\wedge} = \delta^R = \delta$.

Exercise 5.20: Show that, for $F \in L^2(\mathcal{R})^\top$, we have $(F')^\wedge(s) = 2\pi is \cdot F^\wedge(s)$.

Exercise 5.21: Let $y_r(s) = e^{-2\pi irs}$. Show that $y_r^\vee(t) = \delta_r(t)$. In particular, $y_0^\vee(t) = \delta_0$.

Solution 5.21: Let $u(t) = 1$ and recall that $u^\wedge(s) = \delta(s)$. Then

$$\begin{aligned} y_r^\vee(t) &= \int_{-\infty}^{\infty} e^{-2\pi irs} e^{-2\pi ist} ds \\ &= \int_{-\infty}^{\infty} 1 \cdot e^{-2\pi is(r-t)} ds \\ &= u^\wedge(r-t) \\ &= \delta(r-t) \\ &= \delta(t-r) \\ &= \delta_r(t). \end{aligned}$$

Exercise 5.22: Show that $[-2\pi it]^\wedge = \delta_0{}' = -\delta_0 D_t^1$.

Also,

$$\boxed{(F(t-r))^\wedge(s) = e^{-2\pi isr} F^\wedge(s)} \text{ and}$$
$$\boxed{(e^{-2\pi itr} F(t))^\wedge(s) = F^\wedge(s+r)} \text{ and}$$
$$\boxed{F(at))^\wedge(s) = \tfrac{1}{|a|} F^\wedge(\tfrac{s}{a})} \text{ for } F \in L^2(\mathcal{R})^\top.$$

Exercise 5.23: Let $x(t) = e^{2\pi irt}$. Show that $x^\wedge(s) = \delta(s-r)$.

Exercise 5.24: Show that $\delta^\wedge(g(s)) = 1$ for $g \in L^2(\mathcal{R})$, and then show that $(\delta(at))^\wedge(s) = \dfrac{1}{|a|}\delta(s)$. Hint: use $\delta(at) = \dfrac{1}{|a|}\delta(t)$.

5.5 Periodic Linear Functionals

Regular linear functionals based on admissible periodic functions in $NL^2(\mathcal{R})$ are members of $L^2(\mathcal{R})^\top$, and hence the Fourier transform on linear functionals subsumes the Fourier transform for periodic functions.

Exercise 5.25: Let $x_f(t) = cos(-2\pi ift)$ and let $u(t) = 1$. Show that $[x_f]^\wedge(s) = \frac{1}{2}(\delta_f + \delta_{-f})$.

Solution 5.25:

$$\begin{aligned} [x_f]^\wedge(s) &= \int_{-\infty}^{\infty} cos(-2\pi ift) e^{-2\pi ist} dt \\ &= \int_{-\infty}^{\infty} \frac{1}{2}\left[e^{2\pi ift} + e^{-2\pi ift}\right] e^{-2\pi ist} dt \\ &= \frac{1}{2}\int_{-\infty}^{\infty} \left[e^{-2\pi i(s-f)t} + e^{-2\pi i(s+f)t}\right] dt \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2}\left[\int_{-\infty}^{\infty} u(t)e^{-2\pi i(s-f)t}dt + \int_{-\infty}^{\infty} u(t)e^{-2\pi i(s+f)t}dt\right] \\
&= \frac{1}{2}(u^\wedge(s-f) + u^\wedge(s+f)) \\
&= \frac{1}{2}(\delta(s-f) + \delta(s+f)) \\
&= \frac{1}{2}(\delta_f + \delta_{-f}).
\end{aligned}$$

Note that $x_0(t) = u(t)$ and thus, again we repeat $[u]^\wedge(s) = \delta_0$.

Let $[x_n(t)]$ be the regular linear functional corresponding to the function $x_n(t) = \sum_{-n\le h\le n} C_h e^{2\pi i(h/p)t}$ where $C_h \in \mathcal{C}$. Then $[x_n(t)]^\wedge(s) = \sum_{-n\le h\le n} C_h \left[e^{2\pi i(h/p)t}\right]^\wedge(s) = \sum_{-n\le h\le n} C_h\delta(s-h/p)$.

Also when, for some fixed real value α and for some fixed integer k, we have $|C_h| \le \alpha|h|^k$ for all $h \in \mathcal{Z}$, then the sequence of regular linear functionals $[x_0], [x_1], \ldots$ functionally-converges to a linear functional X. If the integer $k \le -1$, then X is the regular linear functional $[x]$ where the function $x(t)$ is the period-p function given by the Fourier series $\sum_{-\infty<h<\infty} C_h e^{2\pi i(h/p)t}$.

If the integer $k \ge 0$, then X is the non-regular linear functional $\left(\sum_{-\infty<h<\infty} C_h\delta(s-h/p)\right)^\vee(t)$ corresponding to the *divergent* series $\sum_{-\infty<h<\infty} C_h e^{2\pi i(h/p)t}$ (!) In either case, X is called a periodic period-p linear functional, since $X(t)(f) = X(t+jp)(f)$ for $j \in \mathcal{Z}$. The linear functional $X^\wedge(s)$ has a "complex area"-C_h "spike" at $s = h/p$ whenever $C_h \ne 0$; and hence $X^\wedge$ is descriptively called a *Dirac comb* linear functional.

Note that, although the sequence $x_1(t), x_2(t), \ldots$, where $x_n(t) = \sum_{-n\le h\le n} C_h e^{2\pi i(h/p)t}$, corresponding to the linear functional X, may be divergent in the $L^2(\mathcal{R})$-norm, it is *convergent* to the linear functional X in the sense that $|[x_n](f) - X(f)| \to 0$ as $n \to \infty$ for $f \in C^\infty_\downarrow(\mathcal{R})$. This is because the integral $[x_n](f)$ is "tempered" by the rapid decay of the test functions in $C^\infty_\downarrow(\mathcal{R})$. Thus, although $\sum_{-\infty<h<\infty} C_h e^{2\pi i(h/p)t}$ may not exist as a function of t, the linear functional that we might digressively denote by $\left[\sum_{-\infty<h<\infty} C_h e^{2\pi i(h/p)t}\right]$ is well-defined when $|C_h| = O(|h|^k)$ for some fixed integer k.

Whether X is regular or not, we can recover the coefficients C_h via the following formula.

$$C_h = \frac{1}{p}\lim_{n\to\infty}\int_{-\infty}^{\infty} U(t/p)x_n(t)e^{-2\pi i(h/p)t}dt,$$

or more succinctly,

$$C_h = \frac{1}{p}\int_{-\infty}^{\infty} U(t/p)X(t)e^{-2\pi i(h/p)t}dt,$$

where $U(t)$ is a so-called *unitary* function; a unitary function U is an infinitely-smooth even bi-constant function in $C^\infty_\downarrow(\mathcal{R})$ with $U(t) = 0$ for $|t| \geq 1$ and $U(t) + U(t-1) = 1$ for $t \in [0,1]$.

Recall that such a function satisfies $U^\wedge(s) = 0$ for $s \in \mathcal{Z} - \{0\}$, and $U^\wedge(0) = 1$.

Then

$$\begin{aligned}
\frac{1}{p}\lim_{n\to\infty}\int_{-\infty}^{\infty} U(t/p)x_n(t)e^{-2\pi i(h/p)t}dt &= \frac{1}{p}\lim_{n\to\infty}\int_{-\infty}^{\infty} U(t/p)\sum_{-n\leq k\leq n} C_k e^{2\pi i(k/p)t}e^{-2\pi i(h/p)t}dt \\
&= \lim_{n\to\infty}\int_{-\infty}^{\infty} U(t/p)\sum_{-n\leq k\leq n} C_k e^{-2\pi i((h-k)/p)t}\frac{1}{p}dt \\
&= \lim_{n\to\infty}\sum_{-n\leq k\leq n} C_k\int_{-\infty}^{\infty} U(t/p)e^{-2\pi i((h-k)/p)t}\frac{1}{p}dt \\
&= \lim_{n\to\infty}\sum_{-n\leq k\leq n} C_k\int_{-\infty}^{\infty} U(r)e^{-2\pi i(h-k)r}dr \\
&= \lim_{n\to\infty}\sum_{-n\leq k\leq n} C_k U^\wedge(h-k) \\
&= \lim_{n\to\infty} C_h U^\wedge(0) \\
&= C_h.
\end{aligned}$$

The following example leads to instances of the Dirac comb linear functional on "both sides" of the $\wedge$ operator.

Let $x_n(t) = \frac{1}{p}\sum_{-n\leq h\leq n} 1\cdot e^{2\pi i(h/p)t}$ and consider the linear functional $X_p = \lim_{n\to\infty}[x_n]$.

Note $x_n(t) = \frac{1}{p} + \frac{2}{p}\sum_{1\leq h\leq n} cos(2\pi(h/p)t)$.

Also, $x_n^\wedge(s) = \frac{1}{p}\delta(s) + \frac{2}{p}\sum_{1\leq h\leq n}\frac{1}{2}(\delta(s-h/p)+\delta(s+h/p)) = \frac{1}{p}\sum_{-n\leq h\leq n}\delta(s-h/p)$.

Thus, $X_p^\wedge(s) = \frac{1}{p}\sum_{-\infty<h<\infty}\delta(s-h/p)$.

We can "sum" the divergent series $X_p(t) = \frac{1}{p}\sum_{-\infty<h<\infty} 1\cdot e^{2\pi i(h/p)t}$ by using a device presented in [Zem87].

Consider the series $g(t) = \frac{1}{p}\sum_{h\neq 0}(2\pi i(h/p))^{-2}e^{2\pi i(h/p)t}$; each term, $\frac{1}{p}(2\pi i(h/p))^{-2}e^{2\pi i(h/p)t}$, can be differentiated twice to ovtain the term $\frac{1}{p}e^{2\pi i(h/p)t}$. Thus $X_p(t) = \frac{1}{p} + g''(t)$.

But the Fourier series $\frac{1}{p}\sum_{h\neq 0} -(2\pi(h/p))^{-2}e^{2\pi i(h/p)t}$ is the Fourier series of the period-p periodic function $f_{[p]}(t)$, where $f(t) = \frac{p}{4\pi^2}\left[\frac{\pi^2}{6} - \frac{1}{2}\left(\frac{2\pi}{p}t - \pi\right)^2\right]$ on $[0, p)$.

Thus, $g(t) = f_{[p]}(t)$, and $g'(t) = (f_{[p]})'(t) = \frac{1}{2} - \frac{1}{p}(t \bmod p)$, and $g''(t) = -\frac{1}{p} + \sum_h \delta(t - hp)$.

Therefore, $X_p(t) = \frac{1}{p} + g''(t) = \sum_h \delta(t - hp)$.

Also, we can verify that $X_p(t) = \sum_h \delta(s - hp) = \sum_h \delta_{hp}$, since the h-th Fourier coefficient of X_p is

$$\begin{aligned}
\frac{1}{p}\int_{-\infty}^{\infty} U(t/p)X_p(t)e^{-2\pi i(h/p)t}dt &= \frac{1}{p}\int_{-\infty}^{\infty} U(t/p)\frac{1}{p}\sum_{-\infty\leq h\leq\infty} \delta_{hp}\, e^{-2\pi i(h/p)t}dt \\
&= \frac{1}{p}\sum_{-\infty<h<\infty} \delta_{hp}\int_{-\infty}^{\infty} U(t/p)e^{-2\pi i(h/p)t}\frac{1}{p}dt \\
&= \frac{1}{p}\sum_{-\infty<h<\infty} \delta_{hp}\int_{-\infty}^{\infty} U(r)e^{-2\pi ihr}dr \\
&= \frac{1}{p}\sum_h \delta_{hp}U^\wedge(h) \\
&= \frac{1}{p}\sum_{0\leq h\leq 0} \delta_{0p}U^\wedge(0) \\
&= \frac{1}{p}U^\wedge(0) \\
&= \frac{1}{p}.
\end{aligned}$$

Thus,

$$X_p^\wedge(s) = \left(\sum_{-\infty<h<\infty} \delta(t - hp)\right)^\wedge(s) = \frac{1}{p}\sum_{-\infty<h<\infty} \delta(s - h/p) = \frac{1}{p}X_{1/p}(s).$$

Exercise 5.26: How is the above identity $X_p^\wedge(s) = \frac{1}{p}X_{1/p}(s)$ related to the Poisson summation formula?

Solution 5.26: We have $\langle X_p^R, f\rangle = \langle X_p^\wedge, f^\wedge\rangle$, and since X_p is even, $X_p^R = X_p$, and thus, $\langle X_p, f\rangle = \langle X_p^\wedge, f^\wedge\rangle$, and this is equivalent to $\sum_h f(hp) = \frac{1}{p}\sum_h f^\wedge(\frac{h}{p})$; this is just the Poisson summation formula.

Exercise 5.27: Let $y \in L^2(\mathcal{R})$ such that the support set of y is contained in $[0, p)$ where $p \in \mathcal{R}^+$. Show that the period-p periodic extension of y, $y_{[p]}$, is $y(t) \circledast (\sum_h \delta(t - hp))$. Also, compute $\left[y_{[p]}\right]^\wedge$.

5.6 Some Fourier Integral Transforms

Recall $H(t) = \begin{cases} 0, & \text{if } t < 0, \\ \frac{1}{2}, & \text{if } t = 0, \\ 1, & \text{if } t > 0, \end{cases}$
and let $\bar{B}_a(t) = H(t + \frac{a}{2}) - H(t - \frac{a}{2})$. Note $sign(t) = 2H(t) - 1 = H(t) - H(-t)$.

Let $sinc_p(s) = \begin{cases} 1, & \text{if } s = 0, \\ sin(\pi ps)/(\pi ps), & \text{otherwise.} \end{cases}$

Below we use the linear functional δ_v on $L^2(\mathcal{R})$ defined such that $\langle \delta_v, f \rangle = \int_{-\infty}^{\infty} \delta_v(t) f(t)\, dt = \int_{-\infty}^{\infty} \delta(t - v) f(t)\, dt := f(v)$. The operator δ_v is often imagined as a unit-area "spike" at v. Note $\delta_v = [H]'(v)$. (We do not write $[H'(v)]$ because H' is not a function.) $[H]'(v)$ is not a regular linear functional. Also, in prefix operator terms, we have $[H]''(v) = \delta_v' = \delta_v(-1)D_a^1$ where a denotes the argument of whatever function the differentiation operator D_a^1 is applied to.

- $x(t) = \begin{cases} be^{-at}, & \text{if } t > 0, \\ \frac{b}{2}, & \text{if } t = 0, \\ 0, & \text{if } t < 0, \end{cases}$ $x^\wedge(s) = b/[a + i2\pi s]$.
- $x(t) = e^{-at^2}$ with $a \in \mathcal{R}$ and $a > 0$, $x^\wedge(s) = \sqrt{\dfrac{\pi}{a}} e^{-\pi^2 s^2/a}$.
- $x(t) = e^{-a|t|}$ with $a \in \mathcal{R}$ and $a > 0$, $x^\wedge(s) = \dfrac{2a}{a^2 + (2\pi s)^2}$.
- $x(t) = \dfrac{1}{2\pi} \dfrac{a}{(t - b)^2 + (a/2)^2}$, $x^\wedge(s) = e^{-2\pi sb - a\pi|s|}$.
- $x(t) = 1$, $x^\wedge(s) = \delta_0$.
- $x(t) = \delta_0$, $x^\wedge(s) = 1$.
- $x(t) = cos(2\pi at)$, $x^\wedge(s) = \frac{1}{2}(\delta_a + \delta_{-a})$.
- $x(t) = sin(2\pi at)$, $x^\wedge(s) = \frac{i}{2}(\delta_{-a} - \delta_a)$.
- $x(t) = \log(|t|)$, $x^\wedge(s) = -1/|2s|$.
- $x(t) = t^k$ with $k \in \mathcal{Z}^+$, $x^\wedge(s) = 2\pi i^k \delta^{(k)}(2\pi s)$.
- $x(t) = H(t)$, $x^\wedge(s) = \frac{1}{2}\left(\delta_0 + \frac{1}{\pi is}\right)$.
- $x(t) = sign(t)$, $x^\wedge(s) = \frac{1}{\pi is}$.

  Exercise 5.28: Show that $(sign(at))^\wedge(s) = \dfrac{sign(a)}{\pi is}$, and that $(H(at))^\wedge(s) = \dfrac{1}{2}\left(\delta_0 + \dfrac{sign(a)}{\pi is}\right)$.

- $x(t) = \bar{B}_a(t)$, $x^\wedge(s) = a\ sinc_a(s)$. Note $\bar{B}_a$ is a bi-constant function, and $x^\wedge(k/a) = 0$ for $k \in \mathcal{Z} - \{0\}$.

- $x(t) = sinc_a(t)$ with $a \in \mathcal{R}$ and $a > 0$, $x^\wedge(s) = \dfrac{1}{a}\bar{B}_a(s)$.
- $(\int_{-\infty}^{t} x(r)\,dr)^\wedge(s) = x^\wedge(s)/(2\pi is) + \dfrac{1}{2}x^\wedge(0)\delta(s)$.
- $\left(\int_0^t x(z)\,dz\right)^\wedge(s) = x^\wedge(s)/(2\pi is) + \left(\dfrac{1}{2}x^\wedge(0) - \int_{-\infty}^{0} x(r)dr\right)\delta(s)$
- $x(t) = tH(t) = (H \circledast H)(t)$, $x^\wedge(s) = \pi i\delta'_{2\pi s} - i/(4\pi^2 s^2)$.
- $x(t) = 1/t$, $x^\wedge(s) = -i\pi \cdot sign(s)$.

6 Finite Record of a Function

Suppose we are given x on an interval of length p, say $[a, p+a]$, and we wish to know $x^\wedge$. Consider $y(t) = x(t) \cdot b(t)$, where $b(t)$ is chosen to be 0 outside $[a, p+a]$. Now $y^\wedge = (xb)^\wedge = x^\wedge \circledast b^\wedge$, so that $y^\wedge$ is just $x^\wedge$ convolved with $b^\wedge$. But $b^\wedge$ is, in principle, known so that we can characterize the effect of convolving with $b^\wedge$ and thereby observe the nature of $x^\wedge$ as seen in $y^\wedge$. The function $y^\wedge$ is an estimate of $x^\wedge$, and differing choices of the "windowing" function, b, will have differing effects on the nature of the estimate $y^\wedge$. (Note if x is known on all of $\mathcal{R}$, then $b(t) = 1$ is appropriate, and this is consistent with $y^\wedge = x^\wedge \circledast \delta_0 = x^\wedge$.)

No matter what choice of b is made, in any case where x is, in fact, not zero outside $[a, p+a]$, the estimate $y^\wedge$ will differ from $x^\wedge$. This error $y^\wedge - x^\wedge$ is said to be due to *leakage*. Values of $y^\wedge(s)$ are really composed of the corresponding values of $x^\wedge(s)$ plus neighboring values of $x^\wedge$ as obtained in the convolution $x^\wedge \circledast b^\wedge$. These neighboring values are said to "leak" into the estimated value for $x^\wedge$.

When $b(t) = 1$ if $a \le t \le p+a$, and 0 otherwise, then $b^\wedge(s) = e^{-2\pi is(a+p/2)} \cdot sin(\pi ps)/(\pi s)$.

Other reasonable choices for b are: $b(t) = \ldots$.

7 Spectral Power Density Function

Let $x(t)$ be the voltage across a one Ohm resistance at time t. By Ohm's law, the current through the resistance at time t is $x(t)/1$ Amperes. Thus, the power being used at time t to heat the resistance is $x(t) \cdot x(t)/1$ Joules/second. Often $x(t) = 0$ for $t < 0$. In any event, $\int_{-\infty}^{\infty} |x(t)|^2\,dt$ Joules is the total amount of energy converted into heat.

Now suppose $x(t)$ belongs to $L^2(\mathcal{R})$ so that $x^\wedge$ exists. By Plancherel's identity, $\|x\|^2 = \|x^\wedge\|^2$, so the energy converted is

$$\|x\|^2 = \|x^\wedge\|^2 = \int_{-\infty}^{\infty} |x^\wedge(s)|^2\,ds.$$

The function $|x^\wedge(s)|^2$ is called the spectral power density function of x. It should, of course, be called the spectral *energy* density, but the word "power" is used in order to correspond to the case where x is periodic. The energy converted due to the complex spectral components of x in the frequency band $[a,b]$ is $\int_a^b |x^\wedge(s)|^2\,ds$.

When x is real-valued, $x^\wedge$ is Hermitian, so $|x^\wedge|^2 = (x \circledast x^R)^\wedge = (x \otimes x)^\wedge$, and the power density function is even. Thus, folding into the positive frequencies results in the energy due to the real spectral components in the positive frequency band $[a,b]$, with $0 \le a \le b$, being $\int_a^b (2 - \delta_{s,0})|x^\wedge(s)|^2\,ds$. Recall that $M(s)$ is the amplitude spectrum function of x, and note $|x^\wedge(s)|^2 = x^\wedge(s)x^\wedge(s)^* = M(s)^2/(2-\delta_{s,0})^2$ when x is real, so $(2-\delta_{s,0})|x^\wedge(s)|^2 = M(s)^2$. The factor $(2-\delta_{s,0})$ in the integrand, which differs from 2 at just one point, can be replaced by 2. The finicky $-\delta_{s,0}$, is not necessary, but it is harmless, and it reminds us of the underlying manipulations which have been performed.

8 Time-Series, Correlation, and Spectral Analysis

Let $x(t)$ be a real (complex-valued extension - ?) stochastic process. Then the *autocovariance* function of x is defined as

$$C_{xx}(r,t) := \mathrm{cov}(x(r), x(t)).$$

When the mean function $m_x(t) := E(x(t))$ and the variance function $v_x(t) := \mathrm{Var}(x(t)) = C_{xx}(t,t)$ are both constant, with $E(x(t)) = \mu_x$ and $\mathrm{Var}(x(t)) = \sigma_x^2$, then x is called *weakly-stationary*, and the autocovariance function $C_{xx}(r,t)$ is, in fact, just a function of the lag $h = r - t$, and we have $C_{xx}(h) = E(x(t+h)x(t)) - \mu_x^2$.

When x is weakly-stationary, C_{xx} is continuous, with $C_{xx}(0) = \sigma_x^2$, and when x is real, C_{xx} is real and even and $|C_{xx}(h)|$ is a decreasing function of h, so $C_{xx}(0) = \max_h C_{xx}(h)$.

The *autocorrelation kernel* function of x is just the non-central second moment $D_{xx}(r,t) = E(x(r) \cdot x(t))$, and if x is weakly-stationary, then the autocorrelation kernel function depends only on the lag $h = r - t$ and we write $D_{xx}(h) = E(x(t+h) \cdot x(t))$.

The autocorrelation kernel function is *not* the same as the correlation coefficient $\rho_x(h) := \mathrm{cor}(x(t+h),\, x(t)) = C_{xx}(h)/(C_{xx}(0))$, but it is linearly-related by $(D_{xx}(h) - \mu_x^2)/C_{xx}(0) = \rho_x(h)$, and it is often easier to work with D_{xx} than with either $\rho_x(h)$ or $C_{xx}(h)$.

It may be that $C_{xx}(h) = \mathrm{cov}(x(t+h),\, x(t))$ can, with probability 1, be computed as an average over time of a time sample $\tilde{x}$ of the weakly-stationary process x, so that

$$C_{xx}(h) = \lim_{p\to\infty} (1/(2p)) \int_{-p}^{p} (\tilde{x}(t+h) - \mu_x)(\tilde{x}(t) - \mu_x)\,dt,$$

and also

$$\mu_x = \lim_{p\to\infty} (1/(2p)) \int_{-p}^{p} \tilde{x}(t)\,dt.$$

If, in general,

$$E(g(x(f_1(t)),\ldots,x(f_n(t)))) = \lim_{p\to\infty}(1/(2p))\int_{-p}^{p} g(x(f_1(t)),\ldots,x(f_n(t)))$$

then the process x is called an *ergodic* process. Note an ergodic process is weakly-stationary.

When x is ergodic, the auto-correlation kernel function $D_{xx}(h) = (x \otimes x)(h)$. When x is ergodic and $\mu_x = 0$, then $\|x\|^2 = C_{xx}(0) = \sigma_x^2$, and thus the variance of x is the total "power" of x, and the *component* of the variance of the form $\int_a^b |x^\wedge(s)|^2\,ds$ is the part of the variance due to the "power" of x in the frequency band $[a,b]$. More generally, $C_{xx}(h) = (x \otimes x)(h)$ and $\int C_{xx} = (\int x)^2$, and $C^\wedge_{xx}(s) = |x^\wedge(s)|^2$. The transform $C^\wedge_{xx}(s)$ is the spectral power density function of the process x.

[Define cross-cov and cor, and their transforms.]

9 Linear Systems

A one-input linear system is an operator, H, which maps an *input* function, x, to a corresponding *output* function, y. Thus $y(t) = (Hx)(t)$. H is a linear operator, so that $H(ax + bz) = a(Hx) + b(Hz)$.

A *shift-invariant* linear system has the property that $H(x(t-s)) = (Hx)(t-s)$. A shift-invariant linear system, H, must be frequency-preserving in the sense that, if x is a period-p periodic function, then $y = Hx$ is also a period-p periodic function. In particular, $H(a \cdot cos(st + b)) = c \cdot cos(st + d)$ for some values c and d. The admissible input functions, x, are just those complex-valued functions which possess a Fourier-Stieljes transform.

An example of primary importance is that where H is defined via a k-th order ordinary differential equation form with constant coefficients; *i.e.*, $H = \alpha_k D^k + \alpha_{k-1} D^{k-1} + \ldots + \alpha_1 D + \alpha_0$, where D denotes the differentiation operator (with respect to the argument t of the input function it is applied to.) In this situation, a solution function x of the non-homogeneous k-th order ordinary differential equation with constant coefficients: $(\alpha_k D^k + \alpha_{k-1} D^{k-1} + \ldots + \alpha_1 D + \alpha_0)x = y$, is the input function to the linear system H corresponding to the output function y, *i.e.*, $Hx = y$ where $H = \alpha_k D^k + \alpha_{k-1} D^{k-1} + \ldots + \alpha_1 D + \alpha_0$.

Every shift-invariant linear system operator H has an associated complex-valued function h, called the *system-weighting* function of H, such that $(Hx) = x \circledast h$. Often h is called the *impulse-response* function of H, since $h = \delta_0 \circledast h$, where δ_0 is the Dirac-δ functional with its spike at 0. Let $Hx = y$. Note that saying $y = x \circledast h$ shows that each value, $y(t)$, is a certain weighted "sum" of the values of x, where the value $x(r)$ is weighted by $h(t-r)$.

In order that $y(t)$ depend only on the x-values $x(r)$ with $r \le t$, we must have $h(t) = 0$ for $t < 0$. A linear system with such a weighting function is called *physically-realizable*; it corresponds to some real-time processor which can input x and output y in real-time. Such a processor may, of course, involve memory, but not a delay due to "reading ahead".

A linear system operator H is *stable* if Hx is bounded when x is bounded; thus finite input cannot cause the output of a stable system to "blow-up". If H is a stable shift-invariant linear system operator with the impulse-response function h, then the Fourier transform of h exists and the complex-valued function $h^\wedge$ is called the *frequency-response* function of H, and is such that $(Hx)^\wedge = x^\wedge h^\wedge$. Often $h^\wedge$ is called the *transfer* function of H.

We may write $h^\wedge$ in polar form as $h^\wedge(s) = |h^\wedge(s)|e^{i\phi(s)}$, where $\phi(s)$ is the phase-shift function of $h^\wedge$. If h is real then $|h^\wedge(s)| = M(s)/(2-\delta_{s,0})$ where $M(s)$ is the amplitude function of h. In any event, $|h^\wedge(s)|$ is called the *gain* function of H and ϕ is called the *phase-shift* function of H, since if the input $x(t)$ is a complex oscillation $Ae^{i(2\pi st+q)}$, then the output $(Hx)(t)$ is $|h^\wedge(s)|Ae^{i(2\pi st+q+\phi(s))}$, which is just an oscillation of the same frequency, s, whose amplitude is multiplied by the gain $|h^\wedge(s)|$ and whose phase is shifted by the phase-shift value $\phi(s)$. This is just a special case of the relation $(Hx)^\wedge = x^\wedge h^\wedge$.

When the system-weighting function, h, is real, then the frequency-response function $h^\wedge$ is Hermitian, *i.e.*, $h^{\wedge R} = h^{\wedge *}$, and the gain and phase-shift functions are real and $|h^\wedge(s)|$ is even and $\phi(s)$ is odd. In this case H preserves real signals, *i.e.*, Hx is real whenever x is real.

Cascading two stable shift-invariant linear systems H_1 and H_2 results in a linear system H_2H_1 whose output is $(H_2(H_1x))$, and the frequency-response function is $h_1^\wedge h_2^\wedge$, so the gain function is $|h_1^\wedge(s)| \cdot |h_2^\wedge(s)|$ and the phase-shift function is $\phi_1(s) + \phi_2(s)$.

Given the input x and the output y of a stable shift-invariant linear system we may determine the frequency-response function $h^\wedge(s)$ as $y^\wedge(s)/x^\wedge(s)$ at each frequency, s, which appears in x, *i.e.*, for which $x^\wedge(s) \neq 0$. A test input function, x, constructed for the purpose of determining $h^\wedge$ should thus have a broad spectrum. In the same way, given the output function y and the frequency-response function $h^\wedge$ (possibly determined by a prior computation based on known input and output,) we can also obtain the input function as $x = (y^\wedge/h^\wedge)^\vee$; this is often called a deconvolution computation.

If we have the output function y and we know the form of the linear operator H (as a differential operator, for example,) apart from the values of some parameters appearing in H, then we can determine these parameter values, and thus determine H, by curve-fitting data-points of the form $(t, y(t))$ where $y(t) = (Hx)(t)$. (Indeed, there is no requirement that H be a linear operator in this situation.)

Exercise 9.1: Show that, if the signal x is input to the linear system H with the transfer function h, then the spectral power density at frequency s of the output is $|h^\wedge(s)|^2 \cdot |x^\wedge(s)|^2$. Thus the input spectral density $|x^\wedge(s)|^2$ is scaled by $|h^\wedge(s)|^2$.

When x has finite duration, the total energy in the input function x is $\|x\|_{L^2(\mathcal{R})}$ and the total energy in the output function x is $\|h^\wedge x^\wedge\|_{L^2(\mathcal{R})}$. Thus the linear system H can produce output with energy in excess of the input energy (if it is plugged-in,) or, it can produce output with energy less than the input energy, or, it can produce output with energy identical to the input energy,

9.1 Linear System Filtering

Given a noisy signal $x(t) = q(t) + n(t)$, where $q(t)$ is the pure signal and $n(t)$ is the noise, we may wish to *filter* x for various purposes. In general, filtering-out noise in a signal is an averaging or *smoothing* process, and using Fourier transforms to compute and suppress the high-frequency components of the signal is often appropriate.

We suppose that the noise n is ergodic ((?) specify q and n more carefully.) and we restrict our attention to shift-invariant linear filters. A linear filter is a linear transformation F such that the output $y(r) = (Fx)(r)$ is computable as $(x \circledast f)(r)$, where f is the system-weighting function of F.

Given x, we may obtain the auto-correlation kernel transform $d_{xx}(s) = (x \otimes x)^\wedge(s)$, and we suppose that the cross-correlation kernel transform $d_{xq}(s) = (x \otimes q)^\wedge(s)$ is known. Then in terms of $d_{xx}(s)$ and $d_{xq}(s)$, the Wiener filter F is that shift-invariant linear transformation which has the system-weighting function $f(t) = (d_{xq}(s)/d_{xx}(s))^\vee(t)$. It is the unique shift-invariant linear filter which minimizes the mean square error $MSE_{qy} = \int_{-\infty}^{\infty}[q(t) - (Fx)(t)]^2\,dt$ between the desired noise-free output q and the realized filtered output, $y(t) = (Fx)(t)$. If the signal q and the noise n are uncorrelated, then

$$f^\wedge(s) = \begin{cases} 0 & \text{if } d_{qq}(s) + d_{nn}(s) = 0, \\ \dfrac{d_{qq}(s)}{d_{qq}(s) + d_{nn}(s)} & \text{otherwise;} \end{cases}$$

and MSE_{qy} reduces to $\int_{-\infty}^{\infty} d_{qq}(s)d_{nn}(s)/(d_{qq}(s) + d_{nn}(s))\,ds$.

If q and n have their respective power concentrated in largely non-overlapping frequency bands, MSE_{qy} is small, since $d_{qq}(s)d_{nn}(s)$ is small. Often q's power is concentrated in a band of relatively low-frequency values, while n is white and has power at high-frequencies also. Then the Wiener filter becomes a low-pass or band-pass filter of optimal shape. Even when d_{xx} and d_{xq}, (or d_{qq} and d_{nn}), are not known, some reasonable estimates may often be employed to advantage[PTVF92].

When the noise in x is, in fact, the result of applying a shift-invariant linear filter H with the system-weighting function h to the signal q, then $x = q \circledast h$, and recovering q from x entails a deconvolution process. In this case, the Wiener filter system-weighting function $(d_{xq}/d_{xx})^\vee = (q^\wedge/x^\wedge)^\vee$ is just the perfect deconvolution system-weighting function $(1/h^\wedge)^\vee$. [Suppose f is to be chosen such that $x \circledast f = q$. Then when $x = q \circledast h$, we have $x^\wedge = q^\wedge h^\wedge$, so $x^\wedge/h^\wedge = q^\wedge$, and hence $x \circledast (1/h^\wedge)^\vee = q$. Therefore f should be chosen as $(1/h^\wedge)^\vee$.] Thus, when the noise n in x is due to passing the signal q through a shift-invariant linear system, then the Wiener filter can recover q exactly.

In the situation where we have the noisy signal x where x is the result of cascading two operators, H_1 and H_2, when we suppose we know the form of the linear operator H_1 apart from the values of some parameters appearing in H_1, and we assume that H_2 is a "reasonable" noise-injecting operator, (for example, H_2 may "encode" the process of measuring a result of H_1,) then curve-fitting data-points of the form $(t, x(t))$ with appropriate weights can both accomodate the noise injected by H_2 and estimate the unknown parameters appearing in H_1.

[Also, consider the spectra of: $x+\text{noise}$, $x\cdot\text{noise}$, x^R. Use $x\cdot\text{noise} = x\cdot y = x(1+(y-1)) = x(1+\varepsilon)$ where $E(y) = 1$ and $E(\varepsilon) = 0$.]

10 Prediction and Control

.... (to add)

References

[BP94] D. Bini and V. Y. Pan. *Polynomial and Matrix Computation, Vol. 1: Fundamental Algorithms.* Birkhauser, Boston, 1994.

[DM72] H. Dym and H. P. McKean. *Fourier Series and Integrals.* Academic Press, New York, 1972.

[Edw67] R. E. Edwards. *Fourier Series - a Modern Introduction, Vol. 1.* Holt, Reinhart and Winston, 1967.

[MP72] J. H. McClellan and T. W. Parks. Eigenvalue and eigenvector decomposition of the discrete fourier transform. *IEEE Trans. on Audio and Electroacoustics*, AU-20:66–74, March 1972.

[PTVF92] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in FORTRAN - The Art of Scientific Computing (2nd. ed.).* Cambridge Univ. Press, N.Y., 1992.

[Rud98] Walter Rudin. *Functional Analysis, 2nd ed.* McGraw-Hill, N.J., 1998.

[Sti86] S.M. Stigler. *The History of Statistics.* Belknap Press, Cambridge MA, 1986.

[Zem87] A. H. Zemanian. *Distribution Theory and Transform Analysis.* Dover Pub., N.Y., 1987.