Fourier Analysis and Signal Processing

John T. Whelan

School of Gravitational Waves, Warsaw, 2013 July 4

1 Continuous Fourier Transforms

You're probably familiar with the continuous Fourier transform

$$\widetilde{x}(f) = \int_{-\infty}^{\infty} dt\, x(t)\, e^{-i2\pi f(t-t_0)} \tag{1.1}$$

and its inverse

$$x(t) = \int_{-\infty}^{\infty} df\, \widetilde{x}(f)\, e^{i2\pi f(t-t_0)} \tag{1.2}$$

Notes:

- Lots of conventions, but note using f instead of ω gets rid of annoying 2π normalizations.
- If $x(t)$ is really a function of time, the origin/epoch t_0 is arbitrary and has no physical meaning. If it's a function of time *difference*, then $t_0 = 0$ makes sense.

The identity

$$\int_{-\infty}^{\infty} df\, e^{i2\pi f(t-t')} = \delta(t-t') \tag{1.3}$$

is useful for proving properties of continuous Fourier transforms.

2 Discrete Fourier Transforms

Real data is neither continuous nor infinite in duration. Consider discretely-sampled time series data of duration $T = N\delta t$:

$$x_j = x(t_j) = x(t + j\delta t) \qquad j = 0, 1, \ldots, N-1 \tag{2.1}$$

Its discrete Fourier transform is

$$\widehat{x}_k = \sum_{j=0}^{N-1} x_j\, e^{-i2\pi f_k(t_j-t_0)} = \sum_{j=0}^{N-1} x_j\, e^{-i2\pi jk/N} \tag{2.2}$$

where $f_k = k\delta f$, and

$$\delta f\, \delta t = \frac{\delta t}{T} = \frac{1}{N}\ . \tag{2.3}$$

We can define $\widehat{x}_k$ for any integer k, but there are only N independent values, thanks to the identifications

$$\widehat{x}_{N+k} = \widehat{x}_k \qquad \text{always} \tag{2.4a}$$

$$\widehat{x}_{-k} = \widehat{x}_k^* \qquad \text{if } \{x_j\} \text{ real} \tag{2.4b}$$

This means, for a real time series $\{x_j\}$, the N real numbers in the Fourier domain are (assuming N even)

- 1 real value x_0
- $\frac{N}{2} - 2$ complex values $\{x_k | k = 1, \ldots \frac{N}{2} - 1\}$
- 1 real value $x_{-N/2} = x_{N/2}$

The identity

$$\sum_{k=0}^{N-1} e^{i2\pi(j-\ell)k/N} = N\, \delta_{j,\ell \bmod N} \tag{2.5}$$

shows us the inverse transform

$$x_j = \frac{1}{N} \sum_{k=0}^{N-1} \widehat{x}_k\, e^{i2\pi jk/N} = \frac{1}{N} \sum_{k=-N/2}^{N/2-1} \widehat{x}_k\, e^{i2\pi jk/N} \tag{2.6}$$

If we consider (2.2) to be an approximation of the integral in (1.1), we'd identify

$$\delta t\, \widehat{x}_k \sim \widetilde{x}(f_k) \tag{2.7}$$

If we plug (1.2) into (2.2) we can get the actual formula

$$\delta t\, \widehat{x}_k = \int_{-\infty}^{\infty} df\, \delta_{N,\delta t}(f_k - f) \widetilde{h}(f) \tag{2.8}$$

with a kernel

$$\delta_{N,\delta t}(x) = \delta t \sum_{j=0}^{N-1} e^{-i2\pi j \delta t\, x} \tag{2.9}$$

this is not quite a Dirac delta function for two reasons:

1. It is periodic with period $\frac{1}{\delta t}$, so it's peaked at $x = 0$, $x = \frac{1}{\delta t}$, $x = -\frac{1}{\delta t}$, etc.
2. It has an oscillatory "ringing" behavior around its peaks.

The second point is related to an issue known as spectral leakage which we won't go into; the first is known as aliasing, and it means that actually $\delta t\, \widehat{x}_k$ is a sum of not only $\widetilde{h}(f_k)$ but also $\widetilde{h}(f_k + \frac{1}{\delta t})$, $\widetilde{h}(f_k - \frac{1}{\delta t}) = \widetilde{h}^*(\frac{1}{\delta t} - f_k)$, etc. This means that to avoid confusion of different frequency components, the original time series $h(t)$ should have undergone some analog processing so that $\widetilde{h}(f)$ is negligible unless

$$-\frac{1}{2\,\delta t} < f < \frac{1}{2\,\delta t} \tag{2.10}$$

which defines the Nyquist frequency $f_{\mathrm{Ny}} = \frac{1}{2\,\delta t}$ which is half the sampling rate $\frac{1}{\delta t}$.

3 Random Data

We'll often be interested in cases where the data $\{x_i\}$ are random with some mean and variance defined by the expectation values

$$E[x_j] = \mu_j \tag{3.1}$$

$$E[x_j x_\ell] = \sigma^2_{j\ell} \tag{3.2}$$

If the data are Gaussian, these are enough to define a probability density

$$P(\mathbf{x}) = (\det 2\pi \boldsymbol{\sigma}^2)^{-1/2} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\sigma}^{-2} (\mathbf{x} - \boldsymbol{\mu}) \right) \tag{3.3}$$

where $\mathbf{x}$ and $\boldsymbol{\mu}$ are column vectors made up out of $\{x_j\}$ and $\{\mu_j\}$, respectively, $\boldsymbol{\sigma}^2$ is a matrix made of $\{\sigma_{j\ell}\}$ and $\boldsymbol{\sigma}^{-2}$ is its inverse. For simplicity we'll assume the data have zero mean. We'll also start in the continuous picture; the random process associated with $x(t)$ is stationary if

$$E\left[x(t), x(t')\right] = K_x(t - t') \tag{3.4}$$

which defines the autocorrelation function $K_x(t-t')$ (in general it would have to be written $K_x(t,t')$). The Fourier transform of the autocorrelation function is the two-sided power spectral density

$$S_x^{\text{2-sided}}(f) = \int_{-\infty}^{\infty} d\tau\, K_x(\tau)\, e^{-i2\pi f\tau} \tag{3.5}$$

We can use (1.3) to show that, formally,

$$E\left[\widetilde{x}(f')^*\, \widetilde{x}(f)\right] = \delta(f-f')\, S_x^{\text{2-sided}}(f) \tag{3.6}$$

Since $S_x^{\text{2-sided}}(f) = S_x^{\text{2-sided}}(-f)$, for real $x(t)$, define one-sided PSD

$$S_x(f) = \begin{cases} S_x^{\text{2-sided}}(0) & f=0 \\ S_x^{\text{2-sided}}(-f) + S_x^{\text{2-sided}}(f) & f>0 \end{cases} \tag{3.7}$$

Unfortunately (?) this is what most GW observers mean by PSD, so formulas have an extra factor of two ($S_x(f) = 2S_x^{\text{2-sided}}(f)$), e.g.,

$$E\left[\widetilde{x}(f)^*\, \widetilde{x}(f)\right] = \delta(f-f')\, \frac{S_x(f)}{2} \tag{3.8}$$

We can translate this into a discrete Fourier transform; just as $\widehat{x}_k \sim \widetilde{x}(f_k)$, we can show

$$E\left[|\widehat{x}_k|^2\right] \sim \frac{N}{2\delta t} S_x(f_k) \tag{3.9}$$

with the usual caveats about leakage and aliasing. Now consider the case of zero-mean Gaussian data: let $\widehat{x}_k = \xi_k + i\eta_k$ and treat ξ_0, $\{\xi_k, \eta_k | k = 1\ldots\frac{N}{2}-1\}$, $\xi_{N/2}$ as independent and Gaussian with

$$E\left[\xi_k^2\right] = E\left[\eta_k^2\right] = \sigma_k^2 = \frac{N}{4\delta t} S_x(f_k) \tag{3.10}$$

so probability density is

$$\begin{aligned} P(\{\xi_k,\eta_k|k=1\ldots\frac{N}{2}-1\}) &= \prod_{k=1}^{N/2-1} \frac{1}{2\pi\sigma_k^2} \exp\left(-\frac{\xi_k^2}{2\sigma_k^2} - \frac{\eta_k^2}{2\sigma_k^2}\right) \\ &\propto \exp(\Lambda) \end{aligned} \tag{3.11}$$

with log-likelihood

$$\Lambda \sim -\sum_{k=1}^{N/2-1} \frac{2\delta t}{N} \frac{|\widehat{x}_k|^2}{S_x(f_k)} \sim -\sum_{k=1}^{N/2-1} 2\delta f \frac{|\widetilde{x}(f_k)|^2}{S_x(f_k)} \sim -2\int_0^{\infty} df\, \frac{|\widetilde{x}(f)|^2}{S_x(f)} \tag{3.12}$$

This means

$$P(x) \propto e^{-\frac{1}{2}\langle x|x\rangle} \tag{3.13}$$

where the inner product is

$$\langle y|z\rangle = 4\,\mathrm{Re} \int_0^{\infty} df\, \frac{\widetilde{y}^*(f)\, \widetilde{z}(f)}{S_x(f)} \tag{3.14}$$

The unfamiliar factor of 4 is one factor of 2 because the integral is only over positive frequencies and one because of the use of the one-sided power spectral density.

If the data vary slowly over the observation time, it may be useful to divide it into pieces of length T and Fourier transform each of them

$$\widetilde{x}_I(f) = \int_{t_{I0}}^{t_{I0}+T} dt\, x(t)\, e^{-i2\pi f(t-t_{I0})} \tag{3.15}$$

In principle, the statistical properties of different segments will be related because of the autocorrelation function $K(t-t')$. But if the correlation length–the time over which $K(\tau)$ is non-negiligible–is small compared to T, we can neglect this, and the log likelihood function would become $P(x) \propto e^{\Lambda(x)}$ with

$$\Lambda = -2\,\mathrm{Re} \sum_I \int_0^{f_{\mathrm{Ny}}} df\, \frac{|\widetilde{x}_I(f)|^2}{S_I(f)} \tag{3.16}$$