# Basics of Signal Detection

Topics covered:

- Fourier Transforms (basic mathematical tool)
- Noise Spectral Density (characterising noise)
- Matched Filters (finding signals)
- Signal-to-Noise Ratio (SNR – signal detectability)

Topics **not** covered:

- Matches
- Template banks
- Parameter estimation
- Probability theory
- Detecting unmodelled signals
- Discretely sampled data (see handout: Ch 12, 13 of "Numerical Recipes")

# Fourier Transforms

"Any" (reasonably behaved) function $x(t)$ can be written as a sum of sines and cosines, or equivalently as exponentials:

$$\cos 2\pi f t = \frac{e^{i2\pi f t} + e^{-i2\pi f t}}{2} \qquad \sin 2\pi f t = \frac{e^{i2\pi f t} - e^{-i2\pi f t}}{2i}$$

Fourier transform definition:

$$x(t) = \int_{-\infty}^{\infty} \underbrace{\tilde{x}(f)}_{\text{"Fourier transform of } x(t)\text{"}} e^{i2\pi f t} \, df$$

where

$$\tilde{x}(f) = \int_{-\infty}^{\infty} x(t)\, e^{-i2\pi f t} \, dt$$

Self-consistency check:

$$x(t) = \int_{f=-\infty}^{\infty} \left[ \int_{t'=-\infty}^{\infty} x(t')\, e^{-i2\pi f t'} dt' \right] e^{i2\pi f t} \, df$$

$$= \int_{t'=-\infty}^{\infty} x(t') \left[ \underbrace{\int_{f=-\infty}^{\infty} e^{+i2\pi f (t-t')} \, df}_{\delta(t-t')} \right] dt'$$

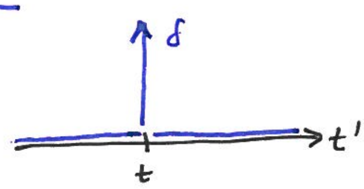
$$= \int_{t'=-\infty}^{\infty} x(t')\, \delta(t-t')\, dt'$$

$$= x(t)$$

☑

Here $\delta(\,)$ is the Dirac delta function:

Roughly speaking: $\delta(t-t') = \begin{cases} 0 & t \neq t' \\ \infty & t = t' \end{cases}$




more rigorous definition: $\int_{-\infty}^{\infty} \delta(t-t')\, f(t')\, dt' = f(t)$ for any function $f$.

## Properties of the Fourier Transform:

① If $x(t)$ is real then $\tilde{x}(-f) = \tilde{x}^*(f)$ so we can focus on positive frequencies only, if we wish.

② Parseval's theorem (or Plancherel's theorem):

$$\int_{-\infty}^{\infty} \tilde{h}(t)\, \tilde{g}^*(t)\, dt = \int_{-\infty}^{\infty} \tilde{h}(f)\, \tilde{g}^*(f)\, df$$

③ The time and frequency domains are interchangeable - any calculation done with $x(t)$ can also be done with $\tilde{x}(f)$.

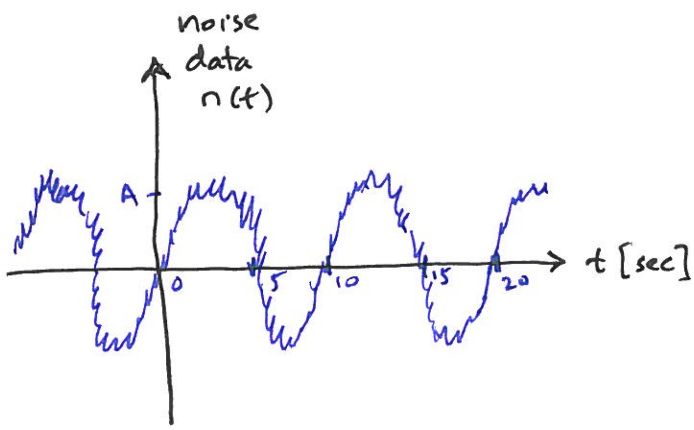
→ See Numerical Recipes, Ch 12 for more properties.

## Why are Fourier Transforms useful in signal detection?

① Many signals and noise models have features which are simpler to model as functions of frequency than of time.

② Formulae in terms of frequencies are often simpler to write and faster to compute than their time-domain equivalents.

## Toy Example: ~Periodic Noise




$n(t) \simeq$ sine wave with frequency $\sim 1/10\,\text{sec}$ + higher-frequency junk

$\simeq \underset{\substack{\text{large}\\ \text{amplitude}\\ \text{"A"}}}{A} \sin(2\pi \times \overbrace{0.1\,\text{Hz}}^{\text{Freq.}} \times t)$ + higher-freq terms. (sines and cosines)

## Noise Power Spectral Density

The simplest case is noise that is stationary and Gaussian. Then $n(t)$ is fully characterised by its autocorrelation function $C(t)$:

$$C(t) \equiv E\left[ n(t_0)\, n(t_0 + t) \right]$$

E: average over all noise realisations

independent of start time $t_0$ ("stationary")

Toy noise example: 10 sec periodicity

if $n(t_0) > 0$ then probably $n(t_0 + 10s) > 0$ as well
" " $< 0$ " " " $< 0$ " "

We'll see it is more convenient to describe the noise in the frequency domain. Consider the Fourier transform, $\tilde{n}(f)$:

$$E\left[\tilde{n}(f)\, \tilde{n}^*(f')\right] = E\left[ \int_{t_1=-\infty}^{\infty} n(t_1)\, e^{-i2\pi f t_1}\, dt_1 \int_{t_2=-\infty}^{\infty} n^*(t_2)\, e^{i2\pi f' t_2}\, dt_2 \right]$$

define new variable "$t$" by $t = t_2 - t_1$ to replace $t_2$

$n$ real

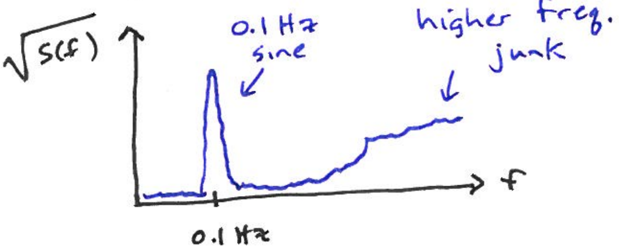
$$= \int_{t_1} \int_{t} \underbrace{E\left[ n(t_1)\, n^*(t_1 + t) \right]}_{C(t)} e^{-i2\pi (f-f')t_1}\, e^{i2\pi f' t}\, dt\, dt_2$$

$$= \underbrace{\left( \int_{t_1=-\infty}^{\infty} e^{-i2\pi (f-f')t_1}\, dt_1 \right)}_{\delta(f-f')} \underbrace{\left( \int_{t=-\infty}^{\infty} C(t)\, e^{i2\pi f' t}\, dt \right)}_{\text{Fourier transform of } C(t)}$$

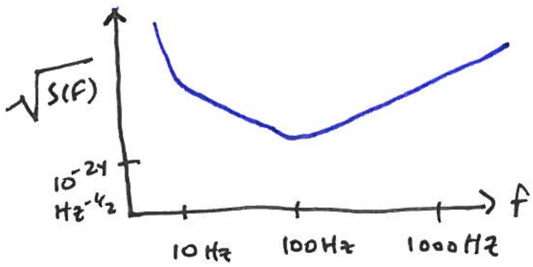
$\Rightarrow$ call "$S(f')$", the noise power spectral density.

$$\therefore E\left[\tilde{n}(f)\, \tilde{n}^*(f')\right] = \delta(f-f')\, S(f) \quad \left[\text{or } S(f')\right]$$

Toy Noise Example:




LIGO Noise:



**BEWARE:** Since $S(-f) = S(f)$, it is common to consider the one-sided spectral density $S_1(f)$, where

$$S_1(f) \equiv 2\,S(f)$$

and work with $f > 0$ only. These notes stick with the two-sided spectrum $S(f)$ because its simpler for the maths. The difference causes <u>many</u> factor of 2 errors in calculations!

Interpretation: $S(f)$ is the average value of $|\tilde{n}(f)|^2$ – the squared noise in a given frequency bin.

## Matched Filters

Suppose we know the signal $h(t)$ we're looking for. How do we determine if it is present in our noisy data?

$$\underbrace{d(t)}_{\text{data}} = \underbrace{n(t)}_{\text{noise}} + \underbrace{h(t)}_{\text{signal}}$$

(may or may not be present)

General approach: Integrate the data with a "filter function" $g(t)$:

$$\begin{aligned}\theta &= \int d(t)\, g(t)\, dt \qquad \text{Parseval} \\ &= \int \tilde{d}(f)\, \tilde{g}^*(f)\, df \\ &= \underbrace{\int \tilde{n}(f)\, \tilde{g}^*(f)\, df}_{\text{noise response}} + \underbrace{\int \tilde{h}(f)\, \tilde{g}^*(f)\, df}_{\text{signal response}}\end{aligned}$$

Rule: Pick $\tilde{g}(f)$ to make the signal response as large as possible relative to the typical noise response. To be precise, pick $\tilde{g}$ to maximise the signal-to-noise ratio (SNR) squared:

$$\begin{aligned}\rho^2 &\equiv \frac{(\text{signal response})^2}{\text{variance of noise response}} \\ &= \frac{\left(\int \tilde{h}(f)\, \tilde{g}^*(f)\, df\right)^2}{E\left[\,\left|\int \tilde{n}(f')\, \tilde{g}^*(f')\, df'\right|^2\right]}\end{aligned}$$

$$\begin{aligned}\text{denominator} &= E\left[\int_{f=-\infty}^{\infty} \tilde{n}(f)\, \tilde{g}^*(f)\, df \int_{f'=-\infty}^{\infty} \tilde{n}^*(f')\, \tilde{g}(f')\, df'\right] \\ &= \int_{f=-\infty}^{\infty}\int_{f'=-\infty}^{\infty} \underbrace{E\left[\tilde{n}(f)\, \tilde{n}^*(f')\right]}_{\delta(f-f')\, S(f)} \tilde{g}^*(f)\, \tilde{g}(f')\, df\, df' \\ &= \int_{f=-\infty}^{\infty} S(f)\, |\tilde{g}(f)|^2\, df\end{aligned}$$

**ASIDE:** A very common integral in these problems is the "noise weighted inner product". For two functions $a(t), b(t)$

$$\langle a,b\rangle \equiv \int_{f=-\infty}^{\infty} \frac{\tilde{a}(f)\,\tilde{b}^*(f)}{S(f)}\,df$$

- Imagine $\tilde{a}(f)$, $\tilde{b}(f)$ as two vectors where $f$ is the element of the vector. Then $\langle a,b\rangle$ is the frequency-by-frequency dot product with weighting $S(f)$ applied to down-weight frequencies with large noise contamination.
- The squared length or magnitude of $a$ is

$$\|a\|^2 \equiv \langle a,a\rangle = \int_{-\infty}^{\infty} \frac{|\tilde{a}(f)|^2}{S(f)}\,df$$

So the denominator of the $\rho^2$ formula is

$$\text{denominator} = \int_{-\infty}^{\infty} S(f)|\tilde{g}|^2 \cdot \frac{S(f)}{S(f)}\,df$$

$$= \langle Sg, Sg\rangle$$

The numerator of the $\rho^2$ formula is

$$\text{numerator} = \left(\int \tilde{h}(f)\,\tilde{g}^*(f)\,df\right)^2$$

$$= \left(\int \frac{\tilde{h}(f)\,(\tilde{g}^*(f)\,S(f))}{S(f)}\,df\right)^2$$

$$= (\langle h, Sg\rangle)^2$$

So

$$\rho^2 = \frac{(\langle h, Sg\rangle)^2}{\langle Sg, Sg\rangle}$$

How do we select $g$?

<u>Note:</u> $\rho^2$ is unchanged by an overall constant rescaling of $g$. Eg for $g \to \lambda g$ ($\lambda$ constant) we get

$$\rho^2 \to \frac{(\langle h, \lambda Sg\rangle)^2}{\langle \lambda Sg, \lambda Sg\rangle} = \frac{\cancel{\lambda^2}}{\cancel{\lambda^2}} \frac{(\langle h, Sg\rangle)^2}{\langle Sg, Sg\rangle} = \rho^2$$

So we can pick $g$ to have unit length without loss of generality: $\langle Sg, Sg\rangle = 1$.

Then

$$\rho^2 = \left( \langle h, Sg \rangle \right)^2$$

signal vector — (unit) filter vector

We can maximise $\rho^2$ by choosing

$$Sg \propto h \quad \Rightarrow \quad \boxed{g(f) = N \frac{h(f)}{S(f)}}$$

signal vector

noise spectrum.

normalisation constant (to get $\langle gS, gS \rangle = 1$)

The normalisation constant is defined by

$$1 = \langle Sg, Sg \rangle = N^2 \int \frac{|\tilde{h}(f)|^2}{S(f)} df = N^2 \langle h, h \rangle$$

$$\therefore \; N = 1 / \sqrt{\langle h, h \rangle} = 1 \Big/ \sqrt{\int \frac{|\tilde{h}(f)|^2}{S(f)} df}$$

Interpretation: our optimal filter $g$ points in the direction of the noise-weighted signal vector.

compare: $\vec{a} \cdot \vec{b} = ab \cos\theta \rightarrow$ maximised for $\theta = 0$ ($\vec{a}$ and $\vec{b}$ parallel)

Our "matched filter" is therefore

$$\begin{aligned} Q &= \int_{-\infty}^{\infty} \tilde{d}(f) \tilde{g}^*(f) \, df \\ &= N \int_{-\infty}^{\infty} \frac{\tilde{d}(f) \tilde{h}^*(f)}{S(f)} df \\ &= \frac{\langle d, h \rangle}{\sqrt{\langle h, h \rangle}} = \left\langle d, \frac{h}{\|h\|} \right\rangle \end{aligned}$$

$\Updownarrow$

matched filter = "dot product of data vector with a unit-length signal vector."

Consider how $\rho$ behaves when a signal is present:

$$d = n + h$$

$$\rho = \frac{\langle d, h\rangle}{\sqrt{\langle h,h\rangle}} = \underbrace{\frac{1}{\sqrt{\langle h,h\rangle}} \int_{-\infty}^{\infty} \frac{\tilde{n}(f)\,\tilde{h}^*(f)}{S(f)}\,df}_{\text{noise term}} + \underbrace{\sqrt{\langle h,h\rangle}}_{\text{signal term}}$$

The signal contribution is just $\rho$:

$$\rho^2 = \left(\langle h, S g\rangle\right)^2 \xrightarrow[g=\frac{Nh}{S}]{} \left(N\langle h,h\rangle\right)^2 = \frac{\langle h,h\rangle^2}{\langle h,h\rangle} = \langle h,h\rangle$$

The noise contribution is a Gaussian with zero mean and unit variance:

$$E\left[\begin{matrix}\text{noise}\\ \text{term}\end{matrix}\right] = \frac{1}{\sqrt{\langle h,h\rangle}} \int_{-\infty}^{\infty} \frac{\overbrace{E[\tilde{n}(f)]}^{\to 0 \text{ for zero-mean noise}}\,\tilde{h}^*(f)}{S(f)}\,df$$

$$\text{variance}\left[\begin{matrix}\text{noise}\\ \text{term}\end{matrix}\right] = E\left[(\text{noise term})^2\right]$$

$$= E\left[\frac{1}{\langle h,h\rangle} \int_{-\infty}^{\infty} \frac{\tilde{n}(f)\,\tilde{h}^*(f)}{S(f)}\,df \int_{-\infty}^{\infty} \frac{\tilde{n}^*(f')\,\tilde{h}(f')\,df'}{S(f')}\right]$$

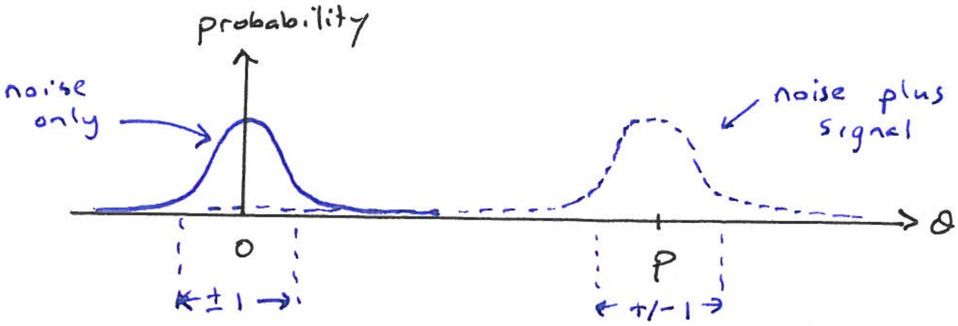
$$= \frac{1}{\langle h,h\rangle} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\underbrace{E\left[\tilde{n}(f)\,\tilde{n}^*(f')\right]}_{\delta(f-f')\,\cancel{S(f)}}\,\tilde{h}^*(f)\,\tilde{h}(f')}{\cancel{S(f)}\,S(f')}\,df\,df'$$

$$= \frac{1}{\langle h,h\rangle} \underbrace{\int_{-\infty}^{\infty} \frac{|\tilde{h}(f)|^2}{S(f)}\,df}_{\langle h,h\rangle}$$

$$= 1$$

So when a signal is present

$$\rho = \underset{\text{SNR}}{\rho} + \underset{\text{unit normal random number}}{N(0,1)}$$



It is conventional to write the formula for the SNR $\rho$ in terms of the one-sided noise spectrum $S_1(f)$:

$$\rho = \langle h,h\rangle = \underbrace{\int_{f=-\infty}^{\infty}}_{=2\int_0^\infty} \frac{|\tilde{h}(f)|^2}{\underbrace{S(f)}_{=\frac{1}{2}S_1(f)}}\,df = 4\int_0^\infty \frac{|\tilde{h}(f)|^2}{S_1(f)}\,df$$