GW data analysis

Edoardo Milotti

January 12, 2023

In this handout I describe some basic techniques in GW data analysis.

Preliminaries:

- the Fourier transform of a signal $x(t)$ and its inverse are defined by the formulas

$$\tilde{x}(f) = \int_{-\infty}^{+\infty} x(t) e^{-2\pi i f t} dt, \quad x(t) = \int_{-\infty}^{+\infty} \tilde{x}(f) e^{2\pi i f t} df \tag{1}$$

- the action of filters in the time domain is described by the convolution

$$(x * y)(t) = \int_{-\infty}^{+\infty} x(t') y(t - t') dt' \tag{2}$$

- consider the Fourier transform of the convolution of two signals $x(t)$ and $y(t)$ with Fourier transforms $\tilde{x}(f)$ and $\tilde{y}(f)$

$$\begin{aligned}
&\int_{-\infty}^{+\infty} (x * y)(t) e^{-2\pi i f t} dt = \\
&\quad = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} x(t') y(t - t') dt' \right] e^{-2\pi i f t} dt \\
&\quad = \int_{-\infty}^{+\infty} x(t') e^{-2\pi i f t'} dt' \int_{-\infty}^{+\infty} y(t - t') e^{-2\pi i f (t - t')} dt \\
&\quad = \int_{-\infty}^{+\infty} x(t') e^{-2\pi i f t'} dt' \int_{-\infty}^{+\infty} y(t'') e^{-2\pi i f t''} dt'' \\
&\quad = \tilde{x}(f) \tilde{y}(f)
\end{aligned} \tag{3}$$

  (convolution theorem)

- a process is *stationary* when all its statistics are constant in time, i.e., when its probability distribution is invariant with respect to time translations. A noise can be weakly stationary when only some of its statistics are time-invariant, for example only mean and variance.

- a process is *ergodic* when time averages are equal to ensemble averages.

- the total square fluctuation of a real stationary signal $s(t)$ is given by

$$\int_{-\infty}^{+\infty} |s(t)|^2 dt = \int_{-\infty}^{+\infty} dt \left| \int_{-\infty}^{+\infty} \tilde{s}(f) e^{2\pi i f t} df \right|^2 \tag{4}$$

$$= \int_{-\infty}^{+\infty} df' \int_{-\infty}^{+\infty} df'' \tilde{s}(f') \tilde{s}^*(f'') \int_{-\infty}^{+\infty} e^{2\pi i (f'-f'')t} dt \tag{5}$$

$$= \int_{-\infty}^{+\infty} df' \int_{-\infty}^{+\infty} df'' \tilde{s}(f') \tilde{s}^*(f'') \delta(f' - f'') \tag{6}$$

$$= \int_{-\infty}^{+\infty} |\tilde{s}(f')|^2 df \tag{7}$$

(Parseval's theorem).

- the *power spectral density* (PSD) of a signal $s(t)$ is usually defined as follows

$$S(f) = \lim_{T\to\infty} \frac{1}{T} \left| \int_{-T/2}^{+T/2} s(t) e^{-2\pi i f t} dt \right|^2 , \tag{8}$$

which is a *two-sided spectral density* where the frequency runs over negative as well as positive values. However, for real signals – as in the case of the $h(t)$ signal recorded by a GW IFO – the PSD is an even function, i.e., $S(-f) = S(f)$, and for this reason it is customary to define and use the *one-sided spectral density*

$$S(f) = \lim_{T\to\infty} \frac{2}{T} \left| \int_{-T/2}^{+T/2} s(t) e^{-2\pi i f t} dt \right|^2 , \tag{9}$$

where $f \geq 0$.

- the definition is similar for a noise signal $n(t)$, the only difference is an ensemble average

$$S(f) = \lim_{T\to\infty} \frac{2}{T} \left\langle \left| \int_{-T/2}^{+T/2} n(t) e^{-2\pi i f t} dt \right|^2 \right\rangle , \tag{10}$$

which is again a one-sided spectral density.

- the autocorrelation function of a zero-mean stationary, ergodic process $s(t)$ is defined by

$$R(\tau) = \langle s(t) s(t+\tau) \rangle = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} s(t) s(t+\tau) dt \tag{11}$$

- autocorrelation function and PSD are closely related. Consider the PSD

$$S(f) = \lim_{T\to\infty} \frac{2}{T} \left| \int_{-T/2}^{+T/2} s(t) e^{-2\pi i f t} dt \right|^2 \tag{12}$$

$$= \lim_{T\to\infty} \frac{2}{T} \int_{-T/2}^{+T/2} s(t') e^{2\pi i f t'} dt' \int_{-T/2}^{+T/2} s(t) e^{-2\pi i f t} dt \tag{13}$$

Now let $t = t' + \tau$, so that $dt = d\tau$ in the second integral (where t' behaves as a constant), then we obtain

$$S(f) = \lim_{T\to\infty} \frac{2}{T} \int_{-T/2}^{+T/2} s(t') e^{2\pi i f t'} dt' \int_{-T/2}^{+T/2} s(t'+\tau) e^{-2\pi i f t'} e^{-2\pi i f \tau} d\tau \tag{14}$$

$$= \int_{-\infty}^{+\infty} e^{-2\pi i f \tau} d\tau \lim_{T\to\infty} \frac{2}{T} \int_{-T/2}^{+T/2} s(t') s(t'+\tau) dt' \tag{15}$$

$$= \int_{-\infty}^{+\infty} R(\tau) e^{-2\pi i f \tau} d\tau \tag{16}$$

i.e., the PSD is the Fourier transform of the autocorrelation function (this is the *Wiener-Kintchine theorem*).

- if the noise is white, stationary and Gaussian, it is completely characterized by its variance N_0 in the time domain. When one takes a two-sided spectral representation, this means (from Parseval's theorem) that

$$S(f) = N_0/2 \tag{17}$$

Gaussian noise

By definition a sampled noise is Gaussian when the samples have a Gaussian distribution. Here we assume that samples of a Gaussian white noise process are taken with time step Δt, and that the variance of each sample is σ^2, so that the correlation function is $R_{jk} = \langle x_j x_k \rangle = \sigma^2 \delta_{jk}$. Using a discretized version of the Wiener-Kintchine theorem, we find the PSD

$$S_x(f) \approx 2 \sum_{j=1,N} R_{jk} e^{2\pi i f (j-k)\Delta t} \Delta t = 2\sigma^2 \Delta t \tag{18}$$

The probability density function of each sample x_i is

$$p(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{x_i^2}{2\sigma^2}\right], \tag{19}$$

therefore the the joint probability density function of N samples is

$$p(\{x_i\}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{\sum_{i=1,N} x_i^2}{2\sigma^2}\right] = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{\sum_{i=1,N} x_i^2 \Delta t}{2\sigma^2 \Delta t}\right] \tag{20}$$

$$\sim \exp\left[-\frac{1}{S_x} \int_{-\infty}^{+\infty} x^2(t) dt\right] \tag{21}$$

$$= \exp\left[-\int_{-\infty}^{+\infty} \frac{|\tilde{x}(f)|^2}{S_x} df\right] \tag{22}$$

Now notice that we can generate non-white noises by proper filtering of a white noise. In the time domain, the filtering operation is described by a convolution

$$y(t) = \int_{-\infty}^{+\infty} k(t-t') x(t') dt' \tag{23}$$

where $k(t)$ is the pulse response function of the filter and its Fourier transform $K(f)$ is the transfer function of the filter. Then, in the frequency domain, we find

$$\tilde{y}(f) = K(f)\tilde{x}(f) \tag{24}$$

from the convolution theorem: this implies that

$$S_y(f) = |K(f)|^2 S_x(f) \tag{25}$$

Therefore

$$p(\{x_i\}) \sim \exp\left[-\int_{-\infty}^{+\infty} \frac{|\tilde{x}(f)|^2}{S_x} df\right] = \exp\left[-\int_{-\infty}^{+\infty} \frac{|K(f)|^2 |\tilde{x}(f)|^2}{|K(f)|^2 S_x} df\right] = \exp\left[-\int_{-\infty}^{+\infty} \frac{|\tilde{y}(f)|^2}{S_y} df\right] \tag{26}$$

and we see that the formula holds also for non-white noise.

In addition, we can use the last result for the argument of the exponential as a motivation to introduce a scalar product in this function space:

$$(x, y) = 2 \text{ Re} \int_{-\infty}^{+\infty} \frac{\tilde{x}(f)\tilde{y}^*(f)}{S} df, \tag{27}$$

or

$$(x, y) = 4 \text{ Re} \int_{0}^{+\infty} \frac{\tilde{x}(f)\tilde{y}^*(f)}{S} df, \tag{28}$$

when we use a single-sided spectral density, so that we can write the probability as a function of such a scalar product

$$p(\{x_i\}) \sim \exp\left[-\frac{(x, x)}{2}\right]. \tag{29}$$

Optimal detection statistic and Bayes' theorem

Recall that for two hypotheses – null hypothesis $\mathcal{H}_0$ and alternative hypothesis $\mathcal{H}_1$ – Bayes' theorem writes

$$P(\mathcal{H}_{0,1}|s) = \frac{P(s|\mathcal{H}_{0,1})P(\mathcal{H}_{0,1})}{P(s|\mathcal{H}_0)P(\mathcal{H}_0) + P(s|\mathcal{H}_1)P(\mathcal{H}_1)} \tag{30}$$

and given the data, we select the hypothesis that maximizes the posterior probability $P(\mathcal{H}_{0,1}|s)$, this is the Maximum A Posteriori (MAP) choice. We can also consider the *odds ratio*

$$\frac{P(\mathcal{H}_1|s)}{P(\mathcal{H}_0|s)} = \frac{P(s|\mathcal{H}_1)P(\mathcal{H}_1)}{P(s|\mathcal{H}_0)P(\mathcal{H}_0)} \tag{31}$$

which reduces to

$$\frac{P(\mathcal{H}_1|s)}{P(\mathcal{H}_0|s)} = \frac{P(s|\mathcal{H}_1)}{P(s|\mathcal{H}_0)} \tag{32}$$

if the prior probabilities of the null and of the alternative hypothesis are equal (in this case the odds ratio is called the *Bayes factor*). It is noteworthy that the Bayes factor is called the *likelihood ratio* in frequentist statistics, in which context it is shown to be the "most powerful test of size α" (Neyman-Pearson lemma). Given that the logarithm is a monotonically increasing function, the argument works also for the log likelihood ratio.

Matched filters

The detection problem that we face in GW data analysis involves a null hypothesis where the signal is just noise, $s(t) = n(t)$, and an alternative hypothesis where the signal is given by the sum of a GW signal plus noise, $s(t) = n(t) + h(t)$. Then, the likelihood ratio is

$$\Lambda = \frac{p(s|\mathcal{H}_1)}{p(s|\mathcal{H}_0)} = \exp\left\{\frac{1}{2}[-(s-h, s-h) + (s, s)]\right\} = \exp\left[(s, h) - (h, h)/2\right]. \tag{33}$$

Since the likelihood ratio depends on data only through the (s, h) product, which is a log likelihood ratio, we conclude that this is the optimal detection statistic, i.e.,

$$(s, h) = 4 \,\mathrm{Re} \int_0^{+\infty} \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)} df \tag{34}$$

where I used tildes to denote Fourier transforms to avoid confusion with other symbols, and where $S_n(f)$ is the noise PSD. Eq. (34) defines the *matched filter*[1].

Here we remark that in eq. (34) the expression $\tilde{s}(f)/\sqrt{S_n(f)}$ is the Fourier transform of the *whitened signal* and $\tilde{h}^*(f)/\sqrt{S_n(f)}$ is the *whitened filter transfer function.*

Now consider the inverse Fourier transform of the conjugate of a Fourier transform

$$\int_{-\infty}^{+\infty} \tilde{x}^*(f) e^{2\pi i f t} df = \left[\int_{-\infty}^{+\infty} \tilde{x}^(f) e^{2\pi i f(-t)} df\right]^* = x(-t), \tag{35}$$

we find that it represents the time-reversed signal. Therefore, when we consider eq. (34), we see that it corresponds to a time convolution where the template signal h is time-reversed (see figure 1).

Signal-to-noise ratio (SNR)

The ideal matched filter has a template signal h that is equal to the detected signal s, and in that case

$$\rho^2_{\mathrm{opt}} = (h, h) = 4 \int_0^{+\infty} \frac{|\tilde{h}(f)|^2}{S_n(f)} df \tag{36}$$

which is the *optimal power signal-to-noise ratio*; its square root $\rho = \sqrt{\rho^2}$ is the *amplitude signal-to-noise ratio.*

[1] This is the Bayesian derivation of the matched filter, which has the advantage of connecting the filter to important statistical concepts and to the Gaussian distribution, however other derivations also exist.

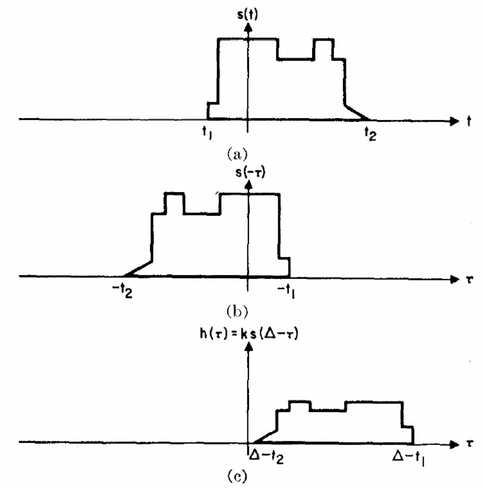

Figure 1: Representation of the template signal in the time-domain. (a) A wave train; (b) the reversed train; (c) a matched-filter impulse response. Figure from G. L. Turin, *An Introduction to Matched Filters*, IRE Trans. on Information Theory **6** (1960) 311.