Appendix B

Statistical quantities

B.1 Signals of finite duration

Consider two complex functions $g(t)$ and $h(t)$, where t is real. Their *convolution* is defined as

$$(g * h)(t) = \int_{-\infty}^{+\infty} g(t-\tau)h(\tau)d\tau = (h * g)(t) \tag{B.1}$$

This function is only defined if one or both of the functions decay rapidly enough for $\tau \to \pm\infty$. A precise condition is a bit tricky. But let us take as an example a function $g(t)$ that does not diverge anywhere but remains substantially non-zero even for $\tau \to \pm\infty$. Take $h(t)$ to be a function that decays rapidly for $\tau \to \pm\infty$ and has normalization $\int_{-\infty}^{+\infty} h(\tau)d\tau = 1$, for instance: The Gauss function $h(\tau) = \exp(-\tau^2)/\sqrt{\pi}$. Then the convolution exists and is for this example a smeared-out version of the original function $g(t)$. In a similar manner, the *correlation* of two complex functions $g(t)$ and $h(t)$ is defined as

$$\mathrm{Corr}[g,h](t) = \int_{-\infty}^{+\infty} g(t+\tau)h^*(\tau)d\tau \tag{B.2}$$

Here the same issues of their existence hold.

Assume from now on that both $g(t)$ and $h(t)$ are signals of finite duration, i.e. their values (which can be complex) go to zero at large $|t|$.

We can then derive many interesting properties for these quantities. For instance convolution and correlation functions have interesting behavior in Fourier space:

$$\mathcal{F}[(g*h)](u) = \mathcal{F}[g](u)\mathcal{F}[h](u) \tag{B.3}$$

$$\mathcal{F}[\mathrm{Corr}[g,h]](u) = \mathcal{F}[g](u)(\mathcal{F}[h](u))^* \tag{B.4}$$

where $\mathcal{F}$ is the Fourier operator:

$$\mathcal{F}[g](u) = \int_{-\infty}^{+\infty} g(t)e^{-2\pi i u t}dt \tag{B.5}$$

If we now take the correlation between $g(t)$ and itself, we obtain the *autocorrelation*[1],

[1] This is often also called more precisely the *autocovariance*, but note that, strictly speaking, autocovariance and autocorrelation are normalized differently, the autocorrelation being the autocovariance normalized by the variance. In this script we will not be so precise in this nomenclature.

which is defined as

$$B_g(t) = \text{Corr}[g,g](t) = \int_{-\infty}^{+\infty} g(t+\tau)g^*(\tau)d\tau \tag{B.6}$$

The *structure function* is defined as

$$D_g(t) = \int_{-\infty}^{+\infty} |g(t+\tau) - g(\tau)|^2 d\tau \tag{B.7}$$

Let us write out the structure function:

$$\begin{aligned} D_g(t) &= \int_{-\infty}^{+\infty} [g(t+\tau)-g(\tau)][g(t+\tau)-g(\tau)]^* d\tau \\ &= \int_{-\infty}^{+\infty} [g(t+\tau)g^*(t+\tau) + g(\tau)g^*(\tau) - g(t+\tau)g^*(\tau) - g(\tau)g^*(t+\tau)]d\tau \\ &= 2B_g(0) - B_g(t) - B_g^*(t) \end{aligned} \tag{B.8}$$

If $g(t)$ is a real function, then we obtain

$$D_g(t) = 2[B_g(0) - B_g(t)] \tag{B.9}$$

From Eq. (B.4) we can also directly obtain the *Wiener-Khinchin theorem*:

$$\mathcal{F}[\text{Corr}[g,g]](u) = |\mathcal{F}[g](u)|^2 \tag{B.10}$$

which can also be derived explicitly as follows:

$$\mathcal{F}[\text{Corr}[g,g]](u) = \int_{-\infty}^{+\infty} e^{-2\pi i u t} \left[\int_{-\infty}^{+\infty} g(t+\tau)g^*(\tau)d\tau\right] dt \tag{B.11}$$

$$= \int_{-\infty}^{+\infty} g^*(\tau) \left[\int_{-\infty}^{+\infty} e^{-2\pi i u t} g(t+\tau)dt\right] d\tau \tag{B.12}$$

$$= \int_{-\infty}^{+\infty} g^*(\tau)e^{+2\pi i u \tau} \left[\int_{-\infty}^{+\infty} e^{-2\pi i u(t+\tau)} g(t+\tau)d(t+\tau)\right] d\tau \tag{B.13}$$

$$= \left[\int_{-\infty}^{+\infty} g^*(\tau)e^{+2\pi i u \tau}d\tau\right] \left[\int_{-\infty}^{+\infty} e^{-2\pi i u t'} g(t')dt'\right] \tag{B.14}$$

$$= (\mathcal{F}[g]^*)\mathcal{F}[g] \tag{B.15}$$

Similarly we can derive *Parseval's theorem*:

$$\int_{-\infty}^{+\infty} |g(t)|^2 dt = \int_{-\infty}^{+\infty} |\mathcal{F}[g](u)|^2 du \tag{B.16}$$

All the above quantities (convolution, correlation, autocorrelation, structure function) all depend on the duration of the signal. Let us assume that the signal exists between $t = 0$ and $t = T$ but is zero before $t = 0$ or after $t = T$. If the stochastic properties of the signal do not change in time but we increase T by a factor 2, then all the above values become roughly twice as large. The definitions we used are therefore not bound if the signal duration gets ever longer. We will discuss this normalization issue in Section B.3. But first we must take a closer look at what a "stochastic signal" means.

B.2 Stochastic signals

Let us take a step back and consider a stochastic signal $g(t)$ a bit more mathematically. Let us define an "experiment" to be one incarnation of the function $g(t)$. But one could do many experiments, yielding many functions $g_i(t)$ for, say, $i \in [1, N]$, where N is the number of experiments. An example is the thermal motion of molecules, where $g_i(t)$ is in that case the velocity vector $\vec{v}_i(t)$ for molecule i. We thus have an *ensemble* of functions $g(t)$. At any time t we thus have N values of $g(t)$: $\{g_i(t)\}$. For $N \to \infty$ the ensemble $\{g_i(t)\}$ defines the statistical properties of the stochastic function $g(t)$. We can now define the propability function of finding the value of g_i (for some arbitrary i) to be between a and $a + da$ at time t:

$$P_g(a;t)da \tag{B.17}$$

It is normalized as

$$\int_{-\infty}^{+\infty} P_g(a;t)da = 1 \tag{B.18}$$

We can define the average:

$$\langle g(t)\rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1,N} g_i(t) = \int_{-\infty}^{+\infty} P_g(a;t)a da \tag{B.19}$$

Let us write for convenience:

$$\eta_g(t) \equiv \langle g(t)\rangle \tag{B.20}$$

The average is therefore the first moment of the probability function. Note that "average" here is not defined in time (it is in fact taken *at* time t): It is an average over the ensemble. However, as we shall see below, the time-average is very closely related. Often this average is written as the *expectation value* $E[]$:

$$\eta_g(t) = E[g(t)] \tag{B.21}$$

Let us continue with defining the standard deviation σ_g^2:

$$\sigma_g^2(t) = \frac{1}{N} \sum_{i=1,N} (g_i(t) - \eta_g(t))^2 \tag{B.22}$$

$$= E[(g(t) - \eta_g(t))^2] \tag{B.23}$$

$$= \int_{-\infty}^{+\infty} P_g(a;t)(a - \eta_g(t))^2 da \tag{B.24}$$

We can also define te probability that we find the value of g_i between a and $a + da$ at time t_1 *and* find it between b and $b + db$ at time t_2:

$$P_g(a, b; t_1, t_2)da\, db \tag{B.25}$$

normalized as

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P_g(a, b; t_1, t_2)da\, db = 1 \tag{B.26}$$

If the value of $g(t)$ at t_2 is totally uncorrelated with the value some time earlier t_1, then $P_g(a, b; t_1, t_2) = P_g(a; t_1)P_g(b; t_2)$, but if there is a correlation, then this is not the case.

Going back to our example of thermal molecular motion: if t_2 is sufficiently close to t_1, then it is unlikely that the velocity has changed very much in between these two times: they are thus correlated over some time interval.

Using this formalism let is now re-define the concepts of *autocorrelation*, *structure function* that we already encountered in Section B.1, but now we do it more thoroughly (and sometimes with a different normalization, but we will come back to this).

So let us define the *autocorrelation* $B_g(t_1, t_2)$ of the stochastic process $g(t)$ as:

$$B_g(t_1, t_2) = E[g(t_1)g(t_2)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P_g(a, b; t_1, t_2)\, a\, b\, da\, db \qquad \text{(B.27)}$$

and the *autocovariance* $C_g(t_1, t_2)$ as

$$C_g(t_1, t_2) = E[(g(t_1) - \eta_g(t_1))(g(t_2) - \eta_g(t_2))] \qquad \text{(B.28)}$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P_g(a, b; t_1, t_2)\, (a - \eta_g(t_1))(b - \eta_g(t_2))\, da\, db \qquad \text{(B.29)}$$

They obey the following properties:

$$C_g(t_1, t_2) = B_g(t_1, t_2) - \eta_g(t_1)\eta_g(t_2) \qquad \text{(B.30)}$$

$$\sigma_g^2(t) = C_g(t, t) = B_g(t, t) - \eta_g(t)^2 \qquad \text{(B.31)}$$

Furthermore the *structure function* $D_g(t_1, t_2)$ is defined as

$$D_g(t_1, t_2) = E[|g(t_1) - g(t_2)|^2] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P_g(a, b; t_1, t_2)\, |a - b|^2\, da\, db \qquad \text{(B.32)}$$

If our stochastic function $g(t)$ is *stationary*, i.e. that its stochastic properties are the same as for $g(t - \tau)$ for any value of τ, then all the above quantities only depend on $t = |t_2 - t_1|$:

$$B_g(t) = E[g(0)g(t)] = E[g(\tau)g(\tau + t)] \qquad \text{(B.33)}$$

$$C_g(t) = E[(g(0) - \eta_g(0))(g(t) - \eta_g(t))]$$

$$= E[(g(\tau) - \eta_g(\tau))(g(\tau + t) - \eta_g(\tau + t))] \qquad \text{(B.34)}$$

$$D_g(t) = E[|g(0) - g(t)|^2] = E[|g(\tau) - g(\tau + t)|^2] \qquad \text{(B.35)}$$

where τ can be any value.

As you can see, as opposed to the definitions in Section B.1, the autocorrelation and the structure function are normalized. But apart from that, they are strongly related, as we shall see in Section B.3.

B.3 Ergodic stochastic signals

In many cases one does not have N incarnations of a stochastic signal $g(t)$, but instead a single signal $g(t)$ for a very long time period. An example is the wave front of light at a telescope: what we observe is the *time average* of the square of the electric field: $\langle EE^* \rangle$. Since the electric field signal is clearly a stationary stochastic signal, the time average over any long enough period of time is sufficient information.

A signal is called *ergodic* if the stochastic properties of an entire ensemble of incarnations of the function $g(t)$ can be derived from one individual incarnation of the function $g(t)$, when it is studied over its entire time domain. A simple example: If we study the velocities of molecules that undergo thermal motion we have two choices: we can study the velocity distribution of a million molecules at a single instance in time, or we can study the velocity evolution of a single molecule over millions of collision times. Since the molecule loses memory of its original velocity in a few collisions, and since the molecules are all the same, the two ways of investigating the stochastics of molecular motion yield the same results: this stochastic variable is ergodic. Counter-example: if we study the time in the day that a human goes to bed, we could study the bed-time of thousands of people on January 1, 2015, or we can study the bed-time of a single person over his/her entire life. The results are, however, *not* the same, because each instance of a human being is different from another. The stochastic variable "bed time" is therefore not ergodoc.

For most of the stochastic variables we are concerned with in the study of techniques of observational astronomy the signals are, fortunately, stationary and ergodic. This explains why we defined autocorrelation, structure function etc. in Section B.1 as an integral over time: we use time as a way to compute expectation values. The problem is, however, that this only works perfectly if we average over an infinitely long time, which is in practice not possible.

So let us redo what we did in Section B.2, but now in a more careful way. We will also take care of normalization. Let $g(t)$ and $h(t)$ be two real or complex-valued stochastic signals, and let us assume that their mean values are zero (for convenience):

$$\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{+T/2} g(t)dt = 0 \qquad \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{+T/2} h(t)dt = 0 \tag{B.36}$$

Let us also assume that both signals are non-zero even for large $|t|$. These are *ever continuing signals*. That means that we would not be able to define their convolution and correlation according to Section B.2, because the integrals would diverge. But we can re-normalize these integrals and compute them over a finite range:

$$(g * h)(t) \simeq \frac{1}{T}\int_{-T/2}^{+T/2} g(t-\tau)h(\tau)d\tau \tag{B.37}$$

$$\mathrm{Corr}[g,h](t) \simeq \frac{1}{T}\int_{-T/2}^{+T/2} g(t+\tau)h^*(\tau)d\tau \tag{B.38}$$

These are only estimates. The true value is for the limit $T \to \infty$. Because of the assumed property of ergodicity and stationarity, the above limit converges quickly with increasing T, and if you would get the same result if one were to shift the integration domain by an arbitrary amount. In other words: the statistical properties of an ergodic signal can be found by studying any finite, but sufficiently long, sample of the signal. Let us, from now on, assume that all our signals are ergodic, which is a good assumption for our purposes. Hence the normalized convolution and correlation can be defined as

$$(g * h)(t) = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{+T/2} g(t-\tau)h(\tau)d\tau \tag{B.39}$$

$$\mathrm{Corr}[g,h](t) = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{+T/2} g(t+\tau)h^*(\tau)d\tau \tag{B.40}$$

and their values can be well approximated by simply taking T large, not infinite.

We now define the autocorrelation function

$$B_g(t) = \text{Corr}[g,g](t) = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{+T/2} g(t+\tau)g^*(\tau)d\tau \tag{B.41}$$

and the structure function

$$D_g(t) = \langle |g(t+\tau) - g(\tau)|^2\rangle_\tau \equiv \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{+T/2} |g(t+\tau) - g(\tau)|^2 d\tau \tag{B.42}$$

These definitions are directly compatible with those from Section B.2. They do not scale with T. Instead, they become more accurate as T increases.

But many properties are the same as in Section B.1, for instance:

$$D_g(t) = 2B_g(0) - B_g(t) - B_g^*(t) \tag{B.43}$$

and again $g(t)$ is a real function, then we obtain

$$D_g(t) = 2[B_g(0) - B_g(t)] \tag{B.44}$$

Applying Fourier analysis is a bit more tricky. We have to use the technique to deal with Fourier transforms on finite domains, discussed in Section A.3. Much remains analogous, but we will not dig deeper into the details here.

B.4 Gaussian signals

Suppose a signal $g(t)$ has a Gaussian probability distribution function with zero mean and variance σ^2:

$$P_g(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{B.45}$$

We can then derive a useful property of Gaussian signals:

$$\begin{aligned}
\langle e^{\alpha x}\rangle &= \int_{-\infty}^{+\infty} e^{\alpha x}P(x)dx && \text{(B.46)}\\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\exp\left(\alpha x - \frac{x^2}{2\sigma^2}\right)dx && \text{(B.47)}\\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\exp\left[-\left(\frac{x^2 - 2\alpha x\sigma^2}{2\sigma^2}\right)\right]dx && \text{(B.48)}\\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\exp\left[-\frac{(x-\alpha\sigma^2)^2}{2\sigma^2}\right]\exp\left(\frac{1}{2}\alpha^2\sigma^2\right)dx && \text{(B.49)}\\
&= \exp\left(\frac{1}{2}\alpha^2\sigma^2\right)\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\exp\left[-\frac{\tilde{x}^2}{2\sigma^2}\right]d\tilde{x} && \text{(B.50)}\\
&= \exp\left(\frac{1}{2}\alpha^2\langle x^2\rangle\right) && \text{(B.51)}
\end{aligned}$$

This holds also for complex α. In the particular case of imaginary α we get a Gaussan expectation value again.