Fourier Analysis of Stochastic Processes

1. Time series

Given a discrete time process $(X_n)_{n\in\mathbb{Z}}$, with $X_n : \Omega \longrightarrow \mathbb{R}$ or $X_n : \Omega \longrightarrow \mathbb{C}$ $\forall n \in \mathbb{Z}$, we define time series a realization of the process, that is to say a series $(x_n)_{n\in\mathbb{Z}}$ of real or complex numbers where $x_n = X_n(\omega) \forall n \in \mathbb{Z}$. In some cases we will use the notation $x(n)$ instead of x_n, with the same meaning. For convenience, we introduce the space l_2 of the time series such that $\sum_{n\in\mathbb{Z}} |x_n|^2 < \infty$. The finite value of the sum can be sometimes interpreted as a form of the energy, and the time series belonging to l_2 are called finite energy time series. Another important set is the space l_1 of the time series such that $\sum_{n\in\mathbb{Z}} |x_n| < \infty$. Notice that the assumption $\sum_{n\in\mathbb{Z}} |x_n| < \infty$ implies $\sum_{n\in\mathbb{Z}} |x_n|^2 < \infty$, because $\sum_{n\in\mathbb{Z}} |x_n|^2 \leq \sup_{n\in\mathbb{Z}} |x_n| \sum_{n\in\mathbb{Z}} |x_n|$ and $\sup_{n\in\mathbb{Z}} |x_n|$ is bounded when $\sum_{n\in\mathbb{Z}} |x_n|$ converges.

Given two time series $f(n)$ and $g(n)$, we define the convolution of the two time series as

$$h(n) = (f * g)(n) = \sum_{k\in\mathbb{Z}} f(n-k)\, g(k)$$

2. Discrete time Fourier transform

Given the realization of a process $(x_n)_{n\in\mathbb{Z}} \in l_2$, we introduce the *discrete time Fourier transform* (DTFT), indicated either by the notation $\widehat{x}(\omega)$ or $\mathcal{F}[x](\omega)$ and defined by

$$\widehat{x}(\omega) = \mathcal{F}[x](\omega) = \frac{1}{\sqrt{2\pi}} \sum_{n\in\mathbb{Z}} e^{-i\omega n} x_n, \qquad \omega \in [0, 2\pi] \ .$$

Note that the symbol $\widehat{\ }$ is used to indicate both DTFT and empirical estimates of process parameters; the variable ω is used to indicate both the independent variable of the DTFT and the outcomes (elementary events) in a set of events. Which of the two meanings is the correct one will be always evident in both cases.

The sequence x_n can be reconstructed from its DTFT by means of the inverse Fourier transform

$$x_n = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} e^{i\omega n} \widehat{x}(\omega)\, d\omega.$$

In fact, assuming that it is allowed to interchange the infinite summation with the integration, we have

$$\frac{1}{\sqrt{2\pi}} \int_0^{2\pi} e^{i\omega n} \widehat{x}(\omega)\, d\omega = \frac{1}{2\pi} \int_0^{2\pi} e^{i\omega n} \sum_{k\in\mathbb{Z}} x_k e^{-i\omega k} d\omega = \frac{1}{2\pi} \sum_{k\in\mathbb{Z}} x_k \int_0^{2\pi} e^{i\omega(n-k)} d\omega = \frac{1}{2\pi} \sum_{k\in\mathbb{Z}} x_k 2\pi\delta(n-k) = x_n$$

Remark 1. *In the definition using $-i\omega n$ as exponent, the increment of the independent variable between two consecutive samples is assumed to be 1. If the independent variable is time t and the*

time interval between two consecutive samples is Δt, in order to change the time scale, the exponent becomes $-i\omega n\Delta t$. The quantity $\dfrac{1}{\Delta t}$ is called sampling frequency.

REMARK 2. *The function $\widehat{x}(\omega)$ can be considered for all $\omega \in \mathbb{R}$, but it is 2π-periodic or $\dfrac{2\pi}{\Delta t}$-periodic. In the last case, ω is called angular frequency.*

REMARK 3. *Sometimes (in particular by physicists), DTFT is defined without the $-$ sign at the exponent; in this case the $-$ sign is obviously present in the inverse transform.*

REMARK 4. *Sometimes the factor $\frac{1}{\sqrt{2\pi}}$ is not included in the definition; in our definition the factor $\frac{1}{\sqrt{2\pi}}$ is included for symmetry with the inverse transform or the Plancherel formula (without $\frac{1}{\sqrt{2\pi}}$, a factor $\frac{1}{2\pi}$ appears in one of them).*

REMARK 5. *Sometimes, it is preferable to use the variant*

$$\widehat{x}(f) = \frac{1}{\sqrt{2\pi}} \sum_{n\in\mathbb{Z}} e^{-2\pi i f n} x_n, \qquad f \in [0,1].$$

where $f = \dfrac{\omega}{2\pi}$. When the factor Δt is present in the exponent, $f \in \left[0, \dfrac{1}{\Delta t}\right]$ is called frequency.

In the following, we will use our definition of DTFT, independently of the fact that in certain applications it is customary or convenient to make other choices. The reader is warned!!

Of course, in practical cases, infinite sampling times do not exist. Let us therefore introduce a sampling window of width $2N$ contaning $2N+1$ samples from time $-N$ to time N (also called boxcar window) $1_{[-N,N]}(n) = \begin{cases} 1 \text{ if } -N \le n \le N \\ 0 \text{ otherwise} \end{cases}$ and the truncation $x_{2N}(n)$ of the time series x_n defined as $x_{2N}(n) = x_n \cdot 1_{[-N,N]}(n) = \begin{cases} x_n \text{ if } -N \le n \le N \\ 0 \text{ otherwise} \end{cases}$. The DTFT of the truncated time series is defined as

$$\widehat{x}_{2N}(\omega) = \frac{1}{\sqrt{2\pi}} \sum_{|n|\le N} e^{-i\omega n} x_n.$$

REMARK 6. *In order to define properly the DTFT for a process, we must assure that it is possible to calculate the DTFT of "every" of its realization. Therefore the above definition, involving only a finite summation, can be used also to define the truncated DTFT of a process:*

$$\widehat{X}_{2N}(\omega) = \frac{1}{\sqrt{2\pi}} \sum_{|n|\le N} e^{-i\omega n} X_n.$$

In fact, each realization $\widehat{x}_{2N}$ of a truncated process $\widehat{X}_{2N}$ is a finite-energy time series that admits DTFT. We underline that $\widehat{X}_{2N}(\omega)$ is a random variable for every $\omega \in [0, 2\pi]$, while $\widehat{x}_{2N}(\omega)$ is a function of ω. The definition of truncated DTFT for a process will be used in the proof of Wiener-Khinchine theorem. Processes whose realizations are l_2 or l_1 time series are never stationary.

Properties of DTFT

1) The L^2-theory of Fourier series guarantees that the series $\sum_{n\in\mathbb{Z}} e^{-i\omega n}x_n$ converges in mean square with respect to ω, namely, there exists a square integrable function $\widehat{x}(\omega)$ such that

$$\lim_{N\to\infty}\int_0^{2\pi}\left|\widehat{x}_{2N}(\omega)-\widehat{x}(\omega)\right|^2 d\omega = 0.$$

2) Plancherel formula

$$\sum_{n\in\mathbb{Z}}|x_n|^2 = \int_0^{2\pi}\left|\widehat{x}(\omega)\right|^2 d\omega$$

PROOF. $\int_0^{2\pi}\left|\widehat{x}(\omega)\right|^2 d\omega = \int_0^{2\pi}\widehat{x}(\omega)\left(\widehat{x}(\omega)\right)^* d\omega = \frac{1}{2\pi}\int_0^{2\pi}\left(\sum_{n\in\mathbb{Z}}e^{-i\omega n}x_n\right)\left(\sum_{m\in\mathbb{Z}}e^{-i\omega m}x_m\right)^* d\omega = \frac{1}{2\pi}\int_0^{2\pi}\left(\sum_{n,m\in\mathbb{Z}}x_n x_m^* e^{-i\omega n}e^{i\omega m}\right)d\omega = \frac{1}{2\pi}\sum_{n,m\in\mathbb{Z}}x_n x_m^*\int_0^{2\pi}e^{-i\omega(n-m)}d\omega = \frac{1}{2\pi}\sum_{n,m\in\mathbb{Z}}x_n x_m^* 2\pi\delta(n-m) = \sum_{n\in\mathbb{Z}}x_n x_n^* = \sum_{n\in\mathbb{Z}}|x_n|^2$ assuming that it is correct to interchange the infinite summation with the integration. □

The meaning of Plancherel formula is that the energy "contained" in a time series and the energy "contained" in its DTFT is the same (a factor 2π can appear in one of the two members when a differnt detinition of DTFT is used).

3) If the time series $g(n)$ is real, then $\widehat{g}(-\omega) = \widehat{g}^*(\omega)$.

PROOF. $\widehat{g}(-\omega) = \frac{1}{\sqrt{2\pi}}\sum_{n\in\mathbb{Z}}g(n)e^{-i(-\omega)n} = \frac{1}{\sqrt{2\pi}}\sum_{n\in\mathbb{Z}}g(n)\left(e^{-i\omega n}\right)^* = \left(\frac{1}{\sqrt{2\pi}}\sum_{n\in\mathbb{Z}}g(n)e^{-i\omega n}\right)^* = g^*(\omega)$ □

4) The DTFT of the convolution of two time series corresponds (a factor can be present,depending on the definition of DTFT) to the product of the DTFT of the two time series

$$\mathcal{F}[f*g](\omega) = \sqrt{2\pi}\widehat{f}(\omega)\widehat{g}(\omega).$$

PROOF. $\mathcal{F}[f*g](\omega) = \frac{1}{\sqrt{2\pi}}\sum_{n\in\mathbb{Z}}(f*g)(n)e^{-i\omega n} = \frac{1}{\sqrt{2\pi}}\sum_{n\in\mathbb{Z}}\left(\sum_{k\in\mathbb{Z}}f(n-k)g(k)\right)e^{-i\omega n} = \frac{1}{\sqrt{2\pi}}\sum_{k\in\mathbb{Z}}g(k)e^{-i\omega k}\sum_{n\in\mathbb{Z}}f(n-k)e^{-i\omega(n-k)} = \frac{1}{\sqrt{2\pi}}\sum_{k\in\mathbb{Z}}g(k)e^{-i\omega k}\sum_{m\in\mathbb{Z}}f(m)e^{-i\omega m} = \sum_{k\in\mathbb{Z}}g(k)e^{-i\omega k}\widehat{f}(\omega) = \sqrt{2\pi}\widehat{f}(\omega)\widehat{g}(\omega)$, assuming that it is correct to commute the two infinite summations. The $\sqrt{2\pi}$ is not present when the definition of DTFT without $\frac{1}{\sqrt{2\pi}}$ is used. □

6) Combining properties 4) and 5), we obtain

$\mathcal{F}\left[\sum_{k\in\mathbb{Z}}f(n+k)g(k)\right](\omega) = \mathcal{F}\left[\sum_{k\in\mathbb{Z}}f(n-k)g(-k)\right](\omega) = \widehat{f}(\omega)\widehat{g}(-\omega) = \sqrt{2\pi}\widehat{f}(\omega)\widehat{g}^*(\omega)$, where

the last equality is valid only for a real time series g. This property is used in the calculation of DTFT of correlation function. Again, the presence of factor $\sqrt{2\pi}$ depends on the definition of DTFT adopted.

7) When $\sum_{n\in\mathbb{Z}} |x_n| < \infty$, then the series $\sum_{n\in\mathbb{Z}} e^{-i\omega n} x_n$ is absolutely convergent, uniformly in $\omega \in [0, 2\pi]$, simply because

$$\sum_{n\in\mathbb{Z}} \sup_{\omega\in[0,2\pi]} \left|e^{-i\omega n} x_n\right| = \sum_{n\in\mathbb{Z}} \sup_{\omega\in[0,2\pi]} \left|e^{-i\omega n}\right| |x_n| = \sum_{n\in\mathbb{Z}} |x_n| < \infty.$$

In this case, we may also say that $\widehat{x}(\omega)$ is a bounded continuous function, not only square integrable.

3. Generalized discrete time Fourier transform

One can do the DTFT also for sequences which do not satisfy the assumption $\sum_{n\in\mathbb{Z}} |x_n|^2 < \infty$, in special cases. Consider for instance the sequence

$$x_n = a \sin(\omega_1 n).$$

Compute the DFTT of the truncated sequence

$$\widehat{x}_{2N}(\omega) = \frac{1}{\sqrt{2\pi}} \sum_{|n|\le N} e^{-i\omega n} a \sin(\omega_1 n).$$

Recall that

$$\sin t = \frac{e^{it} - e^{-it}}{2i}.$$

Hence $\sin(\omega_1 n) = \frac{e^{i\omega_1 n} - e^{-i\omega_1 n}}{2i}$

$$\sum_{|n|\le N} e^{-i\omega n} a \sin(\omega_1 n) = \frac{1}{2i} \sum_{|n|\le N} e^{-i(\omega-\omega_1)n} - \frac{1}{2i} \sum_{|n|\le N} e^{-i(\omega+\omega_1)n}.$$

The next lemma makes use of the concept of generalized function or distribution, which is outside the scope of these notes. We still given the result, to be understood in some intuitive sense. We use the generalized function $\delta(t)$ called Dirac delta (not to be confused with the dircrete time δ defined in example 1 for white noise), which is characterized by the property

$$\int_{-\infty}^{\infty} \delta(t - t_0) f(t)\, dt = f(t_0) \tag{3.1}$$

for all continuous compact support functions f. No usual function has this property. A way to get intuition is the following one. Consider a function $\delta_n(t)$ which is equal to zero for t outside $\left[-\frac{1}{2n}, \frac{1}{2n}\right]$, interval of length $\frac{1}{n}$ around the origin; and equal to n in $\left[-\frac{1}{2n}, \frac{1}{2n}\right]$. Hence $\delta_n(t - t_0)$ is equal to zero for t outside $\left[t_0 - \frac{1}{2n}, t_0 + \frac{1}{2n}\right]$, equal to n in $\left[t_0 - \frac{1}{2n}, t_0 + \frac{1}{2n}\right]$. We have

$$\int_{-\infty}^{\infty} \delta_n(t)\, dt = 1.$$

Now,

$$\int_{-\infty}^{\infty} \delta_n (t - t_0) f(t) \, dt = n \int_{t_0 - \frac{1}{2n}}^{t_0 + \frac{1}{2n}} f(t) \, dt$$

which is the average of f around t_0. As $n \to \infty$, this average converges to $f(t_0)$ when f is continuous. Namely. we have

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} \delta_n (t - t_0) f(t) \, dt = f(t_0)$$

which is the analog of identity (3.1), but expressed by means of traditional concepts. In a sense, thus, the generalized function $\delta(t)$ is the limit of the traditional functions $\delta_n(t)$. But we see that $\delta_n(t)$ converges to zero for all $t \neq 0$, and to ∞ for $t = 0$. So, in a sense, $\delta(t)$ is equal to zero for $t \neq 0$, and to ∞ for $t = 0$; but this is a very poor information, because it does not allow to deduce identity (3.1) (the way $\delta_n(t)$ goes to infinity is essential, not only the fact that $\delta(t)$ is ∞ for $t = 0$).

Lemma 1. *Denote by $\delta(t)$ the generalized function such that*

$$\int_{-\infty}^{\infty} \delta (t - t_0) f(t) \, dt = f(t_0)$$

for all continuous compact support functions f (it is called the delta Dirac distribution). Then

$$\lim_{N \to \infty} \sum_{|n| \leq N} e^{-itn} = 2\pi \delta(t).$$

From this lemma it follows that

$$\lim_{N \to \infty} \sum_{|n| \leq N} e^{-i\omega n} a \sin(\omega_1 n) = \frac{\pi}{i} \delta(\omega - \omega_1) - \frac{\pi}{i} \delta(\omega + \omega_1).$$

In other words,

Corollary 1. *The sequence*

$$x_n = a \sin(\omega_1 n)$$

has a generalized DTFT

$$\widehat{x}(\omega) = \lim_{N \to \infty} \widehat{x}_{2N}(\omega) = \frac{\sqrt{\pi}}{\sqrt{2} i} (\delta(\omega - \omega_1) - \delta(\omega + \omega_1)).$$

This is only one example of the possibility to extend the definition and meaning of DTFT outside the assumption $\sum_{n \in \mathbb{Z}} |x_n|^2 < \infty$. It is also very interesting for the interpretation of the concept of DTFT. If the signal x_n has a periodic component (notice that DTFT is linear) with angular frequency ω_1, then its DTFT has two symmetric peaks (delta Dirac components) at $\pm\omega_1$. This way, the DTFT reveals the periodic components of the signal.

Exercise 1. *Prove that the sequence*

$$x_n = a \cos(\omega_1 n)$$

has a generalized DTFT

$$\widehat{x}(\omega) = \lim_{N \to \infty} \widehat{x}_{2N}(\omega) = \frac{\sqrt{\pi}}{\sqrt{2}} (\delta(\omega - \omega_1) + \delta(\omega + \omega_1)).$$

4. Power spectral density

Given a stationary process $(X_n)_{n\in\mathbb{Z}}$ with correlation function $R(n) = E[X_n X_0]$, $n \in \mathbb{Z}$, we call *power spectral density* (*PSD*) the function

$$S(\omega) = \frac{1}{\sqrt{2\pi}} \sum_{n\in\mathbb{Z}} e^{-i\omega n} R(n), \qquad \omega \in [0, 2\pi].$$

Alternatively, one can use the expression

$$S(f) = \frac{1}{\sqrt{2\pi}} \sum_{n\in\mathbb{Z}} e^{-2\pi i f n} R(n), \qquad f \in [0, 1]$$

which produces easier visualizations because we catch more easily the fractions of the interval $[0, 1]$.

Remark 7. *In principle, to define properly the PSD, the condition* $\sum_{n\in\mathbb{Z}} |R(n)| < \infty$ *should be required, or at least* $\sum_{n\in\mathbb{Z}} |R(n)|^2 < \infty$. *In practice, on a side the convergence may happen also in unexpected cases due to cancellations, on the other side it may be acceptable to use a finite-time variant, something like* $\sum_{|n|\leq N} e^{-i\omega n} R(n)$, *for practical purposes or from the computational viewpoint.*

A priori, one may think that $S(f)$ may be not real valued. However, the function $R(n)$ is *non-negative definite* (this means $\sum_{i=1}^{n} R(t_i - t_j) a_i a_j \geq 0$ for all $t_1, ..., t_n$ and $a_1, ..., a_n$) and a theorem states that the Fourier transform of non-negative definite function is a non-negative function. Thus, at the end, it turns out that $S(f)$ is real and also non-negative. We do not give the details of this fact here because it will be a consequence of the fundamental theorem below.

4.1. Example: white noise. We have

$$R(n) = \sigma^2 \cdot \delta(n)$$

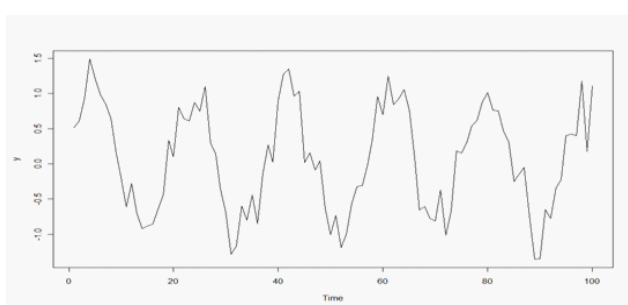
hence

$$S(\omega) = \frac{\sigma^2}{\sqrt{2\pi}}, \qquad \omega \in \mathbb{R}.$$

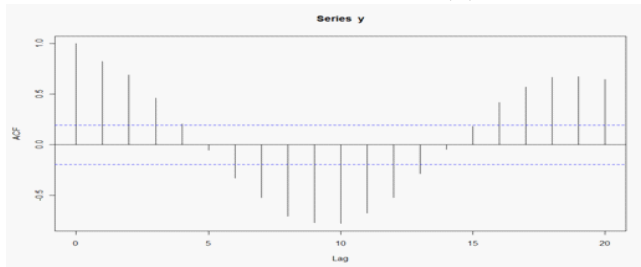
The spectra density is constant. This is the origin of the name, *white noise*.

4.2. Example: perturbed periodic time series. This example is numeric only. Produce with R software the following time series:

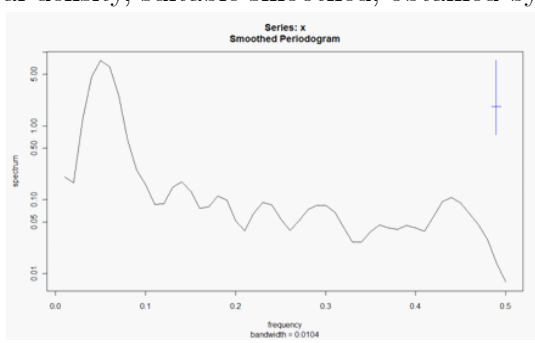
```
t <- 1:100
y<- sin(t/3)+0.3*rnorm(100)
ts.plot(y)
```



The empirical autocorrelation function, obtained by acf(y), is



and the power spectral density, suitable smoothed, obtained by spectrum(y,span=c(2,3)), is



4.3. Pink, Brown, Blue, Violet noise. In certain applications one meets PSD of special type which have been given names similarly to white noise. Recall that white noise has a constant PSD. Pink noise has PSD of the form

$$S(f) \sim \frac{1}{f}.$$

Brown noise:

$$S(f) \sim \frac{1}{f^2}.$$

Blue noise

$$S(f) \sim f \cdot 1_\Lambda.$$

Violet noise

$$S(f) \sim f^2 \cdot 1_\Lambda.$$

where, chosen Λ such that $0 < \Lambda < 1$, $1_\Lambda(f) = \begin{cases} 1 & \text{if} \quad 0 \le f \le \Lambda \\ 0 & \text{otherwise} \end{cases}$

5. A fundamental theorem on PSD

The following theorem is often stated without assumptions in the applied literature. One of the reasons is that it can be proved under various level of generality, with different meanings of the limit operation (it is a limit of functions). We shall give a rigorous statement under a very precise assumption on the autocorrelation function $R(n)$; the convergence we prove is rather strong. The assumption (a little bit strange, but satisfied in all our examples) is

$$\sum_{n\in\mathbb{N}} R(n)^p < \infty \text{ for some } p \in (0,1). \tag{5.1}$$

This is just a little bit more restrictive than the condition $\sum_{n\in\mathbb{N}} |R(n)| < \infty$ which is natural to impose if we want uniform convergence of $\frac{1}{\sqrt{2\pi}}\sum_{n\in\mathbb{Z}} e^{-i\omega n} R(n)$ to $S(\omega)$.

THEOREM 1 (Wiener-Khinchin). *If $(X(n))_{n\in\mathbb{Z}}$ is a wide-sense stationary process satisfying assumption (5.1), then*

$$S(\omega) = \lim_{N\to\infty} \frac{1}{2N+1} E\left[\left|\widehat{X}_{2N}(\omega)\right|^2\right].$$

The limit is uniform for $\omega \in [0, 2\pi]$. Here X_{2N} is the truncated process $X \cdot 1_{[-N,N]}$. In particular, it follows that $S(\omega)$ is real an non-negative.

PROOF. **Step 1**. Let us prove the following main identity:

$$S(\omega) = \frac{1}{2N+1} E\left[\left|\widehat{X}_{2N}(\omega)\right|^2\right] + r_N(\omega) \tag{5.2}$$

where the remainder r_N is given by

$$r_N(\omega) = \frac{1}{2N+1}\mathcal{F}\left[\sum_{n\in\Delta(N,t)} E\left[X(t+n)X(n)\right]\right](\omega)$$

and $\Delta(N,t)$ is a subinterval of $[-N,N]$ to be defined later. Since $R(t) = E[X(t+n)X(n)]$ for all n, we obviously have, for every t,

$$R(t) = \frac{1}{2N+1}\sum_{|n|\le N} E\left[X(t+n)X(n)\right].$$

Thus

$$S(\omega) = \widehat{R}(\omega) = \frac{1}{2N+1}\mathcal{F}\left[\sum_{|n|\le N} E\left[X(t+n)X(n)\right]\right](\omega) = \frac{1}{2N+1}\sum_{|n|\le N}\sum_{t\in\mathbb{Z}}\left[E\left[X(t+n)X(n)\right]\right]e^{-i\omega t} \tag{5.3}$$

To invert the infinite sum over t with the calculation of mean value, we need to introduce truncated process $X_{2N}(n) = X(n)\cdot 1_{[-N,N]}(n)$, for which we have, for $0 \le t \le 2N$

$$\sum_{n\in\mathbb{Z}} X_{2N}(t+n)X_{2N}(n) = \sum_{n=-N}^{N-t} X(t+n)X(n) \ ;$$

for $-2N \leq t < 0$ ve have

$$\sum_{n\in\mathbb{Z}} X_{2N}(t+n)\, X_{2N}(n) = \sum_{n=-N-t}^{N} X(t+n)\, X(n) \ ;$$

for $t > 2N$ or $t < -2N$, we have $X_{2N}(t+n)\, X_{2N}(n) = 0 \forall n$. In general,

$$\sum_{n\in\mathbb{Z}} X_{2N}(t+n)\, X_{2N}(n) = \sum_{n\in[N_t^-, N_t^+]} X(t+n)\, X(n)\,.$$

with

$$[N_t^-, N_t^+] = \begin{cases} \emptyset & \text{if} \quad t < -2N \\ [-N-t, N] & \text{if} \quad -2N \leq t < 0 \\ [-N, N-t] & \text{if} \quad 0 \leq t \leq 2N \\ \emptyset & \text{if} \quad t > 2N \end{cases}$$

Therefore, remembering property 6) of DTFT,

$$\mathcal{F}\left[\sum_{n=N_t^-}^{N_t^+} X(t+n)\, X(n)\right](\omega) = \mathcal{F}\left[\sum_{n\in\mathbb{Z}} X_{2N}(t+n)\, X_{2N}(n)\right](\omega) = \widehat{X}_{2N}(\omega)\, \widehat{X}_{2N}^*(\omega) = \left|\widehat{X}_{2N}(\omega)\right|^2.$$

Taking the mean value of first and last term, we obtain

$$\mathcal{F}\left[\sum_{n=N_t^-}^{N_t^+} E\left[X(t+n)\, X(n)\right]\right](\omega) = E\left[\mathcal{F}\left[\sum_{n=N_t^-}^{N_t^+} X(t+n)\, X(n)\right](\omega)\right] = E\left[\left|\widehat{X}_{2N}(\omega)\right|^2\right].$$

Now, it is possible to decompose the sum running from $-N$ to $+N$ into a sum running from N_t^- to N_t^+ and a remainder r_N containing the other terms, from $-N$ to N_t^- or from N_t^+ to N.

$$S(\omega) = \frac{1}{2N+1}\mathcal{F}\left[\sum_{|n|\leq N} E\left[X(t+n)\, X(n)\right]\right](\omega) = \frac{1}{2N+1}\mathcal{F}\left[\sum_{n=N_t^-}^{N_t^+} E\left[X(t+n)\, X(n)\right]\right](\omega) +$$

$$\frac{1}{2N+1}\mathcal{F}\left[\sum_{n\in\Delta(N,t)} E\left[X(t+n)\, X(n)\right]\right](\omega) = \frac{1}{2N+1}E\left[\left|\widehat{X}_{2N}(\omega)\right|^2\right] + r_N(\omega) \ , \text{ with}$$

$$\Delta(N,t) = \begin{cases} [-N, N] & \text{if} \quad t < -2N \\ [-N, -N-t-1] & \text{if} \quad -2N \leq t < 0 \\ \emptyset & \text{if} \quad t = 0 \\ [N-t+1, N] & \text{if} \quad 0 < t \leq 2N \\ [-N, N] & \text{if} \quad t > 2N \end{cases}$$

Step 2. The proof is complete if we show that $\lim_{N\to\infty} r_N(\omega) = 0$ uniformly in $\omega \in [0, 2\pi]$. Let $\varepsilon_n = R(n)^{1-p}$. As $\sum_{n\in\mathbb{N}} R(n)^p < \infty$, we have $R(n)^p \to 0$, and therefore also $\varepsilon_n \to 0$. Moreover,

$$\sum_{n\in\mathbb{N}} \frac{R(n)}{\varepsilon_n} = \sum_{n\in\mathbb{N}} R(n)^p < \infty.$$

We can write

$$\sum_{n\in\Delta(N,t)} E[X(t+n)X(n)] = \sum_{n\in\Delta(N,t)} R(t) = \varepsilon_{|t|}\frac{R(t)}{\varepsilon_{|t|}}|\Delta(N,t)|$$

where $|\Delta(N,t)|$ denotes the cardinality of $\Delta(N,t)$. If $(2N+1)\wedge|t|$ denotes the smallest value between $(2N+1)$ and $|t|$, we have

$$|\Delta(N,t)| = (2N+1)\wedge|t|$$

hence

$$\frac{1}{2N+1}\left|\sum_{n\in\Delta(N,t)} E[X(t+n)X(n)]\right| = \frac{|R(t)|}{\varepsilon_{|t|}}\frac{((2N+1)\wedge|t|)\,\varepsilon_{|t|}}{2N+1}.$$

Given $\delta > 0$, let t_0 be such that $\varepsilon_{|t|} \le \delta$ for all $t \ge t_0$. Then take $N_0 \ge t_0$ such that $\frac{t_0}{2N+1} \le \delta$ for all $N \ge N_0$. It is not restrictive to assume $\varepsilon_{|t|} \le 1$ for all t. Then, for $N \ge N_0$, if $t \le t_0$ then

$$\frac{((2N+1)\wedge|t|)\,\varepsilon_{|t|}}{2N+1} \le \frac{t_0\varepsilon_{|t|}}{2N+1} \le \frac{t_0}{2N+1} \le \delta$$

and if $t \ge t_0$ then

$$\frac{((2N+1)\wedge|t|)\,\varepsilon_{|t|}}{2N+1} \le \frac{((2N+1)\wedge|t|)}{2N+1}\delta \le \delta.$$

We have proved the following statement: for all $\delta > 0$ there exists N_0 such that

$$\frac{((2N+1)\wedge|t|)\,\varepsilon_{|t|}}{2N+1} \le \delta$$

for all $N \ge N_0$, uniformly in t. Then also

$$\frac{1}{2N+1}\left|\sum_{n\in\Delta(N,t)} E[X(t+n)X(n)]\right| \le \frac{|R(t)|}{\varepsilon_{|t|}}\delta$$

for all $N \ge N_0$, uniformly in t. Therefore

$$|r_N(\omega)| = \left|\frac{1}{2N+1}\frac{1}{\sqrt{2\pi}}\sum_{t\in\mathbb{Z}} e^{-i\omega t}\left[\sum_{n\in\Delta(N,t)} E[X(t+n)X(n)]\right]\right|$$

$$\le \frac{1}{2N+1}\frac{1}{\sqrt{2\pi}}\sum_{t\in\mathbb{Z}}\left|\sum_{n\in\Delta(N,t)} E[X(t+n)X(n)]\right| \le \frac{1}{\sqrt{2\pi}}\sum_{t\in\mathbb{Z}}\frac{|R(t)|}{\varepsilon_{|t|}}\delta = \frac{C}{\sqrt{2\pi}}\delta$$

where $C = \sum_{t\in\mathbb{Z}}\frac{|R(t)|}{\varepsilon_{|t|}} < \infty$. This is the definition of $\lim_{N\to\infty} r_N(\omega) = 0$ uniformly in $\omega \in [0, 2\pi]$. The proof is complete. □

This theorem gives us the interpretation of PSD. The Fourier transform $\widehat{X}_T(\omega)$ identifies the frequency structure of the signal. The square $\left|\widehat{X}_T(\omega)\right|^2$ drops the information about the phase and keeps the information about the amplitude, but in the sense of energy (a square). It gives us the energy spectrum, in a sense. So the PSD is the average amplitude of the oscillatory component at frequency $f = \frac{\omega}{2\pi}$.

Thus PSD is a very useful tool if you want to identify oscillatory signals in your time series data and want to know their amplitude. By PSD, one can get a "feel" of data at an early stage of time series analysis. PSD tells us at which frequency ranges variations are strong.

Remark 8. *A priori one could think that it were more natural to compute the Fourier transform* $\widehat{X}(\omega) = \sum_{n\in\mathbb{Z}} e^{i\omega n} X_n$ *without a cut-off of size* $2N$. *But the process* (X_n) *is stationary. Therefore, it does not satisfy the assumption* $\sum_{n\in\mathbb{Z}} X_n^2 < \infty$ *or similar ones which require a decay at infinity. Stationarity is in contradiction with a decay at infinity (it can be proved, but we leave it at the obvious intuitive level).*

Remark 9. *Under more assumptions (in particular a strong ergodicity one) it is possible to prove that*

$$S(\omega) = \lim_{N\to\infty} \frac{1}{2N+1} \left|\widehat{X}_{2N}(\omega)\right|^2$$

without expectation. Notice that $\frac{1}{2N+1}\left|\widehat{X}_{2N}(\omega)\right|^2$ *is a random quantity, but the limit is deterministic.*