Annals of Physics 383 (2017) 285–311

Contents lists available at ScienceDirect

Annals of Physics

journal homepage: www.elsevier.com/locate/aop

Asymptotic behaviour of time averages for non-ergodic Gaussian processes

Jakub Ślęzak

Wroclaw University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

HIGHLIGHTS

- Ergodic criteria for Gaussian models which use Fourier transform are provided.
- Smooth, field Hamiltonian models are ergodic, discrete models are non-ergodic.
- Fractal Fourier structure can induce recurring correlations and non-mixing.
- Non-ergodic models exhibit non-linear dynamics and complex constants of motion.
- This non-linearity can be studied using time-averaged characteristic function.

ARTICLE INFO

Article history:
Received 10 February 2017
Accepted 16 May 2017
Available online 31 May 2017

Keywords:
Ergodicity breaking
Gaussian process
Statistical analysis
Generalised Langevin equation

ABSTRACT

In this work, we study the behaviour of time-averages for stationary (non-ageing), but ergodicity-breaking Gaussian processes using their representation in Fourier space. We provide explicit formulae for various time-averaged quantities, such as mean square displacement, density, and analyse the behaviour of time-averaged characteristic function, which gives insight into rich memory structure of the studied processes. Moreover, we show applications of the ergodic criteria in Fourier space, determining the ergodicity of the generalised Langevin equation's solutions.

1. Introduction

1.1. The goal

The relation between the time averages and ensemble averages is one of the most important topics of statistical physics and this area of research is under intense development. The abstract,

E-mail address: jakub.slezak@pwr.edu.pl.

http://dx.doi.org/10.1016/j.aop.2017.05.015

mathematical ergodic theory has become a very wide subject, but in the recent years a new trend has emerged which concentrates on very practical questions.

One example is the behaviour of the time-averaged mean square displacement

$$\overline{\delta^2}(\Delta) := \frac{1}{T-\Delta}\int_0^{T-\Delta} \mathrm{d}\tau \, \big(X(\tau+\Delta) - X(\tau)\big)^2. \tag{1}$$

This quantity can be estimated using only one (sufficiently long) trajectory in contrast to the ensemble-averaged mean square displacement

$$\delta^2(\Delta) := \int \mathrm{dP} \, \big(X(t+\Delta) - X(t)\big)^2, \tag{2}$$

which requires many trajectories to estimate. The comparison of these two types of mean square displacement is central in study of the weak ergodicity breaking and can be used to distinguish between different models of classical and anomalous dynamics [1,2].

In this work we discuss the middle-ground between the abstract ergodic theory and the notions used in applications, concentrating on three main areas:

I. We characterise the behaviour of time-averages for non-ergodic processes, including explicit formulae for useful quantities such as the time-averaged mean square displacement (1) and time-averaged density (Section 3.2).
II. We analyse the behaviour of the time-averaged characteristic function, which is a less known statistics providing deeper insight into the dependence structure of the studied processes, unobtainable using only covariance-based methods (Section 3.2).
III. We determine the ergodicity and mixing of various physical models, the most important being generalised and classical Langevin equation (Sections 3.1 and 4.2).

Together, these results form a basis of methodology for statistical analysis of ergodicity-breaking processes, which appear in the well-used physical models.

Such an intent is impossible to realise in full generality, so we concentrate on a class crucial from the point of view of modelling: the Gaussian processes. (However, we do remark briefly on how to treat non-Gaussian case in Section 5.) Under this assumption, one can use the elegant and practical description of the ergodic behaviour in Fourier space. Calculating the Fourier transform of a function

$$\hat{f}(\omega) := \int_{\mathbb{R}} \mathrm{d}t \, e^{i\omega t} f(t) \tag{3}$$

in order to study its properties is a very common technique. However, its relation to the ageing phenomenon (Section 2) and ergodicity (Section 3) may be surprising, and is severely underrated in the physical literature.

1.2. *Basics of ergodic theory*

We consider a continuous-time stochastic process $X = (X_t)_{t\in\mathbb{R}}$ which may be real or even complex valued. In this section the results are general and apply also to non-Gaussian processes. The variable X can be position, velocity, intensity of light, *etc*. We are interested in the behaviour of various averages of $f(X)$, where f may be in general a function of the whole trajectory X. We assume that $\mathbb{E}|f(X)| < \infty$ and call a function f with this property an observable. The expected value is an average under the probabilistic measure associated with X, $\mathbb{E}[f(X)] = \int \mathrm{dP}\, f(X)$; in physical literature the notion $\langle f(X)\rangle$ is also in use. Examples include observable of mean $f(X) = X(t)$, mean square displacement $f(X) = \big(X(t+\Delta) - X(t)\big)^2$, covariance $f(X) = X(t+\Delta)X(t)$, and others. Take note that for the above examples the observables seemingly depend on time t, however further on our assumptions will make this choice irrelevant; it will be possible to take $t = 0$ without any loss of generality.

If the process has a time-varying mean $m(t) := \mathbb{E}[X(t)] \neq$ const. we can always decompose it as a random, zero-mean part and deterministic non-zero part. Because we will study systems in which complex behaviour will be contained in the random part, we assume $m(t) = 0$; the case $m(t) \neq 0$ would be a straightforward generalisation.

For every process there exists associated family of time-shift operators $\mathcal{S}_\tau$ which describe the temporal evolution of the system, i.e. $\mathcal{S}_\tau X = (X(t+\tau))_{t\in\mathbb{R}}$.

With that in mind we can formulate the ergodic theorem. We will use general form of this result: the probabilistic variant of the Birkhoff ergodic theorem, which gives deeper insight into behaviour of both ergodic and non-ergodic processes.

Theorem 1. *If the shift operators $\mathcal{S}_\tau$ are measure-preserving, then the time average exists almost surely and is a $\mathcal{S}$-invariant random variable $\mathbb{E}[f(X)|\mathcal{C}]$ [3,4],*

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}\tau\, f\left(\mathcal{S}_\tau X\right) = \mathbb{E}[f(X)|\mathcal{C}]. \tag{4}$$

By the variable $\mathbb{E}[\cdot|\mathcal{C}]$ we understand the conditional expected value under condition $\mathcal{C}$. The condition $\mathcal{C}$ is formally the σ-algebra of time invariant sets, i.e. sets invariant under transformations $\mathcal{S}_\tau$; essentially we calculate the expected value assuming that all time invariant properties of X are fixed. The physical interpretation of $\mathcal{C}$ is that this is a set of constants of motion associated with X.

By the measure preserving transformation we mean that $\mathrm{P}(f(X) \in A) = \mathrm{P}(f(\mathcal{S}_\tau X) \in A)$ for all observables and measurable events A. This essentially means that X and time shifted $\mathcal{S}_\tau X$ are statistically indistinguishable, which is called the stationarity condition in the probabilistic literature [3], and non-ageing in many physical papers [2].

Therefore the Birkhoff ergodic theorem essentially states that if only process X is stationary, the time-average of any observable converges to a random variable $\mathbb{E}[f(X)|\mathcal{C}]$ which we can precisely determine if we can identify $\mathcal{C}$, i.e. all time invariant properties of X.

It is now clear that for the classical ergodic theorem to hold, the process should be stationary and $\mathcal{C}$ must be sufficiently weak, so that $\mathbb{E}[f(X)|\mathcal{C}] = \mathbb{E}[f(X)]$; there should be no significant time-invariant properties of X.

One immediate and important consequence of the Birkhoff theorem is that for a stationary process the ensemble average of the time average of any observable is equal to the ensemble average without any time averaging

$$\mathbb{E}\left[\lim_{T\to\infty}\frac{1}{T}\int_0^T \mathrm{d}\tau\, f\left(\mathcal{S}_\tau X\right)\right] = \mathbb{E}[f(X)]. \tag{5}$$

It is a direct consequence of the so-called tower property $\mathbb{E}\big[E[\cdot|\mathcal{C}]\big] = \mathbb{E}[\cdot]$. In particular Eq. (5) shows that for stationary processes there is no possibility for a phenomena such as the weak ergodicity breaking [2], because we must observe

$$\mathbb{E}\left[\overline{\delta^2}(\Delta)\right] = \mathbb{E}\left[\left(X(t+\Delta) - X(t)\right)^2\right]. \tag{6}$$

Another concept closely related to the ergodicity is mixing. We say that the system is mixing if for any observables f, g such that $\mathbb{E}|f(X)g(X)| < \infty$,

$$\lim_{T\to\infty}\mathbb{E}[f(X)g(\mathcal{S}_T X)] = \mathbb{E}[f(X)]\mathbb{E}[g(X)]. \tag{7}$$

This is basically a statement that the process X and the shifted process $\mathcal{S}_T X$ are asymptotically independent; the dynamics of X leaves no persisting memory. It implies ergodicity and is often easier to study than the ergodicity itself; however, there are examples of non-mixing ergodic processes, even in the Gaussian case, which we will show in Section 3.

1.3. Gaussian processes

The main part of our considerations is true only for the class of Gaussian processes. Gaussian process is a process for which any finite sum of a type $\sum_k a_k X(t_k)$ has a Gaussian distribution. It is not enough that for any t the values $X(t)$ have Gaussian distribution as it is very easy to construct counterexamples using copula theory. The sufficient and necessary condition for the process to be

Gaussian is that all $X(t)$ are Gaussian (we also admit the degenerate case $X(t) = \text{const.}$) and they are only linearly dependent [5]. The presence of non-linear dynamics excludes Gaussianity. Nevertheless, very large class of widely applied models is still Gaussian which will become apparent in the next sections.

The limitation on the possible memory type has large consequences: the Gaussian variables are fully described by their linear dependence structure which is reflected in second moments. Any Gaussian process is uniquely determined by the mean function $m(t) = \mathbb{E}[X(t)]$ (which we later assume to be 0 without loss of generality), the and covariance function $r_X(s,t) := \mathbb{E}[X(s)X(t)]$. Using these functions, a Gaussian process is stationary if and only if $m(t) = \text{const.}$ and $r_X(s,t) = r_X(t-s)$.

Another consequence of the purely linear structure of Gaussian processes is that the general mixing condition (7) reduces to the much simpler requirement that [4]

$$\lim_{T\to\infty} r_X(T) = 0, \tag{8}$$

which is often straightforward to check. The ergodicity itself can also be expressed in the language of the covariance function. Instead of ergodicity, often the equivalent notion of metric transitivity is used in this context [4,6]. The main part of this theory was completed by Maruyama in 1970 [7].

Theorem 2 (*Maruyama*)**.** *A Gaussian stationary process X is ergodic if and only if*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{d}\tau\, |r_X(\tau)| = 0. \tag{9}$$

Take note that the presence of modulus $|r_X(\tau)|$ is crucially important, because it excludes periodic oscillations of r_X. Generally this condition may also be easy to check, as it is enough to know the asymptotic tail behaviour of the covariance function. But, at the same time, it does not give much insight into the memory structure of non-ergodic Gaussian processes, which we will study later.

2. Stationarity in Fourier space

2.1. Harmonisable representation

The Fourier transform of a stationary Gaussian process X cannot be a process well-defined in classical sense as the stationary process cannot decay to zero at infinity. However, one can define the Fourier transform in the weak sense [8]

$$S(\omega) := \lim_{T\to\infty} \frac{1}{2\pi}\int_{-T}^{T} \mathrm{d}t\, \frac{\mathrm{e}^{-\mathrm{i}\omega t} - 1}{-\mathrm{i}t} X(t). \tag{10}$$

The above limit exists in the mean-square sense, moreover, the process X can be expressed as

$$X(t) = \int_{\mathbb{R}} \mathrm{d}S(\omega)\, \mathrm{e}^{\mathrm{i}\omega t}, \tag{11}$$

where the integral over increments $\mathrm{d}S(\omega)$ is also understood in the mean-square sense. Any process which can be expressed as (11) is called a weakly harmonisable process. The process S is called a spectral process. Often the values of the spectral process are taken to be complex conjugate $S(-\omega) = S(\omega)^*$ if we need the resulting X to be real valued.

If a process S has independent increments, i.e. is a scaled Brownian motion in the Fourier space, we call X harmonisable, or strongly harmonisable, and, as one can directly calculate,

$$r_X(s, s+t) = r_X(t) = \int_{\mathbb{R}} \sigma_X(\mathrm{d}\omega)\, \mathrm{e}^{\mathrm{i}\omega t} \tag{12}$$

where the measure σ_X, called the spectral measure, is defined as the mean-square amplitude of the increments of S

$$\sigma_X(\mathrm{d}\omega) := \mathbb{E}|\mathrm{d}S(\omega)|^2. \tag{13}$$

Because we have shown that r_X depends only on one parameter, any harmonisable Gaussian process is stationary. This is actually also the sufficient condition (see proof in [4]). In other words, the following theorem holds

Theorem 3. *A Gaussian process X is stationary if and only if the corresponding spectral process S has independent increments.*

The measure σ_X is a non-negative and has total mass $\sigma_X(\mathbb{R}) = \mathbb{E}|X(t)|^2 < \infty$. One might think that for important physical cases it is enough to limit ourselves to a case when σ_X is absolutely continuous, that is $\sigma_X(\mathrm{d}\omega) = \mathrm{d}\omega s(\omega)$, and has a density s called the power spectral density [8]. The next section shows that it is not true.

2.2. Harmonic processes

Consider elementary example of a motion in the harmonic potential, governed by the equation

$$\ddot{X} = -\omega_0^2 X, \quad X(0) = X_0, \dot{X}(0) = V_0, \tag{14}$$

which has the solution

$$X(t) = X_0 \cos(\omega_0 t) + \frac{V_0}{\omega_0} \sin(\omega_0 t). \tag{15}$$

The evolution of the system is purely deterministic. But, if we assume that the beginning of the evolution system interacted with the heat bath, the initial conditions X_0 and V_0 are random and have Gibbs distribution given by the density

$$\rho(x_0, v_0) \sim \exp\left(-\omega_0^2 \frac{x_0^2}{2k_B T}\right) \exp\left(-\frac{v_0^2}{2k_B T}\right). \tag{16}$$

Therefore the values $X(t)$ for any t are also random, and simple calculation shows that the resulting stochastic process has the covariance function

$$r_X(t) = \frac{k_B T}{\omega_0^2} \cos(\omega_0 t), \tag{17}$$

so it is a stationary Gaussian process. Its spectral representation is therefore given by the measure

$$\sigma_X(\mathrm{d}\omega) = \frac{k_B T}{2\omega_0^2} \big(\delta(\mathrm{d}\omega - \omega_0) + \delta(\mathrm{d}\omega + \omega_0)\big) \tag{18}$$

which is concentrated in 2 points, which we denote by two Dirac deltas. In a natural way a question about ergodicity arises. Whereas the time average of the observable of the position $f(X) = X(t)$

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{d}\tau\, X(\tau) = \lim_{T\to\infty} \frac{1}{T} \left(\frac{X_0}{\omega_0} \sin(\omega_0 T) = \frac{V_0}{\omega_0^2} (\cos(\omega_0 T) - 1)\right) = 0 \tag{19}$$

converges to the ensemble mean $0 = \mathbb{E}[X(t)]$, the observable of the mean square displacement does not, as

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{d}\tau\, \big(X(\tau + \Delta) - X(\tau)\big)^2 = 2\left(X_0^2 + \frac{V_0^2}{\omega_0^2}\right)\left(\sin\left(\frac{\omega_0 \Delta}{2}\right)\right)^2 \tag{20}$$

differs from the ensemble average

$$\mathbb{E}\left[\big(X(t + \Delta) - X(t)\big)^2\right] = 4\frac{k_B T}{\omega_0^2}\left(\sin\left(\frac{\omega_0 \Delta}{2}\right)\right)^2. \tag{21}$$

From the point of view of dynamical system theory this lack of ergodicity is expected; after initial contact with heat bath the system evolves as microcanonical ensemble and the trajectories are trapped on the surface of constant energy, which prohibits ergodicity. Indeed, the term $X_0^2 + V_0^2/\omega_0^2$

on the left side of Eq. (20) is the total energy of the system, it is random, but constant on the fixed trajectories of X. On the other hand, the factor $2k_BT/\omega_0^2$ in (21) is the mean total energy, which is ensemble averaged.

The similar reasoning applies to more general process of the form

$$X(t) = \sum_k A_k e^{i\omega_k t}, \tag{22}$$

where the sum can even be infinite if A_k are independent, complex Gaussian variables and $\sum_k \mathbb{E}|A_k|^2 < \infty$. The random functions of this class, called harmonic processes, are appearing e.g. in the phonon theory [9], where it is straightforward to recognise normal modes in sums of type (22).

The covariance function and spectral measure which correspond to a given harmonic process are

$$r_X(t) = \sum_k \mathbb{E}|A_k|^2 \cos(\omega_k t), \quad \sigma_X(d\omega) = \sum_k \frac{\mathbb{E}|A_k|^2}{2}\big(\delta(d\omega - \omega_k) + \delta(d\omega + \omega_k)\big). \tag{23}$$

If one calculates the ensemble- and time-average of the mean-square displacement for such process, the different nodes of oscillation prove to be uncoupled in both time- and ensemble-average sense; the corresponding formulae are sums of terms as in Eq. (21) or (20) (for details see Appendix A). Therefore, any process within this class is stationary, but non-ergodic. The next section will show that it is actually the only case of non-ergodic Gaussian stationary process.

3. Ergodicity in Fourier space

3.1. Spectral form of Maruyama's theorem

All the properties of a Gaussian process can be described interchangeably by its covariance function or its spectral measure; this very specific property of the Gaussian class is caused by its linear structure. The Maruyama theorem also can be expressed in the language of the spectral measure, and this reformulation leads to a surprisingly elegant statement [4].

Theorem 4. *A stationary Gaussian process is ergodic if and only if its spectral measure has no points.*

To fully understand this theorem note that any measure can be decomposed as a sum of three distinct components: the absolutely continuous, singular and discrete measures. For a stochastic process the corresponding decomposition of the spectral measure $\sigma = \sigma_{ac} + \sigma_s + \sigma_d$ causes also the process itself to decompose into three independent components

$$X(t) = X_{ac}(t) + X_s(t) + X_d(t), \tag{24}$$

which is guaranteed by the harmonic representation (11). Now:

- The component $X_d(t)$ is non-ergodic and is a Gaussian harmonic process.
- The component X_{ac} is mixing. It has a power spectral density. The Riemann–Lebesgue lemma shows that in this situation the covariance (and all other memory functions) of X_{ac} decays at infinity, i.e. the values of the process become asymptotically independent at long time scales.
- The last, singular component X_s is ergodic, but its memory structure may be not typical. Its covariance function does not necessarily decay to 0. It may oscillate, but must be aperiodic and the high correlation events must become more scarce as $t \to \infty$.

The set of the measures, for which the covariance function decays, called Rajchman measures, generally does not have convenient description [10]. As a demonstration let us consider an example using the most well-known singular measure: the Cantor measure.

The Cantor set is obtained by removing from the middle one-third the interval [0, 1], then repeating this procedure at two remaining intervals [0, 1/3], [1/3, 1] and recursively applying this procedure infinitely many times. The points which will remain are Cantor points. Elementary calculation proves

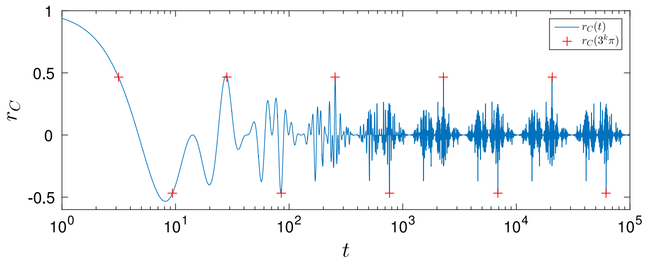

Fig. 1. Plot of covariance function r_C of the process with the Cantor spectral measure.

that the length of the intervals removed during construction of the Cantor set is 1, therefore the Cantor measure cannot have a density and must be singular.

The points in Cantor set can be conveniently characterised as points in interval $[0, 1]$ which have no digits "1" in their ternary representation, i.e. have the representation $\sum_{k=1}^{\infty} d_k 3^{-k}$ for some sequence $d_k \in \{0, 2\}$.

The Cantor measure σ_C is the uniform measure on Cantor set. We move this measure left to the interval $[-1/2, 1/2]$ so that the corresponding process will be real-valued. The Cantor numbers in this interval can be represented as $\sum_{k=1}^{\infty} d_k 3^{-k}$ where $d_k \in \{-1, 1\}$. The corresponding Cantor measure is probably most simple to understand as a discrete uniform distribution on the i.i.d. series D_k, $\mathrm{P}(D_k = -1) = \mathrm{P}(D_k = 1) = 1/2$ mapped onto interval $[-1/2, 1/2]$ using formula $Y = \sum_{k=1}^{\infty} D_k 3^{-k}$. The process X_C which has the Cantor measure as a spectral measure has the covariance function

$$r_{X_C}(t) = \int_{\mathbb{R}} \sigma_C(\mathrm{d}\omega)\, \mathrm{e}^{\mathrm{i}\omega t} = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}t \sum_{k=1}^{\infty} D_k 3^{-k}}\right] = \prod_{k=1}^{\infty} \mathbb{E}\left[\mathrm{e}^{\mathrm{i}t D_k 3^{-k}}\right] = \prod_{k=1}^{\infty} \cos\left(t 3^{-k}\right). \tag{25}$$

In the contrast to more well-known classes of covariance functions, r_{X_C} has a specific property close to a self-similarity

$$r_{X_C}(3t) = \cos(t) r_{X_C}(t), \tag{26}$$

which also guarantees that r_{X_C} does not decay to zero. The extremal points of r_C are located at $t_k = 3^k\pi$, where it attains values $r_C(t_k) = (-1)^{k+1} r_C(\pi) \approx \pm 0.47$, see Fig. 1. It may be not clear that function r_C can be easily calculated numerically, however it can be shown that taking $N \geq \log_3 t$ terms from the product (25), we overshoot no more than to the level $\exp(2t^2 9^{-N}) r_C$, which is a very fast convergence, see proposition in Appendix B and the proof thereof.

This demonstrates that the process X_C has a recurring correlation and cannot be mixing; however, the correlation events are becoming exponentially more rare as the time delay increases, which allows for ergodicity.

Physically, the X_C can be interpreted as a dynamical process generated by the heat bath with Cantor-like geometry in which we observe macroscopic collective average of the oscillators' positions. The fractal structure of such system is recognisable only in the Fourier space. In the position space it is visible only as the recurring correlation. It does not affect e.g. the regularity of the trajectories, which is determined by the asymptotics of the covariance function near $t = 0$. In fact, in the above case the trajectories of X_C are smooth. We comment more on the relation between the heat bath models and ergodicity in Section 4.2.

Unfortunately, there is no simple correspondence between the fractal structure of the spectral measure and the recurring correlation or mixing, which becomes evident even with the slight generalisation of the model. If instead of removing one third we perform the recursive removal procedure such that at any step the remaining intervals on left and right have length one-ηth of the previous one (η being real number bigger than 2), the obtained singular measure and the process is non-mixing for natural η, but it is mixing for any η which is not a Pisot–Vijayaraghavan number [11].

The Pisot–Vijayaraghavan numbers are a closed countable set which causes even infinitesimally small changes of η to change mixing behaviour.

The complex ergodic behaviour complicates the analysis of the models with singular spectral measures, but it is worth stressing the erratic behaviour of covariance functions may be useful for describing the observations which could be otherwise accounted for as an experimental errors. It is also worth noting that the singular measures are gaining attention for their relation to the fractal dynamics and self-similarity [12–14].

3.2. Generalised Maruyama theorem for non-ergodic stationary processes

Any real stationary Gaussian process can be written as

$$X(t) = X_{\mathrm{erg}}(t) + \sum_{k=1}^{N} R_k \cos(\Theta_k + \omega_k t) + X_0, \quad \omega_k \neq 0, \tag{27}$$

which follows from the ergodic decomposition (24), after taking the real part of the harmonic process (22). Variables $R_k := |A_k|$ are amplitudes of the spectral points at frequencies ω_k and have Rayleigh distribution with scale parameters $\sigma_k = \mathbb{E}|A_k|^2$; Θ_k are phases of A_k and have uniform distribution on $[0, 2\pi)$, a consequence of rotational invariance of i.i.d. Gaussian vectors. In full generality the number of spectral points may be infinite, $N = \infty$, however in this case the process X may exhibit complicated aperiodic behaviour; as it is not very important for most of the applicational purposes, further on in this section we limit our considerations to the case $N < \infty$.

The decomposition into non-ergodic and ergodic components yields a useful and straightforward description of the statistical properties of the Gaussian processes. It is made possible by the full characterisation of the invariant sets of this dynamical system.

Theorem 5. *For any stationary Gaussian process X with N spectral points at rationally incommensurable frequencies $\{\omega_k\}_{k=1}^{N}$, the family of invariant sets $\mathcal{C}$ is the σ-algebra $\sigma(X_0, \{R_k\}_k)$.*

The proof is given in Appendix A. The assumption that spectral points at $\{\omega_k\}_{k=1}^{N}$ are rationally incommensurable means that they cannot be represented as $\omega_k = q_k\alpha$ for any rational q_k's. It is fulfilled in most of the real physical systems, in which ω_k are self-frequencies of the harmonic oscillators and depend on the complex set of the system's parameters.

In such case, the aforementioned theorem guarantees that for any observable f, the time average converges to

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{d}\tau\, f(\mathcal{S}_\tau X) = \mathbb{E}\big[f(X)|, X_0, \{R_k\}_k\big], \tag{28}$$

i.e. to the ensemble average calculated under condition that the amplitudes of the spectral points and the constant term X_0 are fixed. For observables which depends on one time moment of X only, $f(X) = f(X(t))$ the above formula simplifies to the explicit integral

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{d}\tau\, f(X(\tau)) = \int_0^{2\pi} \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\theta_2 \dots \int_{\mathbb{R}} \mathrm{d}x \frac{1}{\sqrt{2\pi}c} \mathrm{e}^{-\frac{x^2}{2c^2}} f\Big(x + \sum_k R_k \cos(\theta_k) + X_0\Big), \tag{29}$$

which depends only on the variance of the ergodic component $c^2 = \mathbb{E}\big[X_{\mathrm{erg}}(t)^2\big]$, R_k's and X_0, which are random, but fixed for each trajectory.

In particular, using the cumulative distribution function method, one can calculate the non-ergodic time-averaged probability density $\overline{\rho}$ of any given discrete spectral component with amplitude $R_k = R$

$$\overline{\rho}(x) = \rho_{X_k}(x|R) = \frac{(\pi R)^{-1}}{\sqrt{1 - (x/R)^2}}, \quad -R \le x \le R. \tag{30}$$

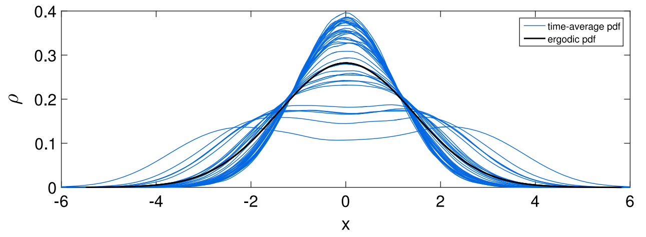

Fig. 2. Time-average kernel density estimation (each blue line corresponds to one trajectory with different random R) and the ergodic density for Gaussian process with one discrete spectral component (black line), $\mathbb{E}\big[R^2\big] = \mathbb{E}\big[X_{\mathrm{erg}}(t)^2\big] = 1$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

This quantity should be observed if one uses time-average estimation of the probability density, e.g. kernel density estimators or histogram [15].

This is an interesting example of a singularity caused by the non-ergodicity. The flat extrema of cosine function are responsible for the probability density divergence of type $x^{-1/2}$ at points $-R$ and $+R$. However, this unusual behaviour is not easy to directly observe as the typical data contains some ergodic component, e.g. some kind of noise from the experimental setup. In such case, the observed empirical probability distribution $\overline{\rho}$ is a convolution of the ergodic component's Gaussian density with a given stationary variance σ^2 and some number of densities of type (30), see Fig. 2. The singular concentration of the probability mass around $-R$ and $+R$ distorts the tails of distribution in specific way: it thickens them by moving the original distribution, but thins them through dividing by a square root factor. More precisely, we can recognise that this convolution of the densities is an example of Laplace transform and use Abelian theorem [16] to obtain asymptotic behaviour of the left tail

$$\begin{aligned}\overline{\rho}(-x) &= \frac{1}{\sqrt{2\pi^3}\sigma R}\int_{-R}^{R}\mathrm{d}y\frac{1}{\sqrt{1-\big(x/R\big)^2}}\mathrm{e}^{-\frac{(y+x)^2}{2\sigma^2}}\\ &= \mathrm{e}^{-\frac{(x-R)^2}{2\sigma^2}}\frac{1}{\sqrt{2\pi^3 R\sigma}}\int_0^{2R}\mathrm{d}y\frac{1}{\sqrt{y}\sqrt{2-y/R}}\mathrm{e}^{-\frac{y^2}{2\sigma^2}}\mathrm{e}^{-\frac{y(x-R)}{\sigma^2}}\\ &\sim \frac{1}{4\pi\sqrt{1-x/R}}\mathrm{e}^{-\frac{(x-R)^2}{2\sigma^2}},\quad x\to\infty;\end{aligned} \tag{31}$$

symmetrically for the right tail. This result does not change significantly for any finite number of spectral points, as it depends only on the presence of singularities of the convolved densities.

The asymptotic formula (31) differs considerably from normal distribution; for most of the realisations the amplitudes R_k are large enough to strongly affect the time-averaged probability density, see Fig. 2. For nearly all realisations statistical tests also show significant non-gaussianity (e.g. Shapiro and Kolmogorov–Smirnov tests). However, for small amplitude of R_k's this effect could be less noticeable.

The fast decay of function $\exp(-x^2)$ may complicate analysis of non-ergodicity through density estimation. More convenient method is to use time-averaged characteristic function. It is a time-average of the observable $f(X) = \exp(\mathrm{i}\theta X(t))$, which for a non-ergodic component equals

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T\mathrm{d}\tau\ \mathrm{e}^{\mathrm{i}\theta X(\tau)} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{d}x\ \mathrm{e}^{\mathrm{i}\theta R\cos(x)} = J_0(R\theta), \tag{32}$$

where J_0 is a Bessel function of the first kind and order 0. Therefore the time-averaged characteristic function $\overline{\phi}$ of any stationary Gaussian process with incommensurable frequencies has form

$$\overline{\phi}(\theta) = \mathrm{e}^{-c^2\theta^2/2}\prod_k J_0\big(R_k\theta\big), \tag{33}$$

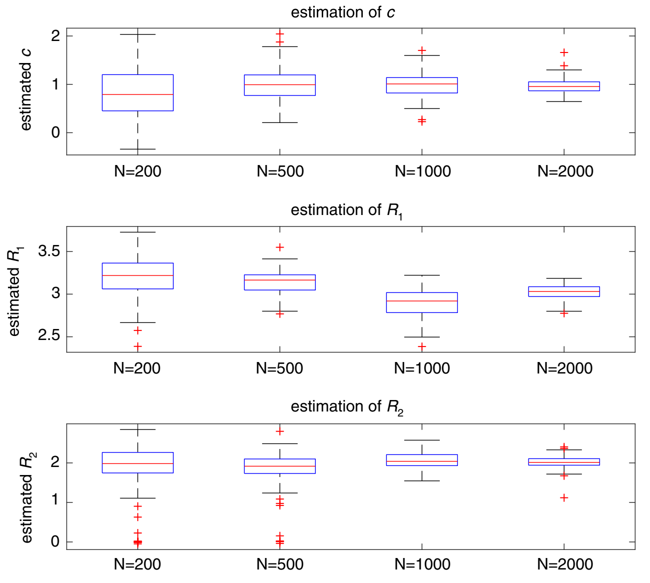

Fig. 3. Estimation of $c = \mathbb{E}[X_{\text{erg}}(t)^2]$, R_1 and R_2 for different sample sizes N, using least-square fit of the time-averaged characteristic function for the process $X_{\text{erg.}}(t) + 3\cos(t) + 2\cos(\sqrt{2}t)$.

where $\exp(-c^2\theta^2/2)$ is the ensemble averaged characteristic function of the ergodic component of the process. The additional, non-ergodic factors $J_0(R_k\theta)$ cause the whole function to have zeros determined by the zeros of function J_0, which may be approximated numerically, the first being $x_1 \approx 2.405$, the second $x_2 \approx 5.520$. Therefore the location of zeros for time-averaged characteristic function may serve to preliminary estimate the number and values of amplitudes R_k. More precise estimation requires least-squares fitting using the formula from Eq. (33). The exemplary results of this procedure are shown in Fig. 3.

The estimators of R_k have tendency to return some undershoot values which cause negative bias, especially for lower lengths of trajectories, but are generally reliable. The results depend on the sample length N, but do not depend on the sampling time Δt as long as $\pi\Delta t$ is incommensurable with $\{\omega_k\}_k$. If this is not true, the measured time series has a periodic component and the proper values of time-averaged observables are obtained in the infill asymptotics, i.e. $\Delta t \to 0$, and $\Delta t N \to \infty$. Such requirements guarantee that the calculated mean converges to the time integral $1/T\int_0^T \mathrm{d}\tau$ where $T \to \infty$.

Let us consider models in which it may happen that $\{\omega\}_k$ are commensurable. This situation may appear in real system when $\{\omega_k\}$ by coincidence or due to some symmetry are close to being commensurable, that is they can be expressed as $\omega_k = \alpha p_k/q_k + \epsilon_k$ where p_k, q_k are small natural numbers and ϵ_k is small compared to $1/T$.

For commensurable $\{\omega_k\}$ the harmonic process is in fact periodic and the length of its period is proportional to the lowest common denominator of ω_k's. We show example in Fig. 4, where the empirical characteristic function of process $R_1\cos(\Theta_1\omega_1 t) + R_2\cos(\Theta_2 + \omega_2 t)$ is presented. For simplicity we fixed $\omega_1 = 1$ and $R_1 = 1$, $R_2 = 2$, as R_k's are constants of motion in this case. Because of the periodicity, the calculated time-average depends on the random initial phases Θ_1, Θ_2.

We actually know the exact dependence on both random amplitudes and phases; the time-averaged characteristic function of any stationary Gaussian process is given by the formula (see Appendix A)

$$\overline{\phi}(\theta) = \mathrm{e}^{-c^2\theta^2/2}\sum_{S\in G}\exp\left(\mathrm{i}\sum_{m_j\in S} m_j\Theta_j\right)\prod_{m_j\in S}\mathrm{i}^{m_j}J_{m_j}(\theta R_j), \tag{34}$$

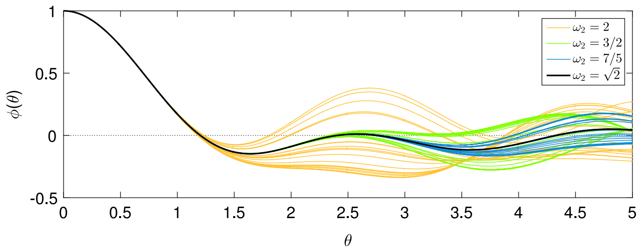

Fig. 4. Time-averaged characteristic function of the process $2\cos(\Theta_1+t)+2\cos(\Theta_2+\omega_2 t)$ for different ω_2 and 20 realisations of Θ_1, Θ_2. Black line, corresponding to the simulation for $\omega_2 = \sqrt{2}$ up to numerical accuracy $\epsilon \sim 10^{-16}$, agrees perfectly with ergodic average $J_0(t)J_0(2t)$, see Eq. (33).

where G is family of sets of m_j's for which $\sum_j m_j\omega_j = 0$. The result depends on linear combinations of the random phases $\sum_j m_j\Theta_j \pmod{2\pi}$. This is true not only for the time-averaged characteristic function, but also for any observable, which is stated in the following theorem:

Theorem 6. *For any stationary Gaussian process X with spectral points $\{\omega_k\}_{k=1}^N$ (some of which may be commensurable), the family of invariant sets $\mathcal{C}$ is the σ-algebra $\sigma(X_0, \{R_k\}_k, \mathcal{M})$, where X_0 is the constant term, R_k are the amplitudes of the atoms of the spectral measure and $\mathcal{M}$ is a family*

$$\mathcal{M} = \left\{ \sum_{m_j} m_j\Theta_j \pmod{2\pi} : \sum_j m_j\omega_j = 0 \right\}. \tag{35}$$

Moreover, $\mathcal{M}$ can be reduced to contain at most $N-1$ integer linear combinations.

The proof is given in Appendix A. This general theorem completely determines the behaviour of time-averaged observables for stationary Gaussian processes and has important practical applications, allowing for statistical analysis of non-ergodic stationary Gaussian models.

As an example, let us come back to the time-averaged characteristic function (34). For the case of process with two spectral points with frequencies $\omega_1 = 1, \omega_2 = 2$, the numerically calculated time-averages are shown as yellow lines in Fig. 4. The integer combinations in family $\mathcal{M}$ from Eq. (35) are exactly $m_1 = 2, m_2 = -1$ and multiples of those. Therefore the time-averaged characteristic function depends only on $2\Theta_1 - \Theta_2 \pmod{2\pi}$. Indeed, any yellow line in Fig. 4 corresponds to many different random choices of Θ_1, Θ_2, and these lines can be parametrised only by the number $c \in [0, 2\pi)$ defined as $2\Theta_1 - \Theta_2 = c \pmod{2\pi}$. Similarly, green lines on the same figure depend only on $3\Theta_1 - 2\Theta_2 \pmod{2\pi}$, because we have $m_1 = 3, m_2 = -2$, and so on.

It may seem counter-intuitive that the rationality or irrationality of the number ω_2, which cannot be experimentally studied, affects numerical simulations and the behaviour of the real systems. This apparent paradox disappears if we carefully analyse the behaviour of the time-averages in the two essential cases

- For $\omega_2 = p/q$ with coprime p, q the trajectory will have period $2\pi q$. For q sufficiently larger than the experiment time ($q \gg T$), the periodicity will be unobservable during the measurement and the time-averaged observables will seem independent of initial phases.
- For irrational $\omega_2 = p/q + \epsilon$ with coprime p, q and $\epsilon q \ll 1$ the time-averaged observables will not depend on initial phases in long-time limit $T \to \infty$, however the process will be very close to a periodic one, therefore the convergence will be slow.

So, we realise that in the real experiment, in which the time of measurement is finite $T < \infty$, the practically significant property is how close ω_2 is to an irreducible fraction with a small denominator. Analogically, for multiple ω_k's we only need to determine how close they are to a set of commensurable numbers with simple rational ratios.

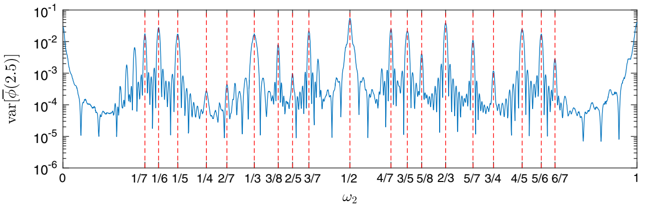

Fig. 5. Numerical estimation of variance var[$\overline{\phi}(2.5)$] obtained using 10^3 samples of trajectories with length $T = 200$ and values of ω_2 taken as one thousand uniformly scattered numbers in interval [0, 1] stored in format double.

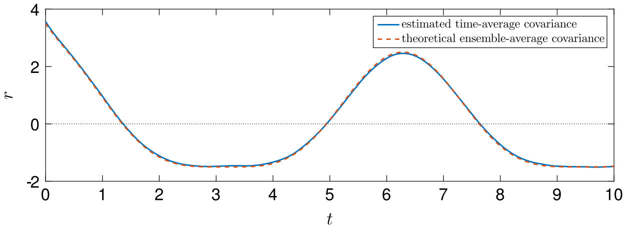

Fig. 6. Comparison between estimated time-average covariance function and a theoretical ensemble one for a stationary non-ergodic process with commensurable frequencies of the spectral points.

In order to illustrate this fact numerically we simulated the process $R_1\cos(\Theta_1+t)+R_2\cos(\Theta_2+\omega_2 t)$ with fixed $R_1 = 2$, $R_2 = 3$ and calculated the variance of the time-averaged characteristic function at one point $\overline{\phi}(2.5)$. For irrational ω_2 and in the limit $T \to \infty$ this quantity should equal zero, but in the finite-time the numerical experiment it is a positive function of ω_2 which indeed measures how close ω_2 is to a simple irreducible fraction. See Fig. 5, where the peaks of the variance indicate the positions of the simplest irreducible fractions like 1/2, 1/3, 2/3, 1/4, 3/4 *etc*.

Another result, which may be somehow unexpected, is that according to the calculation in Section 2.2, even for commensurable $\{\omega_k\}_k$ the time-average second-order properties depend only on $\{R_k\}_k$ and the dependence of initial phases $\{\Theta_k\}_k$ is lost. Indeed, it is sufficient to prove the agreement of the time- and ensemble-averaged covariance function, which we show in Appendix A. It means that the additional memory structure induced by the commensurability of $\{\omega_k\}_k$ is purely non-linear and it is possible to detect it only using higher-order statistics, e.g. the time-averaged characteristic function. This fact do not contradict the purely linear dependence structure of the Gaussian processes, because this property applies only to the time-averages, which in non-ergodic case make use only of a part of the full information contained in the process. The additional non-linear dependences can be interpreted as shadows of the full, linear but inaccessible information.

The exemplary covariance estimation is shown in Fig. 6, where we sampled the stationary Ornstein–Uhlenbeck process with mean-returning parameter $\lambda = 3$, and the addition of two spectral components $2\cos(t)$ and $\cos(2t)$. The covariance conditioned by $R_1 = 2$, $R_2 = 1$ is

$$r(t) = \mathrm{e}^{-\lambda t} + 2\cos(t) + \frac{1}{2}\cos(2t) \tag{36}$$

and it agrees with the estimated time-averaged covariance, as shown in Fig. 6.

4. Ergodicity of the linear response systems

4.1. Linear filters

In this section we will make use of three basic facts.

Proposition 1. *Let us consider any finite measure σ and a measurable function f, defined σ-almost everywhere. In this case:*

1. *If σ is absolutely continuous, then σf is absolutely continuous.*
2. *If σ is Rajchman, then σf is Rajchman.*
3. *If σ is continuous, then σf is continuous.*

The function f is often called the spectral gain.

Fact 1 fallows from the definition of the measure σf. For any measurable set A it equals, by definition

$$(\sigma f)(A) = \int_A \sigma(\mathrm{d}\omega) f(\omega), \tag{37}$$

in short $(\sigma f)(\mathrm{d}\omega) = \sigma(\mathrm{d}\omega)f(\omega)$. So, if σ has form $\sigma(\mathrm{d}\omega) = \mathrm{d}\omega s(\omega)$ then σf has a form $\mathrm{d}\omega s(\omega)f(\omega)$. Fact 2 is a known result from the measure theory and can be proven using trigonometric polynomials [10,17]. The proof of Fact 3 is also simple: $\sigma(\{x_0\})f(x_0) \neq 0$ only if $\sigma(\{x_0\}) \neq 0$.

Considered together, these three facts guarantee that if Gaussian process X has spectral measure σ, then Gaussian process Y with spectral measure σf inherits all ergodic properties (ergodicity, mixing) from X. Process Y can only have more ergodic properties than X.

The above proposition can be used to determine the ergodic behaviour of various transformations of a given process. If process X has spectral process S (i.e. is given by Eq. (11)), then the time-shifted process $t \mapsto X(t-T)$ has harmonisable representation

$$X(t-T) = \int_{\mathbb{R}} \mathrm{d}S(\omega)\, \mathrm{e}^{-\mathrm{i}\omega T} \mathrm{e}^{\mathrm{i}\omega t}, \tag{38}$$

in other words its spectral process has increments $\mathrm{d}S(\omega)\mathrm{e}^{-\mathrm{i}\omega T}$. Consequently, it has the same spectral measure and the same distribution as X. The direct generalisation of this fact is that any process given by

$$Y(t) = \sum_k a_k X(t - T_k), \tag{39}$$

with deterministic $\sum_k |a_k|^2 < \infty$, has a harmonisable representation

$$Y(t) = \int_{\mathbb{R}} \mathrm{d}S(\omega) \sum_k a_k \mathrm{e}^{-\mathrm{i}\omega T_k} \mathrm{e}^{\mathrm{i}\omega t}, \tag{40}$$

therefore has spectral measure $\sigma(\mathrm{d}\omega)\left|\sum_k a_k \mathrm{e}^{-\mathrm{i}\omega T_k}\right|^2$ and inherits the ergodic properties of the process X. Process of this form appears directly e.g. in the biological applications [18].

Because the mean square displacement is the variance of the process $Y = X(t+\Delta) - X(t)$ we can use the above results and obtain

$$\delta^2(\Delta) = \mathbb{E}|X(t+\Delta) - X(t)|^2 = \int_{\mathbb{R}} \sigma(\mathrm{d}\omega) \left|\mathrm{e}^{\mathrm{i}\omega\Delta} - 1\right|^2 = 4 \int_{\mathbb{R}} \sigma(\mathrm{d}\omega) \sin\left(\frac{\omega}{2}\right)^2. \tag{41}$$

Moreover, taking limit $\lim_{h\to 0}(X(t+h) - X(t))/h$ one obtains the harmonisable representation of the mean-square derivative

$$\frac{\mathrm{d}}{\mathrm{d}t} X(t) = \mathrm{i} \int_{\mathbb{R}} \mathrm{d}S(\omega)\, \omega \mathrm{e}^{\mathrm{i}\omega t}, \tag{42}$$

which exists if and only if $\int \sigma(\mathrm{d}\omega)\, \omega^2 < \infty$. Analogically, any process given by

$$Y(t) = \sum_k a_k \frac{\mathrm{d}^k}{\mathrm{d}t^k} X(t), \tag{43}$$

with $\sum_k |a_k|^2 \omega^{2k} < \infty$ for ω in support of σ, has harmonic representation

$$Y(t) = \int_{\mathbb{R}} \mathrm{d}S(\omega) \sum_k a_k (\mathrm{i}\omega)^k \mathrm{e}^{\mathrm{i}\omega t}, \tag{44}$$

and spectral measure $\sigma(\mathrm{d}\omega)\left|\sum_k a_k(\mathrm{i}\omega)^k\right|^2$. Because $\left|\sum_k a_k(\mathrm{i}\omega)^k\right|^2$ is a continuous function defined on the whole $\mathbb{R}$, the process Y inherits ergodic properties of X. The non-ergodicity of X can be not present in Y, because the function $\omega \mapsto \left|\sum_k a_k(\mathrm{i}\omega)^k\right|^2$ may have zeros and if such zero agrees with position of spectral point of X, the process Y does not contain this spectral point. The simplest such case is when X has exactly one spectral point at $\omega = 0$, that is it contains a time-independent Gaussian constant X_0. In this situation any time-averaged observable which depends on the mean of X does not converge to the ensemble-average. However, the derivative $\frac{\mathrm{d}}{\mathrm{d}t}X(t)$, corresponding to the spectral gain function $\omega \mapsto \omega^2$, does not contain X_0 and is ergodic.

Similar reasoning generalises formula (39). For the convolution

$$Y(t) = g * X(t) = \int_{\mathbb{R}} \mathrm{d}s\, g(t-s)X(s) \tag{45}$$

it yields

$$Y(t) = \int_{\mathbb{R}} \mathrm{d}S(\omega)\, \hat{g}(\omega)\mathrm{e}^{\mathrm{i}\omega t}, \tag{46}$$

where $\hat{g}$ is the Fourier transform of $g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Analogically to the previous case, Y inherits ergodic properties of X. If zeros of $\hat{g}$ agree with spectral points of X, the process Y may be ergodic whereas X is not. Moreover, if singular non-mixing measure of X is contained in the domain outside of $\hat{Y}$ support, Y may be mixing when X is non-mixing.

Special care must be taken in more general case $g \in L^2(\mathbb{R})$ but $g \notin L^1(\mathbb{R})$. The Fourier transform of such g has jumps. The well-known result from the Fourier theory [19] states that if $\hat{g}$ has a jump at ω_0, then $\int_{-T}^{T} \mathrm{d}t\, g(t)\mathrm{e}^{\mathrm{i}\omega_0 t}$ converges as $T \to \infty$ to the value $(\hat{g}(\omega_0^-)+\hat{g}(\omega_0^+))/2$, i.e. to the exact middle of the discontinuity of $\hat{g}$. If the process X has spectral point at the exact frequency ω_0, then for the process Y this spectral point will be modulated by $\left|\hat{g}(\omega_0^-)+\hat{g}(\omega_0^+)\right|^2/4$ as long as the filter is applied symmetrically as a limit of convolutions with functions g which are supported on interval $[-T, T]$.

The representation (46) vastly increases the number of models for which it is easy to study ergodicity using Fourier methods, as using convolution is one of the most often chosen methods to model time-invariant linear responses of the system (see also Section 4.2). One of the most common examples that appear in practice is g being one or two sided exponent decaying with rate λ_d. These choices correspond to the spectral responses $1/(\lambda_d^2 + \omega^2)$ or $4\lambda + d^2/(\lambda_d^2 + \omega^2)$, respectively.

One other practical consequence is that one can filter out non-ergodicity from the data. Using estimators of power spectral density (e.g. periodogram [20]) the locations of spectral points can be estimated and the corresponding non-ergodicity removed by using any filter with zeros as its spectral gain function at their frequencies. The simplest choice of such filter is the smoothing

$$\tilde{X}(t) = \int_{t-\pi/\omega_0}^{t+\pi/\omega_0} \mathrm{d}s\, X(s), \tag{47}$$

which integrates the spectral component at ω_0 over its period, therefore removing it. It corresponds to the filter g_{ω_0} and spectral gain $\hat{g}_{\omega_0}$

$$g_{\omega_0}(t) = \begin{cases} 1, & |t| \le \pi/\omega_0, \\ 0, & |t| > \pi/\omega_0, \end{cases} \quad \hat{g}_{\omega_0}(\omega) = \sqrt{\frac{2\omega_0^2}{\pi^3}}\, \mathrm{sinc}\left(\frac{\omega\pi}{\omega_0}\right). \tag{48}$$

For multiple number of spectral points one may use filter $g_{\omega_1} * g_{\omega_2} \ldots g_{\omega_N}$ or any other with the suitable spectral gain. The gain $|\hat{g}|^2$ behaves like $\sim(\omega-\omega_0)^2$ near ω_0, therefore it removes the spectral point ω_0 in a numerically stable manner. If one is more interested in sure removal of the non-ergodicity near the location ω_0 than not distorting the spectrum, he can use spectral gain more flat around ω_0, e.g. using triangular function filter guarantees asymptotical behaviour $\sim(\omega-\omega_0)^4$. Other useful choice is spectral gain

$$\hat{g}(\omega) = \begin{cases} 1 - \mathrm{e}^{-(\omega-\omega_0)^2/c}, & |\omega| \le L; \\ 0, & |\omega| > L, \end{cases} \tag{49}$$

which allows for calibration of level of distortion around ω_0 (parameter c) and frequency cut-off (parameter L). The corresponding filter can be expressed using error, Gaussian and trigonometric functions, so it can be easily computed for the purpose of statistical usage.

The above approach can be understood as, instead of using original observables $f(X)$, use the modified observable $\tilde{f}(X) = f(Y)$ for which the time- and ensemble-averages coincide, even when this is not generally the case. Analysing the transformed process Y instead of X may be more difficult, as properties of X are distorted by filtering, albeit in controlled manner. However, for a small number of spectral points it is manageable, moreover it can be used as an effective method of localising spectral points: if the observables of the filtered process behave ergodically, it is a statistical verification of good choice of these locations. Further analysis can be performed on the filtered process Y, which is ergodic, or by staying with the original X and using methods from Section 3.2.

4.2. Linear response systems and Langevin equations

Probably the most basic example of an ergodic Gaussian process is the solution of the classical Langevin equation

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t) = -\lambda X(t) + \xi(t), \tag{50}$$

governed by white Gaussian noise $\xi(t)$ defined as increments of the Brownian motion $\mathrm{d}t\xi(t) = \mathrm{d}B(t)$. The stationary solution of this equation is given by the convolution

$$X(t) = \int_{\mathbb{R}} \mathrm{d}s\xi(s)\, G(t-s) = \int_{\mathbb{R}} \mathrm{d}B(s)\, G(t-s), \tag{51}$$

where the Green function G is given by the corresponding deterministic problem

$$\frac{\mathrm{d}}{\mathrm{d}t}G(t) = -\lambda G(t) + \delta(t). \tag{52}$$

Here δ is Dirac delta considered as a distribution and the whole equation is interpreted in the distributional sense. The solution is called casual when $G(t) = 0$ for $t < 0$ and the process X at time t depends only on past values of the noise, i.e. $\mathrm{d}B(s)$ for $s \leq t$. In such case the solution is one-sided exponential decay $G(t) = \mathrm{e}^{-\lambda t}$ for $t \geq 0$. It is stationary, therefore it must have harmonisable representation. The Fourier transform of the Green function must satisfy $\mathrm{i}\omega\hat{G} = -\lambda\hat{G} + 1$, so it must be equal to $\hat{G}(\omega) = (\lambda + \mathrm{i}\omega)^{-1}$, and the harmonisable representation is

$$X(t) = \int_{\mathbb{R}} \mathrm{d}\hat{B}(\omega)\, \frac{1}{\lambda + \mathrm{i}\omega}\mathrm{e}^{\mathrm{i}\omega t}, \tag{53}$$

where $\mathrm{d}\hat{B}$ is white Gaussian noise in the Fourier space (which can be interpreted as a generalised Fourier transform of $\mathrm{d}B$, see [21]). Therefore the solution is mixing and has power spectral density $(\lambda^2 + \omega^2)^{-1}$.

The above result is much more general. Any stationary solution of the linear differential system

$$a_k \frac{\mathrm{d}^k}{\mathrm{d}t^k}X(t) + a_{k-1}\frac{\mathrm{d}^{k-1}}{\mathrm{d}t^{k-1}}X(t) + \cdots + a_0 X(t) = \xi(t), \tag{54}$$

has a casual solution given by a proper Green's function and harmonisable representation

$$X(t) = \int_{\mathbb{R}} \mathrm{d}S(\omega)\, \frac{1}{\sum_{j=1}^{k} a_j(\mathrm{i}\omega)^j}\mathrm{e}^{\mathrm{i}\omega t}, \tag{55}$$

where $\mathrm{d}S(\omega) = \mathrm{d}\omega\hat{\xi}(\omega)$ is a spectral process of the stationary Gaussian noise ξ, which in general may not be white noise. It is clear that X inherits ergodic properties of ξ. On the contrary to the case considered in Section 4.1 it cannot be ergodic when ξ is not, as the rational function $1/|\sum_j a_j(\mathrm{i}\omega)^j|^2$ does not have zeros.

Very similar reasoning applies to systems with more rich memory structure, namely described by the generalised Langevin equation [22]

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t) = -\int_0^t \mathrm{d}s\, K(t-s)X(s) + \xi(t). \tag{56}$$

The equation in the above form does not have stationary solution. However, its non-stationary solution, determined, again, by the corresponding Green function

$$X(t) = X_0 G(t) + \int_0^t \mathrm{d}s \xi(s)\, G(t-s), \tag{57}$$

converges pointwise as $t \to \infty$ to the stationary process given by

$$X(t) = \int_{-\infty}^t \mathrm{d}s \xi(s)\, G(t-s). \tag{58}$$

The above integrals and convergence are defined in mean-square or almost sure sense, depending on the regularity of memory kernel K and noise ξ [23,24]. This process is a solution of the stationary generalised Langevin equation

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t) = -\int_{-\infty}^t \mathrm{d}s\, K(t-s)X(s) + \xi(t). \tag{59}$$

In this form we recognise the convolution of the process X with the kernel K understood as a casual function. For $K \in L^2(\mathbb{R})$, the harmonisable representation of solution is, analogically to the previous results,

$$X(t) = \int_{\mathbb{R}} \mathrm{d}S(\omega)\, \frac{1}{\mathrm{i}\omega + \hat{K}(\omega)}, \tag{60}$$

where once again $\mathrm{d}S(\omega) = \mathrm{d}\omega\hat{\xi}(\omega)$. Without any change to the previous considerations, the solution X inherits the ergodic properties from the noise ξ. The formula (60) is valid even for some kernels $K \notin L^2(\mathbb{R})$. The well-used model of subdiffusion is generalised Langevin equation in which ξ is fractional Brownian noise [23,25] and the kernel is $K(t) = t^{-\alpha}$, $0 < \alpha < 2$. One still can use the Fourier theory for functions outside $L^2(\mathbb{R})$, where $\hat{K}(\omega) = |\omega|^{\alpha-1}\Gamma(1-\alpha)\mathrm{e}^{\mathrm{i}\,\mathrm{sgn}(\omega)\pi(\alpha-1)/2}$ [19], substitute it into Eq. (60) and obtain valid result [23].

The above considerations apply also to the full form of the generalised Langevin equation [22,23,25] with the external potential

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}X(t) = -\lambda X(t) - \int_{-\infty}^t \mathrm{d}s\, K(t-s)\frac{\mathrm{d}}{\mathrm{d}t}X(s) + \xi(t), \tag{61}$$

in which the similarity to the Newton equation is more visible. When modelling diffusion by Eq. (59) the process X describes velocity and its integral, the position, is not ergodic, and even not stationary (the particle is not confined). For Eq. (61) the process X is position itself, and due to confining term $-\lambda X(t)$ it is stationary and ergodic when the driving noise ξ is.

The Fourier space approach gives also additional insight into the physical origin of the non-ergodicity. In many applications based on classical or quantum statistical models, the generalised Langevin equation is derived from the bath of harmonic oscillators model. It is often called Kaz–Zwanzig model for classical systems [22,23,26,27] and Caldeira–Leggett model for quantum systems [28,29]. In any case the Hamiltonian depends on macroscopic coordinate X, P, microscopic oscillators $\{q_j, p_j\}_j$ and is given by

$$\mathcal{H} = \frac{P^2}{2M} + \sum_j \frac{p_j^2}{2m_j} + \sum_j \frac{m_j\omega_j^2}{2}q_j^2 + \sum_j \gamma_j q_j X, \tag{62}$$

with the possible addition of the macroscopic potential acting on X. The above sum may be infinite as long as the total energy $\mathcal{H}$ is finite. The analytic formula of the noise ξ may be obtained solving the corresponding Hamilton equations [22,23], and it has a purely point spectrum supported on self-frequencies of the bath harmonic oscillators

$$\sigma_\xi(\mathrm{d}\omega) = \sum_j m_j \frac{\gamma_j^2}{2\omega_j^2}\big(\delta(\mathrm{d}\omega - \omega_j) + \delta(\mathrm{d}\omega + \omega_j)\big). \tag{63}$$

This fact already prohibits ergodicity of solution. It is nothing surprising from the point of view of abstract ergodic theory, which excludes ergodicity for systems governed by quadratic Hamiltonians such as of discrete heat bath [10]. Moreover, even the stationary solution does not exist. Eq. (59) is not directly derived from the Kaz–Zwanzig model. Instead, one can strictly derive only its non-stationary variants like Eq. (56), in which the convolution integral is taken from 0 to t and initial conditions must be provided. The stationary equation (59) is a long-time limit of the former one, but only under assumption that the stationary solution exists. This is not generally the case.

In this model, the fluctuation–dissipation theorem holds and states that kernel K is equal to the covariance function of ξ. For a finite number of oscillators this is a finite sum of cosines, which is periodic. Even for infinite number of oscillators the process ξ is non-ergodic, so both its covariance function and kernel K do not decay at infinity (see mixing condition (8)). It means that even in the thermodynamical limit $N \to \infty$ the Kaz–Zwanzig model cannot describe the long-time asymptotics of the most often used memory functions. These can be nonetheless approximated by this model in finite time scales, most simply by the spectral points which form the Fourier series of a given covariance function in a fixed interval.

Non-decaying kernel also causes problems with divergence of the convolution integral in Eq. (59). There is no spectral measure that would fulfil Eq. (59) or (61) so there cannot be stationary solution in the classical sense. The model, however, is still a valid derivation of the stationary but non-ergodic stochastic noise ξ. Only the solution of the Generalised Langevin equation cannot be stationary if the memory kernel is given by fluctuation–dissipation theorem.

The amplitudes of the spectral points depend on energies of the harmonic oscillators. If these may be considered small, under proper rescaling the bath may replaced by a smooth field ϕ, π. Under this assumption the system is described by field Hamiltonian [30]

$$\mathcal{H} = \frac{P^2}{2M} + \frac{1}{2}\int \mathrm{d}x\, |\pi(x)|^2 + \frac{1}{2}\int \mathrm{d}x\, |\phi'(x)|^2 + \int \mathrm{d}x\, d(x)\phi'(x)X, \tag{64}$$

where the function d, called the density of states, couples the field to the coordinate X and the conjugated momentum P. Function d must be square integrable for the last term to be finite. The generalised Langevin equation can still be derived and has the same form as in the previous case, but this time the noise ξ has a power spectral density $|\hat{d}|^2$. The system is mixing. Besides that there are no essential restrictions on the type of memory determined by this model. The Fourier transform is a bijection from $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$, so in this model $|\hat{d}|^2$ can be any $L^1(\mathbb{R})$ function [31]. The possible corresponding covariance functions form a wide space which encompasses commonly used exponential and power law decays.

The most important conclusion from this comparison of these two related models is that the ergodicity is determined by the type of the heat bath in the physical model: the discrete or the field one.

5. Comments about non-Gaussian case

In previous sections we concentrated on Gaussian processes. Here we want to briefly comment on the limitations of this methodology. The non-linear memory structure of non-Gaussian process is in general so complex that analysing ergodicity is complicated. Even negative results are not commonly available. Nevertheless, some generalisations can be made. First, any harmonic process

$$X(t) = \sum_k R_k \cos(\Theta_k + \omega_k t) + X_0, \tag{65}$$

is non-ergodic for random R_k, as the analysis performed in Section 2.2 does not depend on the precise distribution of R_k's. These processes are also stationary, because under any time-shift τ, the joint distribution of the phases does not change: $[\Theta_1 + \tau\omega_1, \Theta_2 + \tau\omega_2 \ldots] \,(\text{mod } 2\pi) \overset{d}{=} [\Theta_1, \Theta_2 \ldots]$. The deterministic R_k's can be treated as a degenerate case of Gaussian variables with zero variance, so in this specific case the process is ergodic if frequencies ω_k are incommensurable.

Ergodic properties of the general class of infinitely divisible processes were described in using generalised memory functions such as codifference and correlation cascade [32–35]. For this class, the notion of a harmonisable process is also in use and is defined by the analogical formula [8,13]

$$X(t) = \int_{\mathbb{R}} \mathrm{d}S(\omega)\, \mathrm{e}^{\mathrm{i}\omega t}, \tag{66}$$

where S is a rescaling of a Lévy processes L, i.e. process with independent stationary increments, for which additionally the increments are rotationally invariant.

When $\mathrm{d}S(\omega) = \mathrm{d}L(\omega)f(\omega)$, the integral (66) may be defined using Poisson random measure [36] and if Lévy process L has second moment, its covariance structure is the same as for a harmonisable Gaussian process with power spectral density $|f|^2$. The process X is generally non-ergodic, but the time-averaged covariance function converges in the mean-squared sense to the ensemble one if and only if

$$\frac{1}{T^2}\int_0^T \mathrm{d}t_1 \int_0^T \mathrm{d}t_2\, \mathbb{E}[X(t_1+\Delta)X(t_2+\Delta)X(t_1)X(t_2)] \xrightarrow{T\to\infty} 0, \tag{67}$$

which is a direct application of the second order ergodic theorem [37]. In practice it is most often expected that this equation is fulfilled and the detection of non-ergodicity in this case requires using higher-order observables.

In the case of stable processes, $\mathrm{d}S(\omega)$ is interpreted as rotationally invariant α-stable random measure with some control measure $\sigma_\alpha(\mathrm{d}\omega)$. The stable harmonisable processes are stationary, however they are distinct from the stable moving-average processes, in particular solutions of Langevin equations [13]. This fact by itself limits their practical applications. Moreover, they are all non-ergodic, as shown in [38]. Still, there is an analogue of the elegant representation of the covariance structure of Gaussian harmonisable processes. For stationary stable processes, instead of covariance, the codifference function

$$\tau(t) := 2\ln\mathbb{E}\left[\mathrm{e}^{\mathrm{i}X(0)}\right] - \ln\mathbb{E}\left[\mathrm{e}^{\mathrm{i}(X(t)-X(0))}\right] \tag{68}$$

is used, together with notions of long- and short-dependence analogical to those used for covariance. For harmonisable processes the codifference is given by relatively simple formula, which generalises the Gaussian result from Section 2.1

$$\tau(t) = 4^\alpha \int_{\mathbb{R}} \sigma_\alpha(\mathrm{d}\omega) \left|\sin\left(\frac{\omega}{2}\right)\right|^\alpha - 2\int_{\mathbb{R}} \sigma_\alpha(\mathrm{d}\omega), \tag{69}$$

which is a particular case of more general result proven in Appendix B. This quantity may be used in practice to study the memory structure of harmonisable stable processes. However, the behaviour the time-averaged version of (68) was not yet sufficiently studied.

6. Summary

The main mathematical results of this work are Theorems 5 and 6 from Section 3.2. These two results determine the asymptotic behaviour of time-averages which correspond to physically important properties for important and large class of models. Using these theorems we show that in order to thoroughly study the ergodicity the non-linear statistics must be used. Therefore we propose a new tool: the time-averaged characteristic function, and show the theoretical properties of this function in some important cases. The validity of our methods is checked using Monte-Carlo simulations.

The general overview of the spectral theory of the Gaussian processes and its links to the ergodic theory is also an important part of the paper. We try to provide physical interpretation of many mathematical results in this field. We use them to explain the behaviour of solutions for the generalised Langevin equation and their relation to the underlying physical models of the heat bath. We also show few other examples of systems described by the spectral theory, such as the process with Cantor spectral measure, which exhibits unusual recurrent correlations.

Acknowledgement

The research was supported by NCN Maestro Grant No. 2012/06/A/ST1/00258.

Appendix A. Proof of the generalised Maruyama ergodic theorem

Proof of Theorem 5. We will combine methods for trigonometric series presented in [39] and Gaussian ergodic theorem from Section 3.9 of [4].

We will use representation

$$X(t) = X_{\text{erg}}(t) + \sum_{k=1}^{N} R_k \cos(\Theta_k + \omega_k t) + X_0, \quad \omega_k \neq 0, \tag{A.1}$$

where ω_k are distinct. The process $X_{\text{erg}}(t)$ has no spectral points. The full distribution of X is generated by the values $X_{\text{erg}}(t_l)$, $R_j \cos(\Theta_j + \omega_j t_j)$ and X_0. Therefore it is sufficient to study the time-average distribution of the sum

$$\sum_{l=1}^{L} \theta_l X_{\text{erg}}(t + t_l) + \sum_{j=1}^{N} \lambda_j R_j \cos(\Theta_j + \omega_j t_j + \omega_j t) + \lambda_0 X_0. \tag{A.2}$$

For brevity we denote $\tilde{\Theta}_j := \Theta_j + \omega_j t_j$. As the distribution is uniquely determined by the corresponding characteristic function, we will compare the time-averaged characteristic function and the ensemble-averaged one given condition $\sigma(X_0, \{R_k\})$. Their equality will prove the theorem.

First we calculate the conditional ensemble-averaged characteristic function. The variables $\tilde{\Theta}_j$ have the same distribution as Θ_j modulo 2π, because they are independent of each other, R_j's, X_0, X_{erg}, and have marginal uniform distribution. We get

$$\begin{aligned}
&\mathbb{E}\left[e^{i\left(\sum_{l=1}^{L} \theta_l X_{\text{erg}}(t+t_l) + \sum_{j=1}^{N} \lambda_j R_j \cos(\tilde{\Theta}_j + \omega_j t) + \lambda_0 X_0\right)} | X_0, \{R_k\} \right] \\
&= \mathbb{E}\left[e^{i\left(\sum_{l=1}^{L} \theta_l X_{\text{erg}}(t_l)\right)} \right] \prod_{j=1}^{N} \mathbb{E}\left[e^{iR_j \cos(\tilde{\Theta}_j)} | R_j \right] e^{i\lambda_0 X_0} \\
&= \phi_{\theta_1,\ldots,\theta_L} \prod_{j=1}^{N} J_0(R_j \lambda_j) e^{i\lambda_0 X_0}, \quad \phi_{\theta_1,\ldots,\theta_L} := \mathbb{E}\left[e^{i\left(\sum_{l=1}^{L} \theta_l X_{\text{erg}}(t_l)\right)} \right],
\end{aligned} \tag{A.3}$$

where J_0 are Bessel functions of first kind and order 0; they stem from the formula

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} dx\, e^{i\lambda R \cos(x)} = J_0(\lambda R). \tag{A.4}$$

The rest of the proof will be the calculation of the time-average.

We denote

$$\Phi(t) := e^{i\left(\sum_{l=1}^{L} \theta_l X_{\text{erg}}(t+t_l)\right)} \tag{A.5}$$

which appear in the integral used during calculations of time-average

$$I_T := \int_0^T d\tau\, \Phi(t) e^{i\left(\sum_{j=1}^{L} \lambda_j R_j \cos(\tilde{\Theta}_j + \omega_j t)\right)} e^{i\lambda_0 X_0}. \tag{A.6}$$

The factor $e^{i\lambda_0 X_0}$ already agrees with the conditional average, so we will assume $X_0 = 0$ later on for brevity.

Next we expand each exponent of cosine using Jacobi–Anger identity

$$e^{iz\cos(w)} = \sum_{m=-\infty}^{\infty} i^m e^{imw} J_m(z), \tag{A.7}$$

obtaining

$$\begin{aligned} I_T &= \int_0^T d\tau\ \Phi(\tau) \prod_{j=1}^N \sum_{m=-\infty}^{\infty} i^m e^{im(\tilde{\Theta}_j+\omega_j\tau)} J_m(\lambda_j R_j) \\ &= \int_0^T d\tau\ \Phi(\tau) \sum_{S\in M_N} \prod_{m_j\in S} e^{im_j(\tilde{\Theta}_j+\omega_j\tau)} i^{m_j} J_{m_j}(\lambda_j R_j) \\ &= \int_0^T d\tau\ \Phi(\tau) \sum_{S\in M_N} \exp\left(i \sum_{m_j\in S} m_j(\tilde{\Theta}_j+\omega_j\tau)\right) \prod_{m_j\in S} i^{m_j} J_{m_j}(\lambda_j R_j) \\ &= \sum_{S\in M_N} \exp\left(i \sum_{m_j\in S} m_j\tilde{\Theta}_j\right) \prod_{m_j\in S} i^{m_j} J_{m_j}(\lambda_j R_j) \int_0^T d\tau\ \Phi(\tau) \exp\left(i\tau \sum_{m_j\in S} m_j\omega_j\right) \end{aligned} \tag{A.8}$$

where M_N is the family of all N-element subsets of integers S. Exchanging the order of infinite sum and integral is possible because the integrated function is bounded by 1, which also justifies commuting the sum and the limit in the next step. We shall denote $\Omega_S := \sum_{m_j\in S} m_j\omega_j$ and check that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T d\tau\ \Phi(\tau) \exp(i\tau\Omega_S) = 0,\ \text{ if } \Omega_S \neq 0. \tag{A.9}$$

The proof of this statement is given in the lemma below. Let us conclude the whole proof. Because $\{\omega_k\}$ are rationally incommensurable, the equality $\Omega_S = \sum_{m_j\in S} m_j\omega_j = 0$ holds for integer m_j only when $m_1 = m_2 = \cdots = m_N = 0$. In the sum $\sum_{S\in M_N}$ only one element $S = \{0, 0, \ldots, 0\}$ remains and

$$\lim_{T\to\infty} \frac{1}{T} I_T = \prod_{j=1}^N J_0(\lambda_j R_j) \lim_{T\to\infty} \frac{1}{T} \int_0^T d\tau\ \Phi(\tau) = \prod_{j=1}^N J_0(\lambda_j R_j) \phi_{\theta_1,\ldots,\theta_L}, \tag{A.10}$$

where the last equality holds due to the ergodicity of X_{erg}. This is the desired conditional mean. ■

Lemma 1.

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T d\tau\ \Phi(\tau) \exp(i\tau\Omega) = 0, \tag{A.11}$$

for $\Omega \neq 0$ and Φ given by (A.5).

Proof. First note that Φ is a strictly stationary random process. Take $U \sim \mathcal{U}([0, 2\pi))$ independent of Φ. Process $t \mapsto \Phi(t)\exp(iU + \Omega t)$ is also stationary and has finite first moment equal to 1. Therefore, the Birkhoff ergodic theorem guarantees the almost sure existence of time-average

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T d\tau\ \Phi(\tau) \exp(iU\tau\Omega) = \exp(i\Omega) \lim_{T\to\infty} \frac{1}{T} \int_0^T d\tau\ \Phi(\tau) \exp(\tau\Omega) = \hat{X}(\Omega). \tag{A.12}$$

So, the limit (A.11) also exists almost surely and equals a random variable $\exp(-iU)\hat{X}(\Omega)$ (for more details on $\hat{X}(\Omega)$ see Doob [3] Chapter XI.2, page 516). We will prove that it is 0.

Let r be covariance function of X_{erg} and σ its continuous spectral measure. We will study $\mathbb{E}|\cdot|^2$ of the above time-average and show its mean-square convergence to 0, which suffices to prove also almost sure convergence to the same limit.

$$
\begin{aligned}
&\frac{1}{T^2}\int_0^T \mathrm{d}\tau_1 \int_0^T \mathrm{d}\tau_2 \mathbb{E}\left[\exp\left(\mathrm{i}\sum_j^L \theta_j\big(X_{\mathrm{erg}}(t_j+\tau_1)-X_{\mathrm{erg}}(t_j+\tau_2)\big)\right)\right]\mathrm{e}^{\mathrm{i}\Omega(\tau_1-\tau_2)}\\
&\quad=\frac{C}{T^2}\int_0^T \mathrm{d}\tau_1 \int_0^T \mathrm{d}\tau_2 \exp\left(\sum_{j,k=1}^L \theta_j\theta_k r(t_k-t_j+\tau_1-\tau_2)\right)\mathrm{e}^{\mathrm{i}\Omega(\tau_1-\tau_2)}\\
&\quad=\frac{C}{T^2}\int_0^T \mathrm{d}\tau_1 \int_0^T \mathrm{d}\tau_2 \exp\left(\int_{\mathbb{R}} \mathrm{d}\tilde{\sigma}(\omega)\mathrm{e}^{\mathrm{i}\omega(\tau_1-\tau_2)}\right)\mathrm{e}^{\mathrm{i}\Omega(\tau_1-\tau_2)},
\end{aligned}
\tag{A.13}
$$

where we denoted by C the factor before the integral and by $\tilde{\sigma}$ the modified spectral measure; it is just multiplied by a continuous function.

$$
C=\exp\left(2\sum_{j,k=1}^L \theta_i\theta_j r(t_k-t_j)\right),\quad \mathrm{d}\tilde{\sigma}(\omega):=\mathrm{d}\sigma(\omega)\left|\sum_{j=1}^L \theta_j \mathrm{e}^{\mathrm{i}\omega t_j}\right|^2.
\tag{A.14}
$$

Next, we expand external $\exp(\boldsymbol{\cdot})$ into Taylor series, obtaining

$$
\begin{aligned}
&\frac{C}{T^2}\int_0^T \mathrm{d}\tau_1 \int_0^T \mathrm{d}\tau_2 \left(1+\sum_{n=1}^\infty \frac{1}{n!}\left(\int_{\mathbb{R}} \mathrm{d}\tilde{\sigma}(\omega)\mathrm{e}^{\mathrm{i}\omega(\tau_1-\tau_2)}\right)^n\right)\mathrm{e}^{\mathrm{i}\Omega(\tau_1-\tau_2)}\\
&\quad=\frac{C}{T^2}\int_0^T \mathrm{d}\tau_1 \int_0^T \mathrm{d}\tau_2 \left(1+\sum_{n=1}^\infty \frac{1}{n!}\int_{\mathbb{R}} \mathrm{d}\tilde{\sigma}^{*n}(\omega)\mathrm{e}^{\mathrm{i}\omega(\tau_1-\tau_2)}\right)\mathrm{e}^{\mathrm{i}\Omega(\tau_1-\tau_2)},
\end{aligned}
\tag{A.15}
$$

where $\tilde{\sigma}^{*n}$ is n-fold convolution power of $\tilde{\sigma}$. This Taylor series is uniformly convergent. We commute limit $T\to\infty$ with the sum and calculate the integrals; for term $n=1$ we have

$$
\frac{1}{T^2}\int_0^T \mathrm{d}\tau_1 \int_0^T \mathrm{d}\tau_2 \mathrm{e}^{\mathrm{i}\Omega(\tau_1-\tau_2)}=\frac{1}{T^2}\frac{1}{\Omega^2}\left|\mathrm{e}^{\mathrm{i}\Omega T}-1\right|^2 \xrightarrow{T\to\infty} 0,
\tag{A.16}
$$

where the assumption $\Omega\neq 0$ is crucial. For any other term

$$
\begin{aligned}
&\lim_{T\to\infty}\frac{1}{T^2}\int_0^T \mathrm{d}\tau_1 \int_0^T \mathrm{d}\tau_2 \int_{\mathbb{R}} \mathrm{d}\tilde{\sigma}^{*n}(\omega)\mathrm{e}^{\mathrm{i}(\omega+\Omega)(\tau_1-\tau_2)}\\
&\quad=\lim_{T\to\infty}\frac{1}{T^2}\int_{\mathbb{R}} \mathrm{d}\tilde{\sigma}^{*n}(\omega-\Omega)\int_0^T \mathrm{d}\tau_1 \int_0^T \mathrm{d}\tau_2 \mathrm{e}^{\mathrm{i}\omega(\tau_1-\tau_2)}\\
&\quad=2\lim_{T\to\infty}\int_{\mathbb{R}} \mathrm{d}\tilde{\sigma}^{*n}(\omega-\Omega)\frac{1-\cos(\omega T)}{(\omega T)^2}.
\end{aligned}
\tag{A.17}
$$

In the list line one recognise the functional which returns the jump of the measure $\tilde{\sigma}^{*n}$ at point Ω. But, the measure σ is continuous and $\tilde{\sigma}^{*n}$ is also continuous; the result is $\tilde{\sigma}^{*n}(\{\Omega\})=0$. ■

Proof of Theorem 6. Combining (A.8) and (A.9) we obtain the formula of time-averaged characteristic function of finite-dimensional distribution in general case, which is

$$
\phi_{\theta_1,\ldots,\theta_L}\mathrm{e}^{\mathrm{i}\lambda_0 X_0}\sum_{S\in G_N}\exp\left(\mathrm{i}\sum_{m_j\in S} m_j\tilde{\Theta}_j\right)\prod_{m_j\in S}\mathrm{i}^{m_j}J_{m_j}(\lambda_j R_j);
\tag{A.18}
$$

here G_N are all N-element subsets of integers m_j for which $\sum_{j=1}^N m_j\omega_j = 0$. What is left is to show that the above quantity equals the conditional ensemble-average characteristic function

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\left(\sum_{l=1}^L \theta_l X_{\mathrm{erg}}(t+t_l)+\sum_{j=1}^N \lambda_j R_j \cos(\tilde{\Theta}_j+\omega_j t)+\lambda_0 X_0\right)}|X_0,\{R_k\},\mathcal{M}\right]$$
$$= \phi_{\theta_1,\dots,\theta_L}\mathrm{e}^{\mathrm{i}\lambda_0 X_0}\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\left(\sum_{j=1}^N \lambda_j R_j \cos(\tilde{\Theta}_j+\omega_j t)\right)}|\{R_k\},\mathcal{M}\right]. \tag{A.19}$$

The factor $\phi_{\theta_1,\dots,\theta_L}\mathrm{e}^{\mathrm{i}\lambda_0 X_0}$ already agrees, so we will omit it later on. Next, we expand the remaining expected value using Jacobi–Anger identity (A.7), obtaining

$$\sum_{S\in M_N}\prod_{m_j\in S}\mathrm{i}^{m_j}J_{m_j}(\lambda_j R_j)\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\sum_{m_k\in S} m_k(\tilde{\Theta}_k+\omega_k t)}|\mathcal{M}\right]. \tag{A.20}$$

Because $\theta\mapsto \mathrm{e}^{\mathrm{i}\theta}$ is injection on $[0,2\pi)$ and $\tilde{\Theta}_k$ differ from Θ_k only by a deterministic constants, the σ-algebra $\sigma(\mathcal{M})$ is equivalent to $\sigma(\tilde{\mathcal{M}})$ generated by

$$\tilde{\mathcal{M}} = \left\{\exp\left(\sum_{m_j} m_j\tilde{\Theta}_j\right) : \sum_j m_j\omega_j = 0\right\}. \tag{A.21}$$

For terms with $S\in G_N\subset M_N$ the random phases in (A.20) are $\tilde{\mathcal{M}}$-, therefore also $\mathcal{M}$-measurable, moreover they agree with the corresponding terms in (A.18). What is left is to show that the expected value of the remaining terms for $S\notin G_N$ is zero.

Now, for any incommensurable ω_j the corresponding $m_j = 0$. Commensurable ω_j's can be divided into subsets of jointly commensurable numbers, i.e. into blocks $\{\omega_{k_i}\}$ for which $\omega_{k_i} = \alpha q_{k_i}/p_{k_i}$, $q_{k_i}\in\mathbb{Z}$, $p_{k_i}\nmid q_{k_i}$; different blocks have different incommensurable factors α. Each such block corresponds to different subset of independent $\tilde{\Theta}_j$, therefore they can be considered separately.

Let us choose one such block and for simplicity of notation, change indices such that these are $\{\omega_1,\omega_2\dots,\omega_r\}$. The condition $\sum_{j=1}^r m_j\omega_j = 0$ is equivalent to condition $\sum_{j=1}^r m_j\eta_j = 0$, where η_j are relatively prime integers obtained by multiplying ω_j by the least common multiple of p_j's. The equation

$$\sum_{j=1}^r m_j\eta_j = 0 \tag{A.22}$$

has exactly $r-1$ linearly independent solutions in integers [39]. For our one chosen block let us name these solutions $\{m_j^1\},\{m_j^2\},\dots,\{m_j^{r-1}\}$. Any other solution is a linear combination of the elementary solutions

$$m_j = \sum_{\rho=1}^{r-1}\nu_\rho m_j^\rho,\quad \nu_j\in.\mathbb{Z}. \tag{A.23}$$

Therefore

$$\sum_{j=1}^r m_j\tilde{\Theta}_j = \sum_{j=1}^r\sum_{\rho=1}^{r-1}\nu_\rho m_j^\rho\tilde{\Theta}_j = \sum_{\rho=1}^{r-1}\nu_\rho\sum_{j=1}^r m_j^\rho\tilde{\Theta}_j, \tag{A.24}$$

and for each block $\mathcal{M}$ depends actually only on $r-1$ variables $\Xi_\rho = \sum_{j=1}^r m_j^\rho\tilde{\Theta}_j \pmod{2\pi}$, the rest of the variables are linear combinations of the elementary ones. For all blocks together $\mathcal{M}$ depends on at most $N-1$ such variables.

The factor in the studied conditional expectancy corresponding to the chosen block is

$$\mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=1}^r m_j\tilde{\Theta}_j\right)|\{\Xi_\rho\}_{\rho=1}^{r-1}\right],\quad m_j\in S. \tag{A.25}$$

Because $m_j \in S \notin G_N$ the sum $\sum_{j=1}^{r} m_j \tilde{\Theta}_j \pmod{2\pi}$ is linearly independent of the set $\{\Xi_\rho\}_{\rho=1}^{r-1}$. We prove in the lemma below that it implies that this sum is also probabilistically independent of $\{\Xi_\rho\}_{\rho=1}^{r-1}$ and has uniform distribution on $[0, 2\pi)$. Therefore

$$\mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=1}^{r} m_j\tilde{\Theta}_j\right)\Big|\{\Xi_\rho\}_{\rho=1}^{r-1}\right] = \mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=1}^{r} m_j\tilde{\Theta}_j\right)\right]$$
$$= \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\Theta'}\right] = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\theta\, e^{\mathrm{i}\theta} = 0, \quad \Theta' \sim \mathcal{U}(0, 2\pi). \tag{A.26}$$

We have proven that all elements in the sum (A.20) which contain combinations of Θ_j linearly independent of elements of $\mathcal{M}$ are zero. Only $\mathcal{M}$-dependent elements remain, which exactly agrees with the time-average characteristic function (A.18). This concludes the proof. ■

Lemma 2. *For i.i.d. $\{\Theta_j\}_{j=1}^{N}$, $\Theta_j \sim \mathcal{U}(0, 2\pi)$ any N linearly independent integer combinations*

$$\Xi_i = \sum_{j=1}^{N} m_{ij}\Theta_j \pmod{2\pi}, \quad m_{ij} \in \mathbb{Z} \tag{A.27}$$

are set of jointly independent random variables with distribution $\Xi_i \sim \mathcal{U}(0, 2\pi)$.

Proof. Because we work in modulo 2π arithmetic, all variables can be considered to have values in torus $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ and $\mathcal{U}(0, 2\pi) \equiv \mathcal{U}(\mathbb{T})$. The continuous dual of $\mathbb{T}^N$ is $\mathbb{Z}^N$ and thus a natural space of parameters of characteristic function of the vector $(\Xi_i)_{i=1}^{N}$.

For uniform distribution on torus the characteristic function has a very simple form: it is the Kronecker delta

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}k\Theta_j}\right] = \delta_k, \quad k \in \mathbb{Z}. \tag{A.28}$$

It is clear if we think about characteristic function as a Fourier series of density $1/(2\pi)$ on $\mathbb{T}$. We will show that the multidimensional characteristic function of $(\Xi_i)_{i=1}^{N}$ is the product $\delta_{k_1}\ldots\delta_{k_N}$ which corresponds to distribution $\mathcal{U}(\mathbb{T}^N)$, that is i.i.d. uniform variables Ξ_i.

Let us choose any $k_1, \ldots, k_N \in \mathbb{Z}$ and consider $\sum_i k_i\Xi_i$. We calculate characteristic function

$$\mathbb{E}\left[\exp\left(\mathrm{i}\sum_{i=1}^{N} k_i\Xi_i\right)\right] = \mathbb{E}\left[\exp\left(\mathrm{i}\sum_{i=1}^{N} k_i\sum_{j=1}^{N} m_{ij}\Theta_j\right)\right]$$
$$= \mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=1}^{N}\Theta_j\sum_{i=1}^{N} k_i m_{ij}\right)\right] = \prod_{j=1}^{N}\mathbb{E}\left[\exp\left(\mathrm{i}\Theta_j\sum_{i=1}^{N} k_i m_{ij}\right)\right]$$
$$= \prod_{j=1}^{N}\delta_{\sum_{i=1}^{N} k_i m_{ij}}. \tag{A.29}$$

The above product equals 1 if, and only if for all j we have $\sum_{i=1}^{N} k_i m_{ij} = 0$. In all other cases it equals 0. But, the linear integer combinations (A.27) are linearly independent which is equivalent to saying that this is true if, and only if $k_1 = k_2 = \cdots = k_N = 0$. So the above formula must be exactly $\delta_{k_1}\ldots\delta_{k_N}$ which was to be demonstrated. ■

Proposition 2. *For any stationary Gaussian process X, the time-average covariance structure is the ensemble-average structure conditioned by σ-algebra $\sigma(X_0, \{R_k\}_k)$, where R_k are the amplitudes of the atoms of the spectral measure and X_0 is the constant term. The result is true even for rationally commensurable frequencies of spectral points.*

Proof. We begin with calculating the conditional ensemble-average covariance. The conditional mean equals

$$\mathbb{E}\left[X(t)|X_0,\{R_k\}\right]=X_0,\quad \text{as } \mathbb{E}\left[\cos(\Theta_k)\right]=0, \text{ and } \mathbb{E}\left[X_{\mathrm{erg}}(t)\right]=0. \tag{A.30}$$

Next we fix t_1, t_2 and use independence of X_{erg} and X_k's.

$$\begin{aligned}
&\mathbb{E}\Big[(X(t_1)-X_0)(X(t_2)-X_0)|X_0,\{R_k\}\Big]\\
&\quad=\mathbb{E}\Big[X_{\mathrm{erg}}(t_1)X_{\mathrm{erg}}(t_2)|X_0,\{R_k\}\Big]+\sum_{k=1}^{\infty}\mathbb{E}\Big[X_k(t_1)X_k(t_2)|X_0,\{R_k\}\Big]\\
&\quad=\mathbb{E}\Big[X_{\mathrm{erg}}(t_1)X_{\mathrm{erg}}(t_2)\Big]+\sum_{k=1}^{\infty}R_k^2\mathbb{E}\Big[\cos(\Theta_k+\omega_k t_1)\cos(\Theta_k+\omega_k t_2)\Big]\\
&\quad=r(t_2-t_1)+\sum_{k=1}^{\infty}R_k^2\frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}x\ \cos(x+\omega_k t_1)\cos(x+\omega_k t_2)\\
&\quad=r(t_2-t_1)+\frac{1}{2}\sum_{k=1}^{\infty}R_k^2\cos(\omega_k(t_2-t_1)).
\end{aligned}\tag{A.31}$$

The time-average consists of integral from the parts $X_{\mathrm{erg}}(t_i+\tau)X_{\mathrm{erg}}(t_j+\tau)$, $X_k(t_i+\tau)X_k(t_j+\tau)$, $X_{k_1}(t_i+\tau)X_{k_2}(t_j+\tau)$, $k_1\neq k_2$ and $X_{\mathrm{erg}}(t_i+\tau)X_k(t_j+\tau)$, $i,j\in\{1,2\}$; we call them I_1, I_2, I_3, I_4, respectively. All sums are absolutely convergent: we can commute summation, integration and taking limit $T\to\infty$.

Time-average $T^{-1}I_1$ converges to $r(t_2-t_1)$ because X_{erg} is ergodic. For $T^{-1}I_2$ we have

$$\begin{aligned}
\frac{1}{T}I_4&=\frac{1}{T}R_k^2\int_0^T\mathrm{d}\tau\ \cos(\tau+\omega_k t_i)\cos(\tau+\Theta_k+\omega_k t_j)\\
&=\frac{1}{2}R_k^2\cos(\omega_k(t_2-t_1))\\
&\quad+\frac{1}{\omega_k T}\cos(2\Theta_k+\omega_k(t_1+t_2+T))\sin(\omega_k T)\xrightarrow{T\to\infty}\frac{1}{2}R_k^2\cos(\omega_k(t_2-t_1)).
\end{aligned}\tag{A.32}$$

Therefore we need to prove that $T^{-1}I_3$ and $T^{-1}I_4$ decay to 0. For $T^{-1}I_3$ it is straightforward, denoting $\omega_\pm:=\omega_{k_2}\pm\omega_{k_1}$, $\Theta_\pm:=\Theta_{k_2}\pm\Theta_{k_1}$ we get

$$\begin{aligned}
\frac{1}{T}I_3&=\frac{1}{2T\omega_-}\Big(\sin(\omega_{k_1}t_i-\omega_{k_2}t_j-\Theta_-)-\sin(\omega_{k_1}t_i-\omega_{k_2}t_j-\Theta_--T\omega_-)\Big)\\
&\quad+\frac{1}{2T\omega_+}\Big(-\sin(\omega_{k_1}t_i+\omega_{k_2}t_j+\Theta_+)-\sin(\omega_{k_1}t_i+\omega_{k_2}t_j+\Theta_++T\omega_+)\Big)\xrightarrow{T\to\infty}0.
\end{aligned}\tag{A.33}$$

As for $T^{-1}I_4$, we will show that time-average of $X_{\mathrm{erg}}(t_i+\tau)R_k\exp(\mathrm{i}\Theta_k+\mathrm{i}\omega_k\tau)$ converges, which is equivalent condition. The factor $R_k\exp(\mathrm{i}\Theta_k)$ can be brought outside integral, therefore only showing convergence of time-average of $X_{\mathrm{erg}}(t_i+\tau)\exp(\mathrm{i}\omega_k\tau)$ is required. The latter is

$$\frac{1}{T}\int_0^T\mathrm{d}\tau X_{\mathrm{erg}}(\tau)\mathrm{e}^{\mathrm{i}\omega_k\tau}. \tag{A.34}$$

The limit $T\to\infty$ of the above formula exists almost surely, argument is the same as at the beginning of the lemma. We will prove it is 0. Let us calculate $\mathbb{E}|\cdot|^2$ of (A.34)

$$\begin{aligned}
&\frac{1}{T^2}\int_0^T\mathrm{d}\tau_1\int_0^T\mathrm{d}\tau_2\ r(\tau_2-\tau_1)\mathrm{e}^{\mathrm{i}\omega_k(\tau_2-\tau_1)}\\
&\quad=\frac{1}{T^2}\int_0^T\mathrm{d}\tau_1\int_0^T\mathrm{d}\tau_2\int_{\mathbb{R}}\mathrm{d}\sigma(\omega)\mathrm{e}^{\mathrm{i}\omega(\tau_2-\tau_1)}\mathrm{e}^{\mathrm{i}\omega_k(\tau_2-\tau_1)}
\end{aligned}$$

$$= \frac{1}{T^2}\int_{\mathbb{R}} \mathrm{d}\sigma(\omega)\int_0^T \mathrm{d}\tau_1\int_0^T \mathrm{d}\tau_2 \mathrm{e}^{\mathrm{i}(\omega+\omega_k)(\tau_2-\tau_1)}$$
$$= 2\int_{\mathbb{R}} \mathrm{d}\sigma(\omega-\omega_k)\frac{1-\cos(\omega T)}{(\omega T)^2} \xrightarrow{\mathrm{T}\to\infty} \sigma(\{\omega_k\}) = 0.$$

That shows mean-square and almost sure convergence to 0. ■

Appendix B. Other auxiliary results

Proposition 3. *For numerically approximated covariance function of the Cantor process*

$$\tilde{r}(t) := \prod_{k=1}^{N} \cos\left(3^{-k}t\right) \tag{B.1}$$

the above value converges to the cantor covariance function r_C and for $N \geq \log_3 t$ is bounded by

$$(1+ct^2 9^{-N})r_C(t) \leq \tilde{r}(t) \leq \mathrm{e}^{dt^2 9^{-N}} r_C(t), \quad c, d > 0 \tag{B.2}$$

for t's where it is positive, and the reverse inequality for t's where it is negative.

Proof. For $N \geq \log_3 t$ we have $x = 3^{-(N+j)}t \leq \pi/2$, $\cos\left(3^{-(N+j)}t\right) > 0$ and

$$1 - x^2/2 \leq \cos(x) \leq 1 - 4/\pi^2 x^2. \tag{B.3}$$

Expressing $\tilde{r}$ in terms of r_C

$$\tilde{r}(t) = r_C(t)\prod_{j=1}^{\infty}\frac{1}{\cos\left(3^{-(N+j)}t\right)} \tag{B.4}$$

we obtain for area of positive values of r_C and $\tilde{r}$

$$r_C(t)\prod_{j=1}^{\infty}\frac{1}{1-3^{-2(N+j)}t^2 4/\pi^2} \leq \tilde{r}(t) \leq r_C(t)\prod_{j=1}^{\infty}\frac{1}{1-3^{-2(N+j)}t^2/2}. \tag{B.5}$$

We need to approximate product of terms

$$\frac{1}{1-a9^{-N}9^{-j}} = 1 + \frac{1}{a^{-1}9^N 9^j - 1} = 1 + 9^{-N}\frac{1}{a^{-1}9^j - 9^{-N}} = 1 + p_j. \tag{B.6}$$

To use the inequality valid for positive p_j's

$$1 + \sum_{j=1}^{\infty} p_j \leq \prod_{j=1}^{\infty}(1+p_j) \leq \exp\left(\sum_{j=1}^{\infty} p_j\right) \tag{B.7}$$

we estimate the sum of p_j's by

$$\sum_{j=1}^{\infty} p_j \geq a9^{-N}\sum_{j=1}^{\infty}\frac{1}{9^j} \geq a9^{-N}\frac{1}{8} \tag{B.8}$$

from the bottom, and writing

$$\sum_{j=1}^{\infty} p_j = 9^{-N}\frac{1}{a^{-1}-9^{-N}} + 9^{-N}\sum_{j=1}^{\infty}\frac{1}{9a^{-1}9^j - 9^{-N}} \leq \frac{1}{9^N a^{-1} - 1} + \frac{1}{9}\sum_{j=1}^{\infty} p_j \tag{B.9}$$

we estimate

$$\sum_{j=1}^{\infty} p_j \leq \frac{9}{8}\frac{1}{9^N a^{-1} - 1} \tag{B.10}$$

from the top. Substituting proper a we obtain the final result

$$\left(1+\frac{t^2}{2\pi^2}9^{-N}\right) r_C(t) \le \tilde{r}(t) \le \exp\left(\frac{9}{8}\frac{1}{2t^{-2}-9^{-N}}9^{-N}\right) r_C(t). \tag{B.11}$$

For the area with negative r_C, $\tilde{r}$ the above inequality is reversed. ■

Proposition 4. *For α-stable harmonisable process*

$$X(t) = \int_{\mathbb{R}} \mathrm{d}S(\omega)\, \mathrm{e}^{\mathrm{i}\omega t}, \tag{B.12}$$

where the spectral process S has control measure σ_α, the general codifference function

$$\tau_{\theta_1,\theta_2}(s,t) := \ln \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\theta_1 X(s)}\right] + \ln \mathbb{E}\left[\mathrm{e}^{\mathrm{i}\theta_2 X(t)}\right] - \ln \mathbb{E}\left[\mathrm{e}^{\mathrm{i}(\theta_1 X(s)+\theta_2 X(t))}\right] \tag{B.13}$$

equals

$$\tau_{\theta_1,\theta_2}(s,t) = \int_{\mathbb{R}} \sigma_\alpha(\mathrm{d}\omega) \left|\theta_1 + \theta_2 \mathrm{e}^{\mathrm{i}\omega(t-s)}\right|^\alpha - \left(|\theta_1|^\alpha + |\theta_2|^\alpha\right) \int_{\mathbb{R}} \sigma_\alpha(\mathrm{d}\omega). \tag{B.14}$$

Proof. We use the following formula [13]

$$\mathbb{E}\left[\exp\left(\mathrm{i}\int_{\mathbb{R}} \mathrm{d}S(\omega) f(\omega)\right)\right] = \exp\left(-\int_{\mathbb{R}} \sigma_\alpha(\omega)\, |f(\omega)|^\alpha\right) \tag{B.15}$$

and note that

$$\theta_1 X(s) + \theta_2 X(t) = \int_{\mathbb{R}} \mathrm{d}S(\omega)\ \left(\theta_1 \mathrm{e}^{\mathrm{i}\omega s} + \theta_2 \mathrm{e}^{\mathrm{i}\omega t}\right). \tag{B.16}$$

Therefore

$$\begin{aligned} \mathbb{E}\left[\mathrm{e}^{\mathrm{i}(\theta_1 X(s)+\theta_2 X(t))}\right] &= \exp\left(-\int_{\mathbb{R}} \sigma_\alpha(\mathrm{d}\omega) \left|\theta_1 \mathrm{e}^{\mathrm{i}\omega s} + \theta_2 \mathrm{e}^{\mathrm{i}\omega t}\right|^\alpha\right) \\ &= \exp\left(-\int_{\mathbb{R}} \sigma_\alpha(\mathrm{d}\omega) \left|\theta_1 + \theta_2 \mathrm{e}^{\mathrm{i}\omega(t-s)}\right|^\alpha\right) \end{aligned} \tag{B.17}$$

and

$$\begin{aligned} \tau_{\theta_1,\theta_2}(s,t) &= |\theta_1|^\alpha \int_{\mathbb{R}} \sigma_\alpha(\mathrm{d}\omega) + |\theta_2|^\alpha \int_{\mathbb{R}} \sigma_\alpha(\mathrm{d}\omega) - \int_{\mathbb{R}} \sigma_\alpha(\mathrm{d}\omega) \left(\left|\theta_1 \mathrm{e}^{\mathrm{i}\omega s} + \theta_2 \mathrm{e}^{\mathrm{i}\omega t}\right|^\alpha\right) \\ &= \int_{\mathbb{R}} \sigma_\alpha(\mathrm{d}\omega) \left|\theta_1 + \theta_2 \mathrm{e}^{\mathrm{i}\omega(t-s)}\right|^\alpha - \left(|\theta_1|^\alpha + |\theta_2|^\alpha\right) \int_{\mathbb{R}} \sigma_\alpha(\mathrm{d}\omega). \qquad ■ \end{aligned} \tag{B.18}$$

References

[1] R. Metzler, Int. J. Mod. Phys. Conf. Ser. 36 (2015) 1560007. http://dx.doi.org/10.1142/S2010194515600071.
[2] R. Metzler, J.-H. Jeon, A.G. Cherstvy, E. Barkai, Phys. Chem. Chem. Phys. 16 (44) (2014) 24128–24164. http://dx.doi.org/10.1039/C4CP03465A.
[3] J.L. Doob, Stochastic Processes, Wiley-Interscience, 1990.
[4] H. Dym, H.P. McKean, Gaussian Processes, Function Theory and the Inverse Spectral Problem, Academic Press, 1976.
[5] S. Janson, Gaussian Hilbert Spaces, Cambridge University Press, 1997.
[6] U. Grenander, Ark. Mat. 1 (1950) 195–277.
[7] G. Maruyama, Theory Probab. Appl. 15 (1970) 1–22.
[8] A.M. Yaglom, Correlation Theory of Stationary and Related Random Functions, Springer, 1987.
[9] R. Fox, Phys. Rep. 48 (3) (1978) 179–283.
[10] R. Lyons, J. Fourier Anal. Appl. Special Issue (1995) 363–377.
[11] R. Salem, Algebraic Numbers and Fourier Analysis, Heath, 1963.
[12] B. Mandelbrot, Gaussian Self-Affinity and Fractals, Springer, 2002.

[13] G. Samorodnitsky, M. Taqqu, Stable Non-Gaussian Random Processes, Chapman & Hall, 1994.
[14] P. Embrechts, M. Mejima, Selfsimilar Processes, Princeton University Press, 2002.
[15] B.W. Silverman, Density Estimation for Statistics and Data Analysis, Chapman and Hall, 1986.
[16] A. Postnikov, Tauberian Theory and Its Applications, American Mathematical Society, 1980.
[17] T. Eisner, B. Farkas, M. Haase, R. Nagel, Operator Theoretic Aspects of Ergodic Theory, Springer, 2015.
[18] R. Barbuti, G. Caravagna, A. Maggiolo-Schettini, P. Milazzo, in: C. Priami, R.-J. Back, I. Petre, E. de Vink (Eds.), Transactions on Computational Systems Biology XIII, Springer, 2011, pp. 61–84.
[19] D. Champeney, A Handbook of Fourier Theorems, Cambridge University Press, 1987.
[20] P. Brockwell, R. Davis, Time Series: Theory and Methods, Springer, 2006.
[21] I.M. Gel'fand, N.Y. Vilenkin, Generalized Functions: Applications of Harmonic Analysis, Academic Press, 1964.
[22] R. Zwanzig, Nonequlibrium Statistical Mechanics, Oxford University Press, 2001.
[23] S.C. Kou, Ann. Appl. Stat. 2 (2) (2008) 501–535.
[24] V. Anh, R. McVinish, J. Appl. Math. Stoch. Anal. 16 (2) (2003) 97–119.
[25] A. Viñales, M. Despósito, Phys. Rev. E 73 (2006) 016111. http://dx.doi.org/10.1103/PhysRevE.73.016111.
[26] H. Mori, Progr. Theoret. Phys. 33 (3) (1965) 423–455.
[27] R. Kubo, Rep. Progr. Phys. 29 (1966) 255–284.
[28] A. Caldeira, Scholarpedia 5 (2) (2010) 9187. http://dx.doi.org/10.4249/scholarpedia.9187.
[29] U. Weiss, Quantum Dissipative Systems, Word Scientific, 2008.
[30] L. Rey-Bellet, in: S. Attal, A. Joye, C.-A. Pillet (Eds.), Quantum Dissipative Systems II, Springer, 2006, pp. 41–78.
[31] Y. Katznelson, An Introduction to Harmonic Analysis, Cambridge University Press, 2004.
[32] A. Janicki, A. Weron, Simulations and Chaotic Behaviour, Marcel Dekker, 1994.
[33] M. Magdziarz, Theory Probab. Appl. 54 (2) (2009) 322–324. http://dx.doi.org/10.1137/S0040585X97984267.
[34] J. Rosiński, T. Żak, Theory Probab. Appl. 61 (2) (1996) 277–288. http://dx.doi.org/10.1016/0304-4149(95)00083-6.
[35] M. Magdziarz, Stochastic Process. Appl. 119 (2009) 3416–3434. http://dx.doi.org/10.1016/j.spa.2009.06.002.
[36] T. Marquardt, Bernoulli 12 (6) (2006) 1099–1126. http://dx.doi.org/10.3150/bj/1165269152.
[37] K. Sobczyk, Stochastic Differential Equations: With Applications to Physics and Engineering, Springer, 2013.
[38] M. Magdziarz, A. Weron, Phys. Rev. Lett. 105 (2010) 260603. http://dx.doi.org/10.1103/PhysRevLett.105.260603.
[39] E. Kimme, SIAM Rev. 7 (1) (1965) 88–99.