1. Stationary Gaussian processes

Below T will denote $\mathbb{R}^d$ or $\mathbb{Z}^d$. What is special about these index sets is that they are (abelian) groups. If $X = (X_t)_{t\in T}$ is a stochastic process, then its translate X^τ is another stochastic process on T defined as $X^\tau(t) = X(t-\tau)$. The process X is called *stationary* (or translation invariant) if $X^\tau \stackrel{d}{=} X$ for all $\tau \in T$.

Let X be a Gaussian process on T with mean $M : T \to \mathbb{R}$ and covariance $K : T \times T \to \mathbb{R}$. It is an easy exercise to see that X is stationary if and only if M is a constant and $K(t,s)$ depends only on $t-s$. In this case we usually write the covariance as $K(t-s)$ for an even function $K : T \to \mathbb{R}$.

Covariance function: What functions $K : T \to \mathbb{R}$ are possible covariance functions? Clearly K is a valid covariance function if and only if $(t,s) \to K(t-s)$ is p.s.d. That is, $\sum_{i,j=1}^n a_i a_j K(t_i - t_j) \geq 0$ for all $n \geq 1$ and $t_i \in T$ and $a_i \in \mathbb{R}$. Such a function is said to be positive (semi)definite.

Example 1. Let $W = (W_t)_{t\in[0\infty)}$ be a standard Brownian motion in one dimension. Define $X(t) = e^{-t/2}W(e^t)$ for $t \in \mathbb{R}$. Then X is clearly Gaussian, has zero mean and $\mathbf{E}[X_t X_s] = e^{-|t-s|/2}$ (check!). Since this depends only on the difference between t and s, it follows that X is a stationary Gaussian process with $K(u) = e^{-|u|/2}$. This is known as the *Ornstein-Uhlenbeck process*. Up to scaling, it is the only stationary Gaussian Markov process!

Example 2. Let ξ_k be i.i.d. $N(0,1)$. Then $\xi = (\xi_k)_{k\in\mathbb{Z}}$ is a stationary GP on $\mathbb{Z}$ with $K(n) = \delta_{n,0}$. If $a \in \ell^2(\mathbb{Z})$ is a fixed sequence, we can define a Gaussian process by $X = a * \xi$ or more explicitly $X_n := \sum_{k\in\mathbb{Z}} a_{n-k}\xi_k$.

Since $a \in \ell^2(\mathbb{Z})$, for each n the series defining X_n converges a.s. Taking a countable intersections, the entire collection $(X_n)_{n\in\mathbb{Z}}$ is well-defined, a.s. It is centered and has covariance $\mathbf{E}[X_n X_m] = \sum_k a_{n-k}a_{m-k} = \sum_k a_{n-m+j}a_j$ which clearly depends only on $n-m$. Hence, X is a stationary GP on $\mathbb{Z}$. They are often called moving-average processes.

Spectral measure: How to generate positive definite functions? Bochner's theorem says that they can be parameterized by finite Borel measures (always measures will mean positive measures) on an appropriate space. Let $\hat{T} = \mathbb{R}^d$ if $T = \mathbb{R}^d$ and let $\hat{T} = (-\pi,\pi]^d$ if $T = \mathbb{Z}^d$. For a finite Borel measure μ on $\hat{T}$, its Fourier transform is the function $\hat{\mu} : T \to \mathbb{C}$ defined as $\hat{\mu}(t) = \int_{\hat{T}} e^{i\lambda\cdot t} d\mu(\lambda)$.

Theorem 3 (Bochner)**.** *The mapping $\mu \to \hat{\mu}$ is a bijection from the space of finite measures on $\hat{T}$ to the space of all continuous complex-valued positive definite functions on T.*

In this correspondence, continuous real-valued positive definite functions on T correspond to symmetric finite measures on $\hat{T}$.

One way implication in this theorem is easy. If μ is a finite measure on $\hat{T}$, then $\int_{\hat{T}} |\sum_{j=1}^n a_j e^{it_j\cdot\lambda}|^2 d\mu(\lambda)$ must be non-negative. When expanded, it follows that $\hat{\mu}$ is a positive definite function. The other way, starting with a positive definite function K and constructing the corresponding measure, is non-trivial and we omit its proof.

If X is a centered Gaussian process with covariance K, then the unique measure μ on $\hat{T}$ such that $K = \hat{\mu}$ is called the *spectral measure* of the process X. The distribution of the Gaussian process is completely determined by the spectral measure. We shall see how the smoothness of the sample paths of the process or the ergodicity and mixing properties are reflected in the properties of the spectral measure. For now a simpler observation.

Exercise 4. Let X and Y be independent centered stationary Gaussian processes with spectral measures μ and ν respectively. Show that $Z = aX + bY$ is a stationary Gaussian process with spectral measure $|a|\mu + |b|\nu$.

This exercise can be used to decompose a given process into independent processes, for example by separating the atomic and continuous parts of the spectral measure, or separating the high frequency and low frequency parts of the spectral measure etc.

Some examples.

Example 5. $T = \mathbb{R}$. The Ornstein-Uhlenbeck process has covariance $K(t) = e^{-|t|}$ (we have scaled it to lose the factor of 2 in the exponent). It can be checked (a good exercise if you have not seen it already!) that the spectral measure is the Cauchy measure $d\mu(\lambda) = \frac{1}{\pi(1+\lambda^2)} d\lambda$ on the real line.

Example 6. Let μ be a purely atomic measure on $\mathbb{R}$ or $[-\pi,\pi]$ given as $\mu = p_0\delta_0 + \sum_k p_k(\delta_{\lambda_k} + \delta_{-\lambda_k})$ where $\sum_k p_k < \infty$. Then it is easy to see that the Gaussian process with this spectral measure can be represented as

$$X_t = \sqrt{p_0}\xi_0 + \sum_k 2\sqrt{p_k}\xi_k \cos(\lambda_k t)$$

where ξ_j are i.i.d. standard Gaussians (assume that the sums are finite if convergence bothers you). This shows that the process X is a random superposition of cosines with frequencies and amplitudes given by the spectral measure. If μ is not atomic, we don't know if an analogous representation can be written, but the interpretation of the spectral measure is the same.

Example 7. $T = \mathbb{Z}$ and $\mu = \text{unif}[-\pi,\pi]$. Then the covariance is $K(n) = \delta_{n,0}$, i.e., the corresponding Gaussian process is the i.i.d. process.

Example 8. $T = \mathbb{R}^2$ and let μ be the uniform measure on the unit circle $S^1 = \{(x,y) : x^2+y^2 = 1\} \subseteq \mathbb{R}^2$. Then the covariance function is $K(t) = (2\pi)^{-1}\int_0^{2\pi} e^{i(t_1\cos\theta + t_2\sin\theta)} d\theta$ for $t = (t_1,t_2) \in \mathbb{R}^2$. It is easy to see that K is radial, i.e., $K(t)$ depends only on $|t|$ and hence $K(t) = \frac{1}{2\pi}\int_0^{2\pi} e^{-i|t|\cos\theta} d\theta$ which is the definition of the Bessel function of order 0, denoted J_0. Thus, $K(t) = J_0(|t|)$. The corresponding process is known as the *random plane wave*.

To understand this further, consider the discrete measure μ_n that puts mass $1/n$ at the nth roots of unity (take n even so that μ is symmetric). Then the Gaussian process can be written as $X_n(t) = \frac{1}{\sqrt{n}}\sum_{k=1}^n \xi_k \cos(t\cdot\lambda_k)$ where $\lambda_k = (\cos(2\pi ik/n), \sin(2\pi ik/n))$. Thus, X_n is a Gaussian superposition of cosine-waves in directions given by λ_k. What is special is that the cosine waves all have the same wave-length 1 (since $|\lambda_k|^1 = 1$). Letting $n \to \infty$, X_n approaches the random plane wave, which can now be interpreted as a Gaussian superposition of cosine waves in all possible directions.

Exercise 9. Suppose $T = \mathbb{R}^d$ and μ is the uniform measure on the unit sphere S^{d-1}. Show that the corresponding covariance function is the Bessel function $J_{\frac{d}{2}-1}(|t|)$.

Exercise 10. If X is a stationary centered Gaussian process on $\mathbb{R}^d$ that is also rotation invariant (i.e., $X \circ P \stackrel{d}{=} X$ for any $P \in O(n)$). Then show that the covariance function of X must be of the form $K(u) = \int_0^\infty J_{\frac{d}{2}-1}(s|u|) d\nu(s)$ for some measure ν on $\mathbb{R}_+$.

2. Ergodicity

Let X be a stochastic on $\mathbb{R}^d$ (we shall use $d = 1$ for simplicity of notation) defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Let $\mathbb{R}^{\mathbb{R}}$, the space of all functions on $\mathbb{R}$, be endowed with the cylinder sigma-algebra $\mathcal{G}$ generated by all cylinder sets. Then X is a measurable mapping from Ω to $\mathbb{R}^{\mathbb{R}}$. Let $Q = \mathbf{P}\circ X^{-1}$ be the push-forward of $\mathbf{P}$ under X. Under the translated process X^τ, the measure $\mathbf{P}$ pushes forward to $Q\theta_\tau^{-1}$ where $\theta_\tau : \mathbb{R}^{\mathbb{R}} \to \mathbb{R}^{\mathbb{R}}$ is given by $[\theta_\tau\omega](t) = \omega(t-\tau)$.

To say that the process X is stationary is the same as saying that $Q\theta_\tau^{-1} = Q$ for all τ. That is, the group of transformations θ_τ act measure-preservingly on $(\mathbb{R}^{\mathbb{R}}, \mathcal{G}, Q)$. One can then ask about the ergodicity and mixing properties of this system. We quickly recall these notions.

A set $A \in \mathcal{G}$ is said to be invariant if $\theta_\tau^{-1} A = A$ for all τ. The collection of all invariant sets forms the invariant sigma-algebra $\mathcal{J}$. To say that X is *ergodic* means $Q(A) = 0$ or 1 for all invariant sets A. To say that X is *(weakly) mixing* means that $Q(A \cap \theta_\tau^{-1} B) \to Q(A)Q(B)$ as $\tau \to \infty$, for all events $A, B \in \mathcal{G}$. Clearly mixing implies ergodicity.

We now state conditions for the ergodicity and mixing of stationary Gaussian processes[14].

Theorem 11 (Maruyama)**.** *Let X be a centered stationary Gaussian process on $\mathbb{R}^d$ with continuous covariance kernel K and spectral measure μ. Then X is ergodic if and only if μ has no atoms.*

Theorem 12. *Let X be a centered stationary Gaussian process on $\mathbb{R}^d$ with continuous covariance kernel K and spectral measure μ. Then X is weakly mixing if and only if $K(t) = o(1)$ as $|t| \to \infty$. In particular, if μ is absolutely continuous, then X is weakly mixing.*

In proving these theorems we shall need what is perhaps the most basic theorem in ergodic theory.

Theorem 13 (Birkoff's ergodic theorem)**.** *In the above setting where θ_τ act measure preservingly on $(\mathbb{R}^{\mathbb{R}}, \mathcal{G}, Q)$,*

(1) For any $G \in L^1(Q)$, we have $\frac{1}{T}\int_0^T G(\theta_\tau \omega) d\tau \overset{a.s.,\, L^1}{\longrightarrow} \mathbf{E}_Q[G \mid \mathcal{J}]$ as $T \to \infty$.

(2) The system is ergodic if and only if $\frac{1}{T}\int_0^T G(\theta_\tau \omega) d\tau \overset{a.s.,\, L^1}{\longrightarrow} \mathbf{E}_Q[G]$ for all $G \in L^1(Q)$.

The strategy in proving Theorem 11 will be to directly show that for any $G \in L^1(Q)$, the averages

$$\frac{1}{T}\int_0^T G(\theta_\tau \omega) d\tau \overset{a.s.}{\to} \mathbf{E}_Q[G] \tag{1}$$

To prove Theorem 12 we shall show that for any bounded measurable functions G, F (here $F_\tau = F \circ \theta_\tau$) we have

$$\mathbf{E}_Q[F \cdot G_\tau] \to \mathbf{E}_Q[F]\mathbf{E}_Q[G] \text{ as } \tau \to \infty. \tag{2}$$

Applying to $F = \mathbf{1}_A$ and $G = \mathbf{1}_B$ we clearly get the mixing property.

As usual in analysis, a useful first step is to reduce the class of functions for which the statement needs to be proved.

Exercise 14. Let $\mathcal{A}$ be a dense subset of $L^1(Q)$. If (1) holds for all $G \in \mathcal{A}$ then it holds for all $G \in L^1(Q)$.

Exercise 15. For each $n \geq 1$ and any $t = (t_1, \ldots, t_n) \in T^n$ and any $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$, define $G_{n,t,a} : \mathbb{R}^{\mathbb{R}} \to \mathbb{R}$ as $G(\omega) = \exp\{i(a_1\omega(t_1) + \ldots + a_n\omega(t_n)\}$. Let $\mathcal{A}$ be the collection of all finite linear combinations of $G_{n,t,a}$ as n, t, a vary. Show that $\mathcal{A}$ is dense in $L^1(Q)$.

Exercise 16. Let $\mathcal{A}$ be as in the previous exercise. If (2) holds for all $F, G \in \mathcal{A}$, show that it also holds for all bounded measurable F, G.

The proof of the mixing theorem is somewhat simpler and we present it first. For notational simplicity we take $d = 1$ everywhere.

Proof of Theorem 12. It suffices to show (2) for $F(\omega) = \exp\{i\sum_{j=1}^n a_j \omega_{t_j}\}$ and $G(\omega) = \exp\{i\sum_{j=1}^n b_j \omega_{t_j}\}$ for some n and some $t_j \in \mathbb{R}$ and $a_j, b_j \in \mathbb{R}$ (without loss of generality we can take the same t_js by setting some a_j and b_j to 0).

[14]This presentation is taken from the book *Gaussian processes, function theory and the inverse spectral problem* of Dym and McKean.

Under the measure Q the variables ω_t are jointly Gaussian with zero mean and $\mathbf{E}[\omega_t\omega_s] = K(t-s)$. Therefore,

$$\mathbf{E}_Q[F] = \exp\left\{-\frac{1}{2}\sum_{j,k=1}^{n} a_j a_k K(t_j - t_k)\right\},$$

$$\mathbf{E}_Q[G_\tau] = \mathbf{E}_Q[G] = \exp\left\{-\frac{1}{2}\sum_{j,k=1}^{n} b_j b_k K(t_j - t_k)\right\}.$$

Next,

$$\begin{aligned}\mathbf{E}_Q[F\cdot G_\tau] &= \mathbf{E}_Q\left[\exp\left\{i\sum_{j=1}^{n} a_j\omega_{t_j} + b_j\omega_{t_j+\tau}\right\}\right]\\ &= \exp\left\{-\frac{1}{2}\sum_{j,k=1}^{n} a_j a_k K(t_j-t_k) + b_j b_k K(t_j-t_k) + a_j b_k K(t_k - t_j + \tau)\right\}\\ &= \mathbf{E}_Q[F]\mathbf{E}_Q[G]\exp\left\{-\frac{1}{2}\sum_{j,k=1}^{n} a_j b_k K(t_k - t_j + \tau)\right\}.\end{aligned}$$

If $K(u) \to 0$ as $u \to \infty$ then it is clear that (2) holds. Conversely, if $K(u)$ does not go to zero, then we can take $n = 1$ and $a_1 = b_1 = 1$ to see that mixing condition fails. This proves that the process is mixing if and only if $K(u) \to 0$ as $|u| \to \infty$.

Since $K = \hat{\mu}$, if μ is absolutely continuous, then the Riemann-Lebesgue lemma yields that $K(u) \to 0$ as $|u| \to \infty$. This proves the second part of the theorem. ■

The proof of Maruyama's theorem is similar, just a few more computations are needed and we shall invoke Birkoff's theorem.

Proof of Maruyama's theorem. Assume that *mu* has no atoms. Fix n, t, a and let $G = G_{n,t,a}$. Let Z_T be the left hand side of (1). By the exercises, it suffices to show that $Z_T \to \mathbf{E}_Q[Z_T]$ a.s.

We shall show that $\mathbf{E}_Q[|Z_T|^2] - |\mathbf{E}_Q[Z_T]|^2 \to 0$ as $T \to \infty$. Then the variance of Z_T goes to zero and hence by Chebyshev's inequality it follows that $Z_T \to \mathbf{E}_Q[Z_T]$ in probability. By Birkoff's theorem we know that Z_T has an almost-sure limit, hence that limit must be $\mathbf{E}_Q[Z_T]$.

We proceed to compute the first and second moments of Z_T. As we have already seen in the proof of the mixing theorem,

$$\mathbf{E}_Q[Z_T] = \mathbf{E}_Q[G] = \exp\left\{-\frac{1}{2}\sum_{i,j=1}^{n} a_i a_j K(t_i - t_j)\right\}.$$

Now for the second moment of Z_T. Write $|Z_T|^2$ as $Z_T\overline{Z}_T$ to get

$$\begin{aligned}\mathbf{E}_Q[|Z_T|^2] &= \frac{1}{T^2}\int_0^T\int_0^T \mathbf{E}_Q\left[\exp\left\{i\sum_{k=1}^{n} a_k(\omega(t_k+\tau) - \omega(t_k+\tau')\right\}\right] d\tau d\tau'\\ &= \frac{1}{T^2}\int_0^T\int_0^T \exp\left\{-\sum_{i,j=1}^{n} a_i a_j[K(t_i-t_j) - K(t_i - t_j + \tau - \tau')]\right\} d\tau d\tau'.\end{aligned}$$

In the last line we used the fact that $\eta_k := \omega(t_k+\tau) - \omega(t_k+\tau')$ are jointly Gaussian and $\mathbf{E}[\eta_i\eta_j] = 2K(t_i - t_j) - K(t_i - t_j + \tau - \tau') - K(t_i - t_j + \tau' - \tau)$. Two terms have reduced to one in the exponent because of the symmetry by relabeling $(i,j) \to (j,i)$ in the last summand and using $K(-t) = K(t)$.

Compare with the expression obtained for $\mathbf{E}_Q[Z_T]$ to get

$$\mathbf{E}_Q[|Z_T|^2] = |\mathbf{E}_Q[Z_T]|^2 \cdot \frac{1}{T^2} \int_0^T \int_0^T \exp\left\{ \sum_{i,j=1}^n a_i a_j K(t_i - t_j + \tau - \tau') \right\} d\tau d\tau'.$$

If we had assumed that K vanishes at ∞ then it would follow that the integral above converges to 1. But we want to find optimal conditions under which the integral converges to 1.

Therefore, we now bring in the spectral measure to write

$$\begin{aligned}
\sum_{i,j=1}^n a_i a_j K(t_i - t_j + \tau - \tau') &= \int_{\mathbb{R}} \sum_{j,k=1}^n a_j a_k e^{i\lambda(t_j - t_k + \tau - \tau')} d\mu(\lambda) \\
&= \int_{\mathbb{R}} e^{i\lambda(\tau-\tau')} \Big| \sum_{j=1}^n a_j e^{it_j\lambda} \Big|^2 d\mu(\lambda) \\
&= \hat{\nu}(\tau - \tau')
\end{aligned}$$

where $d\nu(\lambda) = |\sum_{j=1}^n a_j e^{it_j\lambda}|^2 d\mu(\lambda)$. Thus,

$$\begin{aligned}
\mathbf{E}_Q[|Z_T|^2] &= |\mathbf{E}_Q[Z_T]|^2 \cdot \frac{1}{T^2} \int_0^T \int_0^T \exp\left\{\hat{\nu}(\tau - \tau')\right\} d\tau d\tau' \\
&= |\mathbf{E}_Q[Z_T]|^2 \cdot \frac{1}{T} \int_{-T}^T \left(1 - \frac{|u|}{T}\right) \exp\left\{\hat{\nu}(u)\right\} du
\end{aligned}$$

We claim that the integral converges to 1. The arguments that follow are standard ones in Fourier analysis.

(1) Firstly, $\exp\{\hat{\nu}(u)\} = \hat{\theta}(u)$ where $\theta = \sum_{k=0}^\infty \frac{1}{k!} \nu^{*k}$ (if ν is the distribution of a random variable Y then θ is the distribution of $Y_1 + \ldots + Y_N$ where Y_i are i.i.d. copies of Y and $N \sim \text{Pois}(1)$ is independent of Y_is).

(2) The probability measure $d\gamma_T(u) = \frac{1}{T}(1 - \frac{|u|}{T}) \mathbf{1}_{[-T,T]}(u) dt$ (it is the convolution of unif$[-T/2, T/2]$ with itself) has characteristic function $\hat{\gamma}_T(\lambda) = \frac{\sin^2(\lambda T/2)}{(\lambda T/2)^2}$. Observe that $\hat{\gamma}_T(\lambda) \to 0$ for all $\lambda \neq 0$, as $T \to \infty$. In fact, $\int \hat{\gamma}_T(\lambda) dM(\lambda) \to M\{0\}$ for any measure M.

(3) Parseval's relation says that for any two finite measures θ and γ we have $\int \hat{\theta}(u) d\gamma(u) = \int \hat{\gamma}(\lambda) d\theta(\lambda)$ (just integrate $e^{i\lambda u}$ against $\nu \otimes \gamma$ in two ways).

Thus, from the third observation we have $\mathbf{E}_Q[|Z_T|^2] = |\mathbf{E}_Q[Z_T]|^2 \int \hat{\gamma}_T(\lambda) d\theta(\lambda)$. As $T \to \infty$, from the second observation we get $\int \hat{\gamma}_T(\lambda) d\theta(\lambda) \to \theta\{0\}$. By assumption, μ has no atom and hence ν has no atom either and then ν^{*n} cannot have atoms either (for any $n \geq 1$). Therefore, the only atom of θ comes from the first term and the size of that atom is 1. Therefore, $\lim_{T\to\infty} \mathbf{E}_Q[|Z_T|^2] = \lim_{T\to\infty} |\mathbf{E}_Q[Z_T]|^2$ which is exactly what we wanted to show. Thus we get ergodicity.

Conversely, assume that μ has an atom, say $\mu\{a\} = p > 0$. By symmetry of μ, it also has an atom of the same size at $-a$. Hence $\mu^{*2}\{0\} \geq p^2$. Now take $n = 1$ and $G = G_{1,0,1}$ in the previous notation, i.e., $G(\omega) = \exp\{i\omega_0\}$. The measure ν obtained earlier is equal to μ and hence $\theta\{0\} \geq 1 + \frac{1}{2}\nu^{*2}\{0\} \geq 1 + \frac{1}{2}p^2$. This implies that $\lim \mathbf{E}_Q[|Z_T|^2] > \lim |\mathbf{E}_Q[Z_T]|^2$. Therefore, $Z := \lim_{T\to\infty} Z_T$ cannot be a constant random variable. Hence the system is not ergodic. ■

The converse part is actually simple and does not rely on all the computations. Here is the argument given by Dym and McKean, which is also of interest as it related the spectral measure to the Hilbert spaces that we introduced earlier.

RKHS of a stationary Gaussian process: Let X be a centered, stationary Gaussian process with continuous covariance K and spectral measure μ. Recall that H_X is the closed span of $\{X_t : t \in \mathbb{R}^d\}$ in $L^2(\Omega, \mathcal{F}, \mathbf{P})$. Let $e_t : \mathbb{R}^d \to \mathbb{R}$ be defined as $e_t(\lambda) = e^{i\lambda \cdot t}$. The mapping $T : H_X \to L^2(\mathbb{R}^d, \mu)$ defined by $T(X_t) = e_t$ is easily checked to define a (surjective) isomorphism (Caution: We are departing from usual convention and considering complex Hilbert spaces here. In other words, we define $L^2(\mu)$ as well as H_X over the complex field. Needless to say, there is no real change or difficulty in doing this). Thus $L^2(\mu)$ is a substitute for H_X (how does $L^2(\mu)$ relate to H_K)?

Define the translations θ_τ acting on H_X by $(\theta_\tau X)_t := X(t - \tau)$. Under the isomorphism this corresponds to the multiplication operator $\theta_\tau : L^2(\mu) \to L^2(\mu)$ defined by $(\theta_\tau e_t)(\cdot) = e_{-\tau}(\cdot) e_t(\cdot)$.

Exercise 17. Let X, K and μ be as usual. Assume that μ has a positive atom at a.

(1) Note that $\mathbf{1}_a$ is a non-zero element in $L^2(\mu)$. Let $V = T^{-1}[\mathbf{1}_a]$, a non-zero element in H_X (as remarked earlier, note that V is a complex random variable).

(2) For any τ we have $[e_\tau \mathbf{1}_a](\cdot) = e^{i\tau a} e_t(\cdot)$. Deduce that $\theta_\tau(V) = e^{i\tau a} V$.

(3) Deduce that the system cannot be ergodic. [**Hint:** Consider the random variable $|V|$.]

3. Smoothness of stationary Gaussian processes

Theorem 18. *Let X be a stationary Gaussian process on $\mathbb{R}^d$ with a continuous covariance kernel and spectral measure μ. The following are equivalent. Here smooth means C^∞.*

(1) X has smooth sample paths, almost surely.

(2) $K : \mathbb{R}^d \to \mathbb{R}$ is smooth.

(3) μ has moments of all orders, i.e., $\int |\lambda|^p d\mu(\lambda) < \infty$ for all $0 < p < \infty$.

Remark 19. Even outside the stationary setting, the first two statements must be equivalent, except that K is a symmetric function on $\mathbb{R}^d \times \mathbb{R}^d$. Of course, the third one does not make sense except for stationary processes.