11 Random real zeroes: no derivatives

11a Exponential concentration in general

11a1 Definition. [1] (a) A sequence $(x_n)_n$ of real numbers is *exponentially decaying,* if

$$\exists \delta > 0,\ C < \infty\ \ \forall n\ \ |x_n| \le C \mathrm{e}^{-\delta n} \,.$$

(b) A sequence $(X_n)_n$ of random variables $X_n : \Omega_n \to \mathbb{R}$ is *exponentially concentrated at zero,* if for every $\varepsilon > 0$ the sequence of numbers $\mathbb{P}\big(|X_n| > \varepsilon \big)$ is exponentially decaying.

(c) A sequence $(X_n)_n$ of random variables $X_n : \Omega_n \to \mathbb{R}$ is *exponentially concentrated,* if there exist $x_n \in \mathbb{R}$ such that $(X_n - x_n)_n$ is exponentially concentrated at zero.

Notation:

$$(X_n)_n \in \mathrm{ExpConZero}\,; \qquad (X_n)_n \in \mathrm{ExpCon}\ .$$

Only the distributions of these X_n matter. For a sequence $(\mu_n)_n$ of probability measures on $\mathbb{R}$ we define the relations $(\mu_n)_n \in \mathrm{ExpConZero}$ and $(\mu_n)_n \in \mathrm{ExpCon}$ evidently, getting $(X_n)_n \in \mathrm{ExpConZero} \iff (\mu_n)_n \in \mathrm{ExpConZero}$ where μ_n is the distribution of X_n; and the same for ExpCon. However, the language of random variables is more appropriate in many cases below.

11a2 Exercise. (a) All exponentially decaying sequences of real numbers are a linear space.

(b) ExpConZero is a linear space (for given $(\Omega_n)_n$).

[1] Not a standard definition.

(c) Let $(X_n)_n \in \text{ExpConZero}$ and $x_n \in \mathbb{R}$. Then $(X_n - x_n)_n \in \text{ExpConZero}$ if and only if $x_n \to 0$.
Prove it.

Thus, the condition $(X_n - x_n)_n \in \text{ExpConZero}$ determines $(x_n)_n$ up to $o(1)$.

Recall that a number x is called a median of a random variable X if

$$\mathbb{P}\big(X < x\big) \le \frac{1}{2} \le \mathbb{P}\big(X \le x\big) .$$

All medians of X are in general a compact nonempty interval (often a single point). Also, x is a median of X if and only if $(-x)$ is a median of $(-X)$.

11a3 Exercise. The following three conditions are equivalent for every sequence of random variables X_n:
(a) $(X_n)_n \in \text{ExpCon}$;
(b) there exist medians x_n of X_n such that $(X_n - x_n)_n \in \text{ExpConZero}$;
(c) all medians x_n of X_n satisfy $(X_n - x_n)_n \in \text{ExpConZero}$.
Prove it.

In this sense,

$$(X_n)_n \in \text{ExpCon} \quad \text{if and only if} \quad \big(X_n - \text{Me}(X_n)\big)_n \in \text{ExpConZero} .$$

The median interval of X_n is of length $o(1)$ whenever $(X_n)_n \in \text{ExpCon}$.

Medians cannot be replaced with expectations...

11a4 Exercise. (a) ExpCon is a linear space (for given $(\Omega_n)_n$).
(b) Let $(X_n)_n, (Y_n)_n \in \text{ExpCon}$, then $\text{Me}(X_n + Y_n) = \text{Me}(X_n) + \text{Me}(Y_n) + o(1)$.
Formulate it accurately, and prove.

11a5 Exercise. ("Sandwich") Let random variables $Y_n : \Omega_n \to \mathbb{R}$ be such that for every $r > 0$ there exist $X_n, Z_n : \Omega_n \to \mathbb{R}$ satisfying

$$\begin{gathered} (X_n)_n, (Z_n)_n \in \text{ExpCon} , \\ \forall n \;\; (X_n \le Y_n \le Z_n \text{ a.s.}) , \\ \forall n \;\; \text{Me}(Z_n) - \text{Me}(X_n) \le r . \end{gathered}$$

Then $(Y_n)_n \in \text{ExpCon}$.

Prove it.

Gaussian concentration usually ensures $\mathbb{E}\,|X_n| < \infty$ (integrability) and $\operatorname{Me}(X_n) - \mathbb{E}\,X_n \to 0$. Thus, we define ExpConInt (for given Ω_n) as the set of all sequences $(X_n)_n$ where $X_n : \Omega_n \to \mathbb{R}$ are integrable, and

$$(X_n - \mathbb{E}\,X_n)_n \in \text{ExpConZero}\,.$$

This is a linear space.

11a6 Lemma. Let random variables $Y_n : \Omega_n \to \mathbb{R}$ be such that for every $\varepsilon > 0$ there exist $X_n, Z_n : \Omega_n \to \mathbb{R}$ satisfying

$$\begin{gathered}(X_n)_n, (Z_n)_n \in \text{ExpConInt}\,,\\ \forall n\ \ (\,X_n \le Y_n \le Z_n \text{ a.s.}\,)\,,\\ \forall n\ \ \mathbb{E}\,Z_n - \mathbb{E}\,X_n \le \varepsilon\,.\end{gathered}$$

Then $(Y_n)_n \in \text{ExpConInt}$.

It can be proved similarly to 11a5. However, we need a quantitative version.

First, we note that the relation $(X_n)_n \in \text{ExpConInt}$ may be reformulated as follows: there exist families $(\delta_\varepsilon)_\varepsilon$ and $(C_\varepsilon)_\varepsilon$ of numbers $\delta_\varepsilon > 0$, $C_\varepsilon < \infty$ given for $\varepsilon > 0$ such that for all n,

$$\forall \varepsilon > 0 \quad \mathbb{P}\big(|X_n - \mathbb{E}\,X_n| > \varepsilon\big) \le C_\varepsilon \mathrm{e}^{-\delta_\varepsilon n}\,.$$

Second, in order to get $\mathbb{P}\big(|Y_n - \mathbb{E}\,Y_n| > \varepsilon\big) \le C_\varepsilon \mathrm{e}^{-\delta_\varepsilon n}$ in the conclusion of "Sandwich", we require $\mathbb{P}\big(|X_n - \mathbb{E}\,X_n| > \varepsilon\big) \le C_{r,\varepsilon} \mathrm{e}^{-\delta_{r,\varepsilon} n}$ (and the same for Z_n) in the assumption; here r is the parameter denoted by r in 11a5.

The lemma below constructs δ_ε and C_ε for given $\delta_{r,\varepsilon}$ and $C_{r,\varepsilon}$. The formulas are simple, but will not be used; rather, their existence will be used.

11a7 Lemma. ("Sandwich") Let positive numbers $\delta_{r,\varepsilon}$ and $C_{r,\varepsilon}$ be given for all positive r and ε. Let random variables $Y_n : \Omega_n \to \mathbb{R}$ be such that for every $r > 0$ there exist $X_n, Z_n : \Omega_n \to \mathbb{R}$ satisfying

$$\begin{gathered}\forall n, \varepsilon \quad \mathbb{P}\big(|X_n - \mathbb{E}\,X_n| > \varepsilon\big) \le C_{r,\varepsilon} \mathrm{e}^{-\delta_{r,\varepsilon} n}\,,\\ \forall n, \varepsilon \quad \mathbb{P}\big(|Z_n - \mathbb{E}\,Z_n| > \varepsilon\big) \le C_{r,\varepsilon} \mathrm{e}^{-\delta_{r,\varepsilon} n}\,,\\ \forall n\ \ (\,X_n \le Y_n \le Z_n \text{ a.s.}\,)\,,\\ \forall n \quad \mathbb{E}\,Z_n - \mathbb{E}\,X_n \le r\,.\end{gathered}$$

Then

$$\forall n, \varepsilon \quad \mathbb{P}\big(|Y_n - \mathbb{E}\,Y_n| > \varepsilon\big) \le C_\varepsilon \mathrm{e}^{-\delta_\varepsilon n}$$

where $\delta_\varepsilon = \delta_{\varepsilon/2,\varepsilon/2}$ and $C_\varepsilon = 2C_{\varepsilon/2,\varepsilon/2}$.

11a8 Exercise. Prove Lemma 11a7.

11a9 Lemma. ("Approximation") Let integrable random variables $X_n : \Omega_n \to \mathbb{R}$ be such that for every $\varepsilon > 0$ there exist $Y_n : \Omega_n \to \mathbb{R}$ satisfying

$$(Y_n)_n \in \operatorname{ExpConInt}\,,$$
$$\text{the sequence of numbers } \mathbb{P}\big(|X_n - Y_n| > \varepsilon\big) \text{ is exponentially decaying}\,,$$
$$\forall n \quad |\mathbb{E}\, X_n - \mathbb{E}\, Y_n| \le \varepsilon\,.$$

Then $(X_n)_n \in \operatorname{ExpConInt}$.

11a10 Exercise. Prove Lemma 11a9.

Here is a quantitative version. The assumption $(Y_n)_n \in \operatorname{ExpConInt}$ is weakened (to a single $\varepsilon\dots$). The same δ_ε, C_ε are used in two assumptions, which is not a problem (just take the minimum of two δ_ε and the sum of two C_ε).

11a11 Lemma. ("Approximation") Let positive numbers δ_ε and C_ε be given for all positive ε. Let random variables $X_n : \Omega_n \to \mathbb{R}$ be such that for every $\varepsilon > 0$ there exist $Y_n : \Omega_n \to \mathbb{R}$ satisfying

$$\forall n \quad \mathbb{P}\big(|Y_n - \mathbb{E}\, Y_n| > \varepsilon\big) \le C_\varepsilon \mathrm{e}^{-\delta_\varepsilon n}\,,$$
$$\forall n \quad \mathbb{P}\big(|X_n - Y_n| > \varepsilon\big) \le C_\varepsilon \mathrm{e}^{-\delta_\varepsilon n}\,,$$
$$\forall n \quad |\mathbb{E}\, X_n - \mathbb{E}\, Y_n| \le \varepsilon\,.$$

Then

$$\forall n, \varepsilon \quad \mathbb{P}\big(|X_n - \mathbb{E}\, X_n| > \varepsilon\big) \le 2C_{\varepsilon/3} \mathrm{e}^{-\delta_{\varepsilon/3} n}\,.$$

11a12 Exercise. Prove Lemma 11a11.

11b Exponential concentration over Gaussian measures

If a function $\xi : \mathbb{R}^d \to \mathbb{R}$ is $\operatorname{Lip}(\sigma)$ for a given $\sigma > 0$ then Theorem 1a2 gives $\xi[\gamma^d] = f[\gamma^1]$ for an increasing $f : \mathbb{R} \to \mathbb{R}$, $f \in \operatorname{Lip}(\sigma)$. Let us denote by $\operatorname{GaussLip}(\sigma)$ the set of all such random variables. Clearly, $f(0)$ is the only median of ξ, and[1]

$$\mathbb{P}\big(|\xi - \operatorname{Me}(\xi)| > \varepsilon\big) = \mathbb{P}\big(|f(\zeta) - f(0)| > \varepsilon\big) \le \\ \le \mathbb{P}\big(|\zeta| > \varepsilon/\sigma\big) \le C \exp\Big(-\frac{\varepsilon^2}{2\sigma^2}\Big)$$

[1] $\zeta \sim \gamma^1$ as before.

for some absolute constant C.[1] Also, $|\mathrm{Me}(\xi) - \mathbb{E}\,\xi| = |f(0) - \int f \, \mathrm{d}\gamma^1| \le C\sigma$ for another absolute constant C.[2] It follows easily that

$$\text{(11b1)} \qquad \mathbb{P}\big(|\xi - \mathbb{E}\,\xi| > \varepsilon \big) \le C \exp\Big(-c \frac{\varepsilon^2}{2\sigma^2} \Big)$$

for some absolute constants c, C;[3] [4] and, of course,

$$\text{(11b2)} \qquad \mathbb{E}\,|\xi - \mathbb{E}\,\xi| \le C\sigma$$

for some absolute constant C.

11c Using assumption A_n

We consider the Gaussian random function $X(\cdot)$ introduced in Sect. 2a as a linear function of the independent $N(0,1)$ random variables $X_1, \dots, X_{2n}$ (via $a_1, \dots, a_N$ and $\lambda_1, \dots, \lambda_N$) under the assumption A_n (also introduced in Sect. 2a). Here is a non-probabilistic property of the linear operator $\mathbb{R}^{2n} \to L_2[0,1]$.[5]

11c1 Proposition.

$$\int_0^1 X^2(t) \, \mathrm{d}t \le \frac{C}{n} (X_1^2 + \dots + X_{2N}^2)$$

for some absolute constant C.

11c2 Remark. Assumption A_n requires also assumption A, namely $\sum_k a_k^2 = 1$, but we do not need it here; we use only the assumption

$$\forall \lambda \in [0,\infty) \qquad \sum_{k:\lambda_k \in [\lambda, \lambda+1]} a_k^2 \le \frac{1}{n} \, .$$

Given $f \in L_2[0,1]$, we consider the random variable

$$\langle f, X \rangle = \int_0^1 f(t) X(t) \, \mathrm{d}t \, ;$$

this is a linear combination of $X_1, \dots, X_{2n}$, thus $\langle f, X \rangle \sim N(0, \mathrm{Var}\langle f, X \rangle)$.

[1] $C = \sup_{t>0} \mathrm{e}^{t^2/2} \cdot 2 \int_t^\infty (2\pi)^{-1/2} \mathrm{e}^{-u^2/2} \, \mathrm{d}u = 2 \sup_{t>0} (2\pi)^{-1/2} \int_0^\infty \exp(-\frac{s^2}{2} - ts) \, \mathrm{d}s = 1$.

[2] $C = (2\pi)^{-1/2} \int_0^\infty t \mathrm{e}^{-t^2/2} \, \mathrm{d}t = 1/\sqrt{2\pi}$.

[3] Here and henceforth, constants c and C (possibly with indices) are positive. They may be different in different formulas.

[4] In fact, $c = 1$ and $C = 2$. Moreover, $\mathbb{P}\big(\xi - \mathbb{E}\,\xi > \varepsilon \big) \le 2\mathbb{P}\big(\sigma\zeta > \varepsilon \big)$ (Cirel'son, Ibragimov, Sudakov 1976), thus, $\mathbb{P}\big(\xi - \mathbb{E}\,\xi > \varepsilon \big) \le \exp\big(-\frac{\varepsilon^2}{2\sigma^2} \big)$.

[5] But under assumption A only, the operator need not be of small norm; just try $N = 1$.

11c3 Exercise. Deduce 11c1 from the following claim (to be proved soon):

$$\mathrm{Var}\langle f, X\rangle \le \frac{C}{n}\|f\|^2 .$$

11c4 Exercise. Prove that

$$\mathrm{Var}\langle f, X\rangle = \sum_{k=1}^N a_k^2 |g(\lambda_k)|^2 ,$$

where $g(\lambda) = \int_0^1 \mathrm{e}^{\mathrm{i}\lambda t} f(t)\,\mathrm{d}t$.

It is well-known that $\|g\|_2^2 = 2\pi\|f\|_2^2$. Thus, the claim in 11c3 boils down to[1]

$$\sum a_k^2 |g(\lambda_k)|^2 \le C\|g\|^2 \sup_\lambda \sum_{k:\lambda_k\in[\lambda,\lambda+1]} a_k^2 ,$$

which may be rewritten as

$$\int |g|^2\,\mathrm{d}\mu \le C\Big(\int |g|^2\,\mathrm{d}m\Big) \sup_\lambda \mu\big([\lambda,\lambda+1]\big) \tag{11c5}$$

where $\mu = \sum_k a_k^2 \delta_{\lambda_k}$ (a discrete measure), and m is the Lebesgue measure.

The idea is, roughly, that g cannot be nearly concentrated on a short interval, because f is concentrated on an interval of length 1. The proof, given below, uses Fourier transform ($\varphi \mapsto \hat\varphi$) and convolution ($*$). If you are familiar with these, keep reading. Otherwise feel free to skip the rest of 11c.

11c6 Lemma. There exist even real-valued functions $\varphi \in L_\infty[-0.5, 0.5] \subset L_1(\mathbb{R})$ and $\psi \in L_1(\mathbb{R})$ such that $\hat\varphi(x)\hat\psi(x) = 1$ for all $t \in [-1, 1]$.

Proof. We take $\varphi(t) = \text{const}$ on $[-0.5, 0.5]$ (and 0 outside), $\hat\varphi(t) = \frac{1}{t}\sin\frac{t}{2}$, note that $\hat\varphi(\cdot)$ does not vanish on $[-1, 1]$, $1/\hat\varphi(\cdot)$ is smooth on $[-1, 1]$ and therefore can be extended to a smooth compactly supported function $\hat\psi(\cdot)$; its Fourier transform is integrable, since it decays fast enough. □

Proof of the proposition. The function $|g(\cdot)|^2$ is the Fourier transform of a function supported on $[-1, 1]$ and therefore invariant under multiplication by $\hat\varphi\hat\psi$. It means that $|g|^2 = |g|^2 * \varphi * \psi$. Thus,

$$\int |g|^2\,\mathrm{d}\mu = \langle |g|^2 * \psi, \mu * \varphi\rangle \le \||g|^2 * \psi\|_1 \|\mu * \varphi\|_\infty ;$$

$$\||g|^2 * \psi\|_1 \le \||g|^2\|_1 \|\psi\|_1 = \|g\|_2^2 \|\psi\|_1 \le C\|g\|_2^2 ;$$

$$\|\mu * \varphi\|_\infty \le \|\varphi\|_\infty \sup_\lambda \mu\big([\lambda - 0.5, \lambda + 0.5]\big) \le C \sup_\lambda \mu\big([\lambda, \lambda + 1]\big) ,$$

which gives (11c5). □

[1] Do not forget that C may be different in different formulas.

11d Proving Theorem 2a2

If $\xi : L_2[0,1] \to \mathbb{R}$ is Lip(1) then $\xi(X)$, treated as a function of $X_1, \dots, X_{2N}$, is a $\operatorname{Lip}(C/\sqrt{n})$ function $\mathbb{R}^{2N} \to \mathbb{R}$ (by 11c1). Thus, $\xi(X) \in \operatorname{GaussLip}(C/\sqrt{n})$. By (11b1),[1]

$$\mathbb{P}\big(|\xi - \mathbb{E}\,\xi| > \varepsilon\big) \le C \exp(-c\varepsilon^2 n) \tag{11d1}$$

for some absolute constants c, C. In this sense, abusing the language, we write (under assumption A_n)

$$\xi \in \operatorname{ExpConInt}(n)$$

whenever ξ is Lip(1) on $L_2[0,1]$, or Lip(C) for some C not depending on n. Usually, a stronger condition will be satisfied: ξ is Lip(C) on $L_1[0,1]$.

11d2 Exercise. Prove Lemma 2a1.

11d3 Exercise. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be Lip(1). Then the function $\xi : L_1[0,1] \to \mathbb{R}$,

$$\xi(x) = \int_0^1 \varphi(x(t))\,\mathrm{d}t\,,$$

is well-defined and Lip(1).

Prove it.

Thus, for such φ the random variable

$$\xi = \int_0^1 \varphi(X(t))\,\mathrm{d}t$$

satisfies

$$\xi \in \operatorname{GaussLip}(C/\sqrt{n})\,; \quad \xi \in \operatorname{ExpConInt}(n)$$

with absolute constants (as in (11d1)).

Now let φ be as in Theorem 2a2 (continuous a.e., of linear growth). We introduce for every k

$$\varphi_k^-(x) = \inf_y \big(\varphi(y) + k|y-x|\big)\,, \quad \varphi_k^+(x) = \sup_y \big(\varphi(y) - k|y-x|\big)\,.$$

11d4 Exercise. (a) φ_k^-, φ_k^+ are Lip(k) functions $\mathbb{R} \to \mathbb{R}$ for all k large enough;[2]

(b) $\varphi_k^- \uparrow \varphi$ and $\varphi_k^+ \downarrow \varphi$ almost everywhere;

[1] I often write just ξ instead of $\xi(X)$.

[2] Do you understand why not just "for all k"?

(c) there exists C_φ such that for all k large enough and all x

$$-C_\varphi(1+|x|) \le \varphi_k^-(x) \le \varphi_k^+(x) \le C_\varphi(1+|x|)\,.$$

Prove it.

It follows (using Fubini and the dominated convergence theorem) that $\mathbb{E}\,\xi_k^- \uparrow \mathbb{E}\,\xi$ and $\mathbb{E}\,\xi_k^+ \downarrow \mathbb{E}\,\xi$ a.s., where $\xi_k^\pm = \int_0^1 \varphi_k^\pm(X(t))\,\mathrm{d}t$. We have a "sandwich"; and so, Theorem 2a2 follows by 11a7. (The upper bound $2\mathrm{e}^{-c_{\varepsilon,\varphi}n}$ is not stronger than $C_{\varepsilon,\varphi}\mathrm{e}^{-c_{\varepsilon,\varphi}n}$ since $c_{\varepsilon,\varphi}$ can be made smaller.)

11e Proving Theorem 2a3

The function T was defined in Sect. 2a on $C[0,1]$, but the same definition works on $L_1[0,1]$ and evidently gives a Lip(1) function $T : L_1[0,1] \to [0,\infty)$. It follows that $T(X) \in \mathrm{ExpConInt}(n)$. However, Theorem 2a3 states that $T(X) \in \mathrm{ExpConZero}(n)$. Thus, it is sufficient to prove that $\mathbb{E}\,T(X) \le \varepsilon_n \to 0$.

We modify T as follows:

$$T_k(f) = \inf_g \|\psi_k(f(\cdot)) - \psi_k(g(\cdot))\|_1\,,$$

where g is as before (distributed γ^1), and $\psi_k(x) = \mathrm{mid}(-k,x,k)$, that is, $-k$ for $x \in (-\infty,-k]$; x for $x \in [-k,k]$; and k for $x \in [k,\infty)$. We have

$$\mathbb{E}\,|T_k(X(\cdot)) - T(X(\cdot))| \le \mathbb{E}\,\|\psi_k(X(\cdot)) - X(\cdot)\|_1 + \|\psi_k(g(\cdot)) - g(\cdot)\|_1 = \\ = 2\int |\psi_k(x) - x|\,\gamma^1(\mathrm{d}x) \to 0$$

as $k \to \infty$. It remains to prove that $\mathbb{E}\,T_k(X) \le \varepsilon_{k,n} \to 0$ as $n \to \infty$.

11e1 Exercise. For every $f \in L_1[0,1]$ and every Lip(1) function $\varphi : \mathbb{R} \to \mathbb{R}$,

$$\Big| \int_0^1 \varphi(f(t))\,\mathrm{d}t - \int \varphi\,\mathrm{d}\gamma^1 \Big| \le T(f)\,.$$

Prove it.

It is well-known that[1]

$$\sup_\varphi \Big| \int_0^1 \varphi(f(t))\,\mathrm{d}t - \int \varphi\,\mathrm{d}\gamma^1 \Big| = T(f)\,,$$

where the supremum is taken over all Lip(1) functions $\mathbb{R} \to \mathbb{R}$. (I use this fact without proof.) Clearly we may demand $\varphi(0) = 0$.

[1] Kantorovich-Rubinstein theorem. This $T(f)$ is nothing but the transportation distance between γ^1 and the distribution of f. This fact is evident when f is a step function. It extends to the whole $L_1[0,1]$ by continuity.

11e2 Exercise. For every k and ε there exists a finite set of Lip(1) functions $\varphi_1, \dots, \varphi_N : [-k,k] \to \mathbb{R}$ such that $\varphi_1(0) = 0, \dots, \varphi_N(0) = 0$, and every Lip(1) function $\varphi : [-k,k] \to \mathbb{R}$ such that $\varphi(0) = 0$ is ε-close to some φ_i uniformly on $[-k,k]$.

Prove it.

11e3 Exercise. Prove that

$$T_k(f) \le 2\varepsilon + \max_{i=1,\dots,N} \biggl| \int_0^1 \varphi_i(\psi_k(f(t)))\,\mathrm{d}t - \int \varphi_i(\psi_k(\cdot))\,\mathrm{d}\gamma^1 \biggr| .$$

The function $\varphi_i(\psi_k(\cdot))$ is Lip(1), thus the random variable $\xi_{i,k} = \int_0^1 \varphi_i(\psi_k(X(t)))\,\mathrm{d}t$ belongs to GaussLip$(C/\sqrt{n})$. By (11b2), $\mathbb{E}\,|\xi_{i,k} - \mathbb{E}\,\xi_{i,k}| \le C/\sqrt{n}$. Thus,

$$\mathbb{E}\,T_k(X) \le 2\varepsilon + \mathbb{E} \max_{i=1,\dots,N} |\xi_{i,k} - \mathbb{E}\,\xi_{i,k}| \le 2\varepsilon + N_{k,\varepsilon} \cdot \frac{C}{\sqrt{n}} \,,$$

which can be made small enough by choosing ε first and n afterwards. That is, $\mathbb{E}\,T_k(X) \le \varepsilon_{k,n} \to 0$ as $n \to \infty$, which completes the proof.[1]

11f Dimension two, and higher

Returning to the definition of $X(\cdot)$ given in Sect. 2a via $a_1, \dots, a_N$ and $\lambda_1, \dots, \lambda_N$, we replace the numbers $a_1, \dots, a_N > 0$ with vectors $a_1, \dots, a_N \in \mathbb{R}^2$, thus getting $X : \mathbb{R} \to \mathbb{R}^2$; we endow $\mathbb{R}^2$ with the Euclidean norm $x \mapsto |x|$. Further, all occurrences of a_k^2 (in assumptions A and A_n, and everywhere) turn into $|a_k|^2$, and all occurrences of $X^2(t)$ (in Prop. 11c1, and everywhere) into $|X(t)|^2$. We also replace the requirement $0 < \lambda_1 < \dots < \lambda_N < \infty$ with a weaker requirement $0 < \lambda_1 \le \dots \le \lambda_N < \infty$, thus allowing a single frequency to cover more than one dimension.[2] The distribution of the process X fails to determine uniquely the vectors a_k, but still determines the measure $\sum_k |a_k|^2 \delta_{\lambda_k}$, since

$$\mathbb{E}\,\langle X(0), X(t) \rangle = \sum_{k=1}^N |a_k|^2 \cos \lambda_k t \,.$$

[1] In fact, $\mathbb{P}\bigl(T(X) \ge \varepsilon\bigr) \le \exp\bigl(-c((\varepsilon - \alpha_n)^+)^2 n\bigr)$ for some absolute constant c and some $\alpha_n \to 0$ (depending on n only). It is like the large deviations principle with the rate function $I(\varepsilon) \ge c\varepsilon^2$.

[2] Think, what does it change in the one-dimensional case.

Still, 11c3 and 11c4 hold, but $f \in L_2[0,1]$ turns into $f \in L_2([0,1] \to \mathbb{R}^2)$, and 11c4 becomes

$$\operatorname{Var}\langle f, X\rangle = \sum_{k=1}^{N} |\langle a_k, g(\lambda_k)\rangle|^2 \le \sum_{k=1}^{N} |a_k|^2 |g(\lambda_k)|^2 \,.$$

Nothing changes in the rest of Sect. 11c (it is about the measure $\mu = \sum_k |a_k|^2 \delta_{\lambda_k}$).[1]

Thus, 11c1 gives us a linear operator $\mathbb{R}^{2N} \to L_2([0,1] \to \mathbb{R}^2)$ of norm $\le C/\sqrt{n}$. If $\xi : L_2([0,1] \to \mathbb{R}^2) \to \mathbb{R}$ is $\operatorname{Lip}(1)$ then $\xi(X) \in \operatorname{GaussLip}(C/\sqrt{n})$.

The function $\varphi : \mathbb{R} \to \mathbb{R}$ in 2a1, 2a2, 11d3, 11d4 (as well as $\varphi_k^{\pm}$ in 11d4) turns into $\varphi : \mathbb{R}^2 \to \mathbb{R}$; γ^1 in 2a1 turns into γ^2. And of course, $L_1[0,1]$ in 11d3 turns into $L_1([0,1] \to \mathbb{R}^2)$.

Theorem 2a2 is thus generalized.

About Theorem 2a3. The definition of $T(f)$ is generalized evidently (γ^1 turns into γ^2); now T is a $\operatorname{Lip}(1)$ function $L_1([0,1] \to \mathbb{R}^2) \to [0,\infty)$. The functions $\psi_k : \mathbb{R}^2 \to \mathbb{R}^2$ may be defined by $\psi_k(x) = x$ if $|x| \le k$, otherwise $\psi_k(x) = kx/|x|$. The Kantorovich-Rubinstein theorem holds for all metric spaces, in particular $\mathbb{R}^2$. Exercise 11e2 generalizes for a disk of $\mathbb{R}^2$ (and in fact for every precompact metric space). Exercise 11e3 and the rest of the proof remain valid.[2]

Theorem 2a3 is thus generalized.

All said about $\mathbb{R}^2$ holds equally well for $\mathbb{R}^d$, $d = 3, 4, \dots$

11g Hints to exercises

11d2: Fubini.

Index

[1] And so, the absolute constant C in 11c1 remains intact.

[2] Still, $\mathbb{P}\big(T(X) \ge \varepsilon\big) \le \exp\big(-c((\varepsilon - \alpha_n)^+)^2 n\big)$ for the same absolute constant c as in dimension one, and another (worse) sequence $\alpha_n \to 0$.