Corollary 6.4 Let X_t be a centered Gaussian process on compact $T \subset \mathbb{R}^n$, with a continuous covariance function. Denote

$$\varrho(u) := \sup_{|s-t| \le u} (\mathbb{E}|X_t - X_s|^2)^{1/2}, \quad u > 0.$$

If for some $\delta > 0$

$$\int_0^\delta (\log \tfrac{1}{u})^{1/2} d\varrho(u) < \infty,$$

the process has a version with continuous paths and the modulus of continuity w.r.t. the standard metric satisfies for small enough $\delta < \delta_0$:

$$\omega_X(\delta) \le K \int_0^\delta (\log \tfrac{1}{u})^{1/2} d\varrho(u)$$

Proof. By compactness and continuity of C, ϱ is continuous, $\varrho(0) = 0$. Assume first that ϱ is strictly increasing so that the inverse is well-defined. By scaling we may assume that $T \in [0,1]^n$.

We may cover $[0,1]^n \supset T$ by $\left(1 + \frac{\sqrt{n}}{2\varrho^{-1}(\varepsilon)}\right)^n$ Euclidean balls of d_X-radius at most ε. Hence

$$H(\varepsilon) \le \log\left(1 + \frac{\sqrt{n}}{2\varrho^{-1}(\varepsilon)}\right)^n$$

$$\le c_n + n\log\left(\frac{1}{\varrho^{-1}(\varepsilon)}\right) \qquad \text{(for small } \varepsilon)$$

$$\le c_n' \log\left(\frac{1}{\varrho^{-1}(\varepsilon)}\right)$$

Thus
$$\int_0^\delta \sqrt{H(\varepsilon)}\, d\varepsilon \lesssim \int_0^\delta \sqrt{\log \frac{1}{\varrho^{-1}(\varepsilon)}}\, d\varepsilon$$

$$= \int_0^{\varrho^{-1}(\delta)} \sqrt{\log \frac{1}{\varepsilon}}\, d\varrho(\varepsilon)$$

If the Euclidean distance is less than δ, then d_X is less than $g(\delta)$ and we obtain

$$\omega_X(\delta) \leq K \int_0^\delta \sqrt{\log(\tfrac{1}{\varepsilon})}\, dg(\varepsilon),$$

as was to be shown. In the general case let X' be e.g. the Levy-Brownian motion on T, so that $\mathbb{E}|X'_t - X'_s|^2 = |t-s|$. Observe that $d'_{X+\varepsilon X'}(s,t) = (d^2_{X}(s,t) + \varepsilon^2 d^2_{X'}(s,t))^{1/2}$, assuming $X' \perp X$. Apply the above on $X+\varepsilon X$ and let $\varepsilon \searrow 0$. $\square$

Corollary 6.5. If $(X_t)_{t\in T}$ is a centered Gaussian process on $T \subset \mathbb{R}^n$ (T compact) with

$$\mathbb{E}|X_t - X_s|^2 \leq \frac{c}{(\log \frac{1}{|s-t|})^{1+\alpha}}, \quad \alpha > 0$$

for $|s-t| < 1/2$, then X_t has a version with continuous paths.

Before proving this, observe:

Exercise $I := \int_0^\delta (\log \frac{1}{u})^{1/2} dg(u) < \infty$

$$\Leftrightarrow \lim_{u\to 0^+} \sqrt{\log \tfrac{1}{u}}\, g(u) = 0 \quad \& \quad \int_{\sqrt{\log 1/\delta}}^\infty g(e^{-u^2})\, du < \infty.$$

If finite, then

$$I = g(\delta)\sqrt{\log 1/\delta} + \int_{\sqrt{\log 1/\delta}}^\infty g(e^{-u^2})\, du.$$

Proof of Cor. 6.5 Direct computation using the exercise. Actually one obtains the modulus of continuity (use $g(\varepsilon) \leq (\log \frac{1}{\varepsilon})^{-\frac{1}{2}-\alpha/2}$)

$$\omega_X(\delta) \leq (\log 1/\delta)^{-\alpha/2}$$

$\square$

<u>Corollary 6.6:</u> If $(X_t)_{t\in T}$ is a centered Gaussian process on $T\subset \mathbb{R}^n$ (T compact), with

$$\mathbb{E}|X_t-X_s|^2 \le c\,|t-s|^{\alpha} \qquad (\alpha<2),$$

then $$\omega_X(\delta) \lesssim \sqrt{\log\tfrac{1}{\delta}}\;\delta^{\alpha/2}$$ for small enough $\delta<\eta$ (random).

Proof. Choose $g(u)=u^{\alpha/2}$, obtain by the exercise on p. 79

$$\omega_X(\delta) \le \delta^{\alpha/2}\sqrt{\log 1/\delta} + \int_{\sqrt{\log 1/\delta}}^{\infty} e^{-\frac{\alpha u^2}{2}}$$

$$\sim \delta^{\alpha/2}\sqrt{\log 1/\delta} + \frac{1}{\sqrt{\alpha}}\int_{\sqrt{(\log 1/\delta)\alpha}} e^{-u^2/2} \sim \delta^{\alpha/2}\sqrt{\log\tfrac{1}{\delta}}$$

$$+\sqrt{\log\tfrac{1}{\delta}}\;\delta^{\alpha/2} + \frac{1}{\sqrt{\alpha}}\frac{1}{\sqrt{\log 1/\delta}}\left(\frac{1}{\delta}\right)^{-\alpha/2} \lesssim \sqrt{\log\tfrac{1}{\delta}}\;\delta^{\alpha/2}$$

$\square$

<u>Remark.</u> The above corollary is very sharp: For the Brownian motion (see Example 1 below) $\alpha=\frac{1}{2}$ and it is known that for the Brownian motion on zero-one one has (Levy)

$$\lim_{\delta\to 0}\frac{\omega_B(\delta)}{\sqrt{2\delta\log 1/\delta}} = 1 \quad \text{a.s.} \;?$$

<u>Example 1.</u> In $\mathbb{R}^n$ there is a Gaussian process*) with

$$\mathbb{E}|X_t-X_s|^2 = |t-s|^{2H} \qquad (H\in(0,1)).$$

This is called the Levy FBM (with Hurst parameter H). By Corollary 8.6

$$\omega_{B^H}(\delta) \lesssim \sqrt{\log 1/\delta}\;\delta^{H}.$$

---

*) A comment to Daniel: hence the <u>snowflaked</u> $\mathbb{R}^n$ embeds isometrically to e.g. $L^2(0,1)$!

Example 2 (For those who already know white noise, we will study it soon ~~later~~). Recall that BM is integrated white noise. How about in n-dimension what replaces BM as the integral? Answer: $(W_t)_{t\in\mathbb{R}_+^n}$, where

$$W_t = \int_{[0,t]} dW \quad \leftarrow \text{white noise}$$

$$[0,t] \leftarrow = [0,t_1]\times\dots\times[0,t_n]$$

One obtains

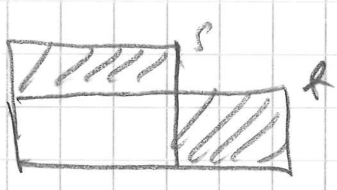
$$\mathbb{E}\, W_t W_s = |[0,t]\cap[0,s]|$$

$$= (s_1\wedge t_1)(s_2\wedge t_2)\dots(s_n\wedge t_n)$$

We will later on verify that in a rigorous sense

$$dW = \frac{d^n}{dt_1\dots dt_n} W_t .$$

Simple geometric reasoning shows that for $s,t\in[0,1]^n$



$$E|W_t - W_s|^2 \le C_n |t-s| ,$$

which yield the same mod. of continuity as for standard BM.

Example 3. If $\mathbb{E}|X_t - X_s|^2 \ge \frac{c}{(\log\frac{1}{|s-t|})^{1/2}}$ (in $\mathbb{R}^1$), and the process is stationary, the process does not have cont. paths.
This follows from the next important result.

Theorem 6.7 (Fernique) Let $(X_t)_{t\in T}$ be a stationary process on a compact Abelian group. Then Dudley's condition (Thm 8.1) is necessary for the existence of continuous paths (or for boundedness)

Proof. See e.g. Kahane's book. □

Rem.
- Thus, in the stationary case continuity and boundedness occur simultaneously.
- In $\mathbb{R}^n$ (one may often use $\mathbb{T}^n = [0,2\pi)^n$ to reduce to 8.7) stationarity means that $\mathbb{E}|X_t - X_s|^2 = F(t-s)$.
- Talagrand gave a complete (albeit more technical) characterization in the nonstationary case by applying majorizing measure's (hence proving a conjecture of Dudley!).

Exercise Assume that $(X_t)_{t\in\mathbb{T}}$ ($\mathbb{T}\subset\mathbb{R}^n$ compact) Gaussian, centered has continuous modification. Denote the covariance $C(t,s) = \mathbb{E}\, X_t X_s$ (now C is continuous). Assume also that $\mathbb{E}|X_t - X_s|^2 > 0$ if $t\neq s$. Consider the linear space

$$S := \Big\{ u: T\to\mathbb{R} \;\Big|\; u(\cdot) = \sum_{j=1}^{m} a_j\, C(s_j,\cdot),\; a_j\in\mathbb{R}\ \forall j\le m,\ m\ge 1 \Big\}.$$

Define the inner product

$$\langle u, v\rangle_H := \sum_{j=1}^{m}\sum_{k=1}^{\ell} a_j b_k\, C(s_j, t_k)$$

if $u(\cdot) = \sum_{j=1}^{m} a_j u(s_j, \cdot)$, $v(\cdot) = \sum_{k=1}^{l} b_j u(k_j, \cdot)$.

Prove:

- $\langle u, v \rangle_H$ is an inner product with the reproducing property:

  $\langle u, C(t, \cdot) \rangle_H = u(t)$ (*)

- The completion of S in this inner product is a Hilbert space that embeds continuously in $C(T)$. Call this space H.

- Show that H is the Cameron-Martin space of X (thought as $C(T)$-valued Gaussian random variable.

  (because of (*) H is often called the RKHS = reproducing kernel Hilbert space).

~~We finally, without proof~~

We finish this section by presenting couple of further important basic results on Gaussian fields. Always below $(X_t)_{t \in T}$ is Gaussian, centered and has continuous version. !

Theorem 6.8. (Borel-Tsirelson-Ibragimov-Sudakov inequality). Let $\sigma_{max} = \max_{t \in T} (\mathbb{E} X_t^2)^{1/2}$. Then

$$\mathbb{P}\left(\sup_{t \in T} X_t - \mathbb{E}(\sup_{t \in T} X_t) > u \right) \le e^{-u^2/2\sigma_{max}^2}$$

We will base the proof of Thm 6.8 on the following tail estimate for functionals of standard d-dimensional Gaussians:

Lemma 6.9 Let $X \sim N(0, Id_d)$ and assume that $F: \mathbb{R}^d \to \mathbb{R}$ is Lipschitz, so that

$$\mathrm{Lip}(F) := \sup_{\substack{x,y \in \mathbb{R}^d \\ x \neq y}} \frac{|F(x)-F(y)|}{|x-y|} < \infty .$$

Then

$$\mathbb{P}\left(|F(X) - \mathbb{E}F(X)| > \lambda \right) \leq 2 e^{-x^2/2(\mathrm{Lip}(F))^2}$$

Proof. We may assume that $\mathrm{Lip}(F) = 1$ by scaling and $F \in C_0^\infty(\mathbb{R}^d)$ by simple approximation.
Recall the heat kernel in $\mathbb{R}^d$:

$$H(x,y,t) := \frac{1}{(2\pi t)}^{d/2} \exp(-|x-y|^2/2t), \quad \begin{matrix} x \in \mathbb{R}^d \\ t > 0 \end{matrix}$$

It solves the heat equation*) and is a good approximation of identity: if

$$g(x,t) = \int_{\mathbb{R}^d} H(x,y,t) F(y) dy$$
$$= \frac{1}{(2\pi t)}^{d/2} \int_{\mathbb{R}^d} e^{-|x-y|^2/2t} F(y) dy ,$$

then

(6) $$\frac{d}{dt} g = \frac{1}{2} \Delta_x g$$

and (7) $$\lim_{t \to 0+} g(x,t) = F(x) \quad \text{(uniformly for } x \in \mathbb{R}^d\text{)}.$$

Our aim is to apply Itô's Formula on the function**)

$$g(B_t, 1-t) , \qquad t \in [0,1] ,$$

---

*) Note the normalization with factor $\frac{1}{2}$!

**) By (7) we set $g(x,0) = F(x)$

where $t \mapsto B(t)$ is a standard d-dimensional Brownian motion. For that end, note that

$$(8) \qquad g(B_0,1) = g(0,1) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-|y|^2/2} F(y)\,dy$$
$$= \mathbb{E}\, F(X),$$

since X has the density $(2\pi)^{-d/2}$ in $\mathbb{R}^d$. Secondly,

$$(9) \qquad g(B_1, 1-1) = g(B_1, 0) = F(B_1) \sim F(X)$$

since $X \sim B_1$. By Ito's formula

$$g(B_1,0) - g(B_0,1)$$
$$= \int_0^t \nabla_x g(B_s, 1-s) \cdot dB_s$$
$$+ \int_0^t \left(-\frac{d}{dt} g(B_s,1-s) + \frac{1}{2}\Delta_x g(B_s,1-s) \right) ds$$
$$= \int_0^t \nabla_x g(B_s,1-s) \cdot dB_s$$

When this is combined with (8) and (9) we get

$$(10) \qquad F(X) - \mathbb{E}\, F(X) \sim \int_0^1 \nabla_x g(B_s, 1-s) \cdot dB_s$$

Since convolution does not increase the norm of the derivative, we have

$$(11) \qquad |\nabla_x g| \le 1 \qquad (|\cdot| \text{ is the Euclidean norm in } \mathbb{R}^d)$$

We then invoke the standard identity for exponentials of Gaussian stochastic integrals:

$$(12) \quad \mathbb{E} \exp\left(\int_0^t h(B_s,s)\cdot dB_s - \frac{1}{2}\int_0^t |h(B_s,s)|^2 ds\right) = 1$$

This applied to (10) and recalling (11) yields for any $y \in \mathbb{R}$

$$1 = \mathbb{E}\exp\left(y\int_0^1 \nabla_x g(B_s,1-s)\cdot dB_s - \frac{y^2}{2}\int_0^1 |\nabla_x g(B_s,1-s)|^2 ds\right)$$

$$\geq \mathbb{E}\exp\left(y\int_0^1 \nabla_x g(B_s,1-s)\cdot dB_s - \frac{y^2}{2}\right)$$

$$= \mathbb{E}\exp\left(y(F(X) - \mathbb{E}F(X)) - \frac{y^2}{2}\right),$$

or in other words

$$\mathbb{E}\exp\left(y(F(X) - \mathbb{E}F(X))\right) \leq e^{y^2/2}$$

By Chebyshev, if $y > 0$, this yields that

$$\mathbb{P}(F(X) - \mathbb{E}f(X) > y) \leq e^{y^2/2 - y^2} = e^{-y^2/2}$$

and taking $y < 0$ gives the same estimate for the negative tail. $\square$

<u>Proof of Theorem 6.8.</u> Assume first that our compact space T is finite, $T = \{t_1, \dots, t_d\}$. Write

$$X \sim C^{1/2} Y_0,$$

where $C_{ij} = \mathbb{E} X_{t_i} X_{t_j}, \quad 1 \leq i,j \leq d.$

Define function $F: \mathbb{R}^d \to \mathbb{R}$ by setting

$$F(x) := \max_{1\leq i\leq d} (C^{1/2}x)_i$$

Denote by e_i the i:th unit vector in $\mathbb{R}^d$. Then for any $x,y\in\mathbb{R}^d$ we obtain that

$$|F(x)-F(y)| \le \max_{1\le i\le d} |(C^{1/2}(x-y))_i| = \max_{1\le i\le d} |\langle e_i, C^{1/2}(x-y)\rangle|$$
$$= \max_{1\le i\le d} |\langle C^{1/2}e_i, x-y\rangle| \le \big(\max_{1\le i\le d} |C^{1/2}e_i|\big)|x-y|$$

Above $|C^{1/2}e_i| = \langle C^{1/2}e_i, C^{1/2}e_i\rangle^{1/2} = \langle Ce_i, e_i\rangle^{1/2} = (C_{ii})^{1/2}$

Thus f is Lipschitz with Lipschitz constant
$$\mathrm{Lip}(F) \le \max C_{ii}^{1/2} = \sup_{t\in T}(\mathbb{E}|X_t|^2)^{1/2},$$
and the claim then follows from Thm. 6.9.

The general case is reduced by standard approximation to the above using a countable dense subset of T and monotonicity (Exercise). Here it is important to note that
$$\mathbb{E}\sup_{t\in T}|X_t|$$
is finite, which follows from Fernique and our assumption on the existence of continuous realizations. (or see the remark on p. 76). $\square$

Remark • Thm 6.1 can be used to estimate

$$\mathbb{E}\sup_{t\in T} X_t \le \mathbb{E}\sup_{t\in T}|X_t - X_{t_0}| + \mathbb{E}|X_{t_0}|$$
$$\le \mathbb{E}|X_{t_0}| + c\int_0^{\mathrm{diam}(T)} \sqrt{H(\varepsilon)}\,d\varepsilon$$

When this is combined with Thm 6.8 we obtain an estimate for the tail of $\sup X_t$ that is quantitative and useful for many purposes.

- The remarkable part in Borell-TIS theorem is that the tail of the deviation of the maximum from the expectation decays as ($\approx$) well as the tail of the single X_{t_0} with the largest variance!

We next move towards Slepian's and other comparision inequalities.

Theorem 6.10 (Slepian's inequality) Let X_t, Y_t, $t \in T$ (again compact, metric ...) be centered Gaussian fields on T with continuous versions. Assume that

$$\mathbb{E}X_t^2 = \mathbb{E}Y_t^2 \quad \forall\, t \in T \text{ and}$$

$$\mathbb{E}(Y_t - Y_s)^2 \le \mathbb{E}(X_t - X_s)^2 \quad \forall (s,t) \in T^2.$$

Then $$\mathbb{P}(\sup_{t\in T} X_t > \lambda) \ge \mathbb{P}(\sup_{t \in T} Y_t > \lambda) \quad \forall\, \lambda > \mathbb{R}$$

Proof. We shall first show that if $h \in C^2(\mathbb{R}^d)$ and h together with its first and second derivatives grows at most exponentially, then for any symmetric $d \times d$ matrix $C = (c_{k,l})_{k,l=1}^d > 0$ it holds that if for the all (k,l), $1 \le k < l \le d$,

(12) $$\frac{\partial^2 h(x)}{\partial x_k \partial x_l} \ge 0 \quad \forall\, x \in \mathbb{R}^d,$$ then

(13) $$\frac{\partial}{\partial c_{k,l}} H(C) \ge 0,$$ where

$$H(C) = \mathbb{E}\, h(Z), \quad \text{with } Z \sim N(0, C).$$

In order to prove (12) we denote by
$$g(x) = (2\pi)^{-d/2} |C|^{-1/2} \exp\left(-\tfrac{1}{2}(C^{-1}x, x)\right)$$
the density of X. Observe that the characteristic function satisfies
$$\hat{g}(\xi) = \exp\left(-\tfrac{1}{2}\langle C\xi, \xi\rangle\right),$$
$$\frac{\partial}{\partial c_{k,l}} \hat{g}(\xi) = -\xi_k \xi_l \hat{g}(\xi)$$
and taking the inverse Fourier transform yields
$$\frac{\partial}{\partial c_{k,l}} g(x) = \frac{\partial^2}{\partial x_k \partial x_l} g(x).$$
Hence we may compute
$$\frac{\partial H(C)}{\partial c_{k,l}} = \frac{\partial}{\partial c_{k,l}} \int_{\mathbb{R}^d} h(x) g(x)$$
$$= \int_{\mathbb{R}^d} h(x) \frac{\partial^2}{\partial x_k \partial x_l} g(x)$$
$$= \int_{\mathbb{R}^d} \frac{\partial^2 h(x)}{\partial x_k \partial x_l} g(x) dx \geq 0,$$
which is (12).

Let us then note that by standard approximation an dominated convergence we may assume that $T = \{t_1, .., t_d\}$ is finite. Define the matrices

$$C_{k,l} := \mathbb{E}\, Y_{t_k} Y_{t_l}, \quad \tilde{C}_{k,l} = \mathbb{E}\, X_{t_k t_l}, \quad 1 \leq k, l \leq d.$$

Thus $\tilde{C}_{k,k} = C_{k,k}$ and $\tilde{C}_{k,l} \leq C_{k,l}$ $\forall\, 1 \leq k, l \leq d$.

We may assume that $\{X_t\} \perp \{Y_t\}$ and define

$$R(t) := \mathbb{E}\, h(\sqrt{1-t}\, Y + t^{1/2} X) \quad , t \in [0,1]$$

$$= H(t C + (1-t)\tilde{C}) ,$$

where h is as before. Assume first that $C > 0$ and $\tilde{C} > 0$, then we may compute by (13):

$$R'(t) = \sum_{1 \le k \le l \le d} \frac{\partial}{\partial c_{k,l}} H \,(C_{k,l} - \tilde{C}_{k,l})$$

$$= \sum_{1 \le k < l \le d} \text{---} \; \text{''} \; \text{---}$$

$$\ge 0 .$$

Especially, $H(1) \ge H(0)$. We choose

$$h(x) = \prod_{k=1}^{d} a(x_k) ,$$

where $a \ge 0$ is bounded, $a \in C^2(\mathbb{R})$ and $a' \le 0$. Then h satisfies (12) and we obtain that

$$\mathbb{E} \prod_{k=1}^{d} a(X_{t_k}) \le \mathbb{E} \prod_{k=1}^{d} a(Y_{t_k})$$

Letting a approximate suitably $\chi_{(-\infty,\lambda)}$ from below it follows that

$$\mathbb{E}\left(\max_k X_{t_k} < \lambda\right) \le \mathbb{E}\left(\max_k Y_{t_k} < \lambda\right) ,$$

which clearly yields the claim. Finally, if $C \not> 0$ or $\tilde{C} \not> 0$ we may apply the above to $X + \varepsilon Z$ and $Y + \varepsilon Z$, where $Z \sim N(0, \mathrm{Id}_d)$ is independent and let $\varepsilon \to 0$. $\square$

<u>Corollary 6.11</u> In the situation of Thm 6.10

$$\mathbb{E}(\sup_{t\in T} X_t) \geq \mathbb{E}(\sup_{t\in T} Y_t).$$

<u>Proof</u> Exercise. $\square$

<u>Exercise:</u> Show by a 2-dim example that if $\sup X_t$ is replaced by $\sup |X_t|$ (same for Y_t), then Thm 6.10 is not true!

We next state a generalization of Slepian's inequality without proof:

<u>Theorem 6.12</u> Let X_t, Y_t be centered and a.s. continuous fields on (compact, metric) T.
If

$$\mathbb{E}(X_t - X_s)^2 \geq \mathbb{E}(Y_t - Y_s)^2 \ \forall\, t,s \in T,$$

then

$$\mathbb{E} \sup_{t\in T} X_t \geq \mathbb{E} \sup_{t\in T} Y_t .$$

$\square$

We next very shortly study the <u>tail of the supremum</u>*. Always below T is a compact metric space, $(X_t)_{t\in T}$ is a centered Gaussian field on T so that X_t has an a.s. continuous version. A basic simple result is given by

<u>Theorem 6.13</u> Assume that the entropy function satisfies $H(\varepsilon) \leq A + \alpha \log(\frac{1}{\varepsilon})$. Then for any $\delta > 0$

$$\mathbb{P}(\sup_{t\in T} X_t \geq \lambda) \leq C \lambda^{\alpha+\delta} e^{-\lambda^2/2\sigma_{max}^2},$$

where $C = C(\delta, \alpha, A, \sigma_{max})$

---

* equivalently: <u>excursion probabilities</u>

Proof. Set for any $\varepsilon > 0$ and $t \in T$

$$g(t,\varepsilon) := \mathbb{E} \sup_{d_X(u,t)\le\varepsilon} X_u$$

and denote $\quad g(\varepsilon) := \sup_{t\in T} \mu(t,\varepsilon)$.

By Thm 6.1 we have

$$g(\varepsilon) \le C_1 \int_0^\varepsilon (H(\varepsilon))^{1/2} d\varepsilon \le \underset{\text{depends on } A}{C_2} \int_0^\varepsilon (\log \tfrac{1}{\varepsilon})^{\frac{1}{2}}$$
$$\le C_3 \, \varepsilon (\log \tfrac{1}{\varepsilon})^{1/2}$$

Assume that $\lambda > 0$ (large enough) is given. By the definition of $N(\lambda^{-1})$ and Borell-TIS thm we obtain by applying the latter in λ^{-1}-balls:

$$\mathbb{P}\left(\sup_{t\in T} X_t \ge \lambda\right) \le N(\lambda^{-1}) \, e^{-\frac{1}{2}(\lambda - g(\frac{1}{\lambda}))^2/\sigma_{max}^2}$$
$$\lesssim \lambda^{\alpha} \exp\left(-\tfrac{1}{2}\left(\lambda - C_3 \lambda^{-1}\sqrt{\log \lambda^{-1}}\right)^2/\sigma_{max}^2\right)$$
$$\lesssim \lambda^{\alpha} \exp\left(-\tfrac{1}{2}\lambda^2/\sigma_{max}^2\right) \exp\left(C_3\sqrt{\log \lambda^{-1}}\right)$$

Here $\exp(C_3\sqrt{\log \lambda}) \lesssim \lambda^{\delta}$ for any δ $\square$

The previous Thm can be improved in many ways. Especially, Talagrand or Samorodnitsky were able to take $\delta = 0$ and gave beautiful explicite expression for the constant C:

$$C = \left(\frac{K e^{A}}{\sqrt{\alpha}\,\sigma_{max}^2}\right)^{\alpha}.$$

- Rem. Note that the theorem above applies e.g. if $g(\varepsilon)$ is polynomial.

Mercer thm and Karhunen-Loeve expansion

We finish this section by recalling the classical Karhunen-Loeve expansion. First we state and prove

Thm 6.14. (Mercer). Assume that $K \in C^2(T)$ is a continuous symmetric kernel on the metric measure space (T, d, μ), where μ is a finite Borel measure on T such that every open ball has positive measure. Define the operator A on $L^2(T, d\mu)$ by setting

$$Af(x) := \int_T K(x,y) f(y)$$

Assume that $A \geq 0$, i.e

$$(14) \qquad \int_{T\times T} f(x) K(x,y) f(y)\, \mu(dy)\mu(dx) \geq 0$$

for all $f \in L^2(T)$ (equivalently, for all $f \in C(T)$). Then A is a compact self-adjoint operator on $L^2(d\mu)$ that has a sequence of eigenfunctions $(\varphi_n)_{n\geq 1}$ (or their number may be finite) corresponding to positive eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \xrightarrow[n\to\infty]{} 0$ $(n=1,\ldots)$. Thus

$$A\varphi_n = \lambda_n \varphi_n .$$

Moreover $\varphi_n \in C(T)$ for all n and we may write

$$K(x,y) = \sum_{n=1}^{\infty} \lambda_n \varphi_n(x)\varphi_n(y)$$

with uniform convergence on $T\times T$.

Finally, T is of trace class and
$$\sum_{k=1}^{\infty} \lambda_k = \int_{T} K(t,t)\,d\mu(t)$$

Proof. By the course 'Functional Analysis', operator A is compact, and since $K(x,y) = K(y,x)$, A is self-adjoint, and hence there is a basis on $L^2(d\mu)$ consisting of eigenfunctions of A. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \underset{n\to\infty}{\to} 0$ be the strictly positive eigenvalues, and φ_n's the corresponding eigenfunctions. The uniform continuity of K, Cauchy-Schwarz and the equality
$$\lambda_n \varphi_n(x) = \int_T K(x,y)\varphi_n(y)\,dy \qquad \text{For } \mu \text{ a.e. } x \in T$$
verify that we may pick continuous representatives of φ_n.

Set
$$K_N(x,y) = \sum_{k=1}^{N} \lambda_n \varphi_n(x)\varphi_n(y)$$
By the basic spectral theory,
$$Af = \sum_{n=1}^{\infty} \lambda_n \langle \varphi_n, f\rangle \varphi_n \quad , \quad f \in L^2(d\mu),$$
where the convergence is in $L^2(d\mu)$. Especially,
$$\langle Af, f\rangle = \sum_{n=1}^{\infty} \lambda_n \langle \varphi_n, f\rangle^2$$
If we denote
$$A_N f = \sum_{n=1}^{N} \lambda_n \langle \varphi_n, f\rangle ,$$
and the kernel of A_N is K_N.

Since $\langle A_N f, f\rangle = \sum_{n=1}^{N} \lambda_n \langle \varphi_n, f\rangle^2$,

we have $A_N \leq A$. We next claim that

(15) $\sup_{x\in T} \sum_{n=1}^{\infty} \lambda_n |\varphi_n(x)|^2 \leq C < \infty$

Fix $x_0 \in T$ and choose $f = \frac{\chi_{B(0,r)}}{\mu(B(0,r))}$ in the inequality $\langle A_N f, f\rangle \leq \langle A f, f\rangle$, which takes the form

$$\sum_{n=1}^{N} \lambda_n \left(\frac{1}{\mu(B(x_0,r))} \int_{B(x_0,r)} \varphi_n(x)\,dx\right)^2$$

$$\leq \left(\frac{1}{\mu(B(x_0,r))}\right)^2 \int_{(B(x_0,r))^2} K(x,y)\,\mu(dx)\,\mu(dy)$$

In the limit $r \to 0^+$, by continuity we obtain

(16) $\sum_{n=1}^{N} \lambda_n |\varphi_n(x_0)|^2 \leq K(x_0,x_0)$.

Letting $N \to \infty$, (15) follows with $C = \sup_{x_0\in T} K(x_0,x_0)$.

Let $\delta > 0$ be given. Choose $y_0 \in T$ and pick $N_0 \geq 1$ so large by (15) that

$$\sum_{n=N_0}^{\infty} \lambda_n |\varphi_n(y_0)|^2 < \delta^2$$

Then by Cauchy-Schwartz for any $x \in T$

$$\sum_{n=N_0}^{\infty} \lambda_n |\varphi_n(x)|\,|\varphi_n(y_0)|$$

$$\leq \left(\sum_{n=N_0}^{\infty} \lambda_n |\varphi_n(x)|^2\right)^{1/2} \left(\sum_{n=N_0}^{\infty} \lambda_n |\varphi_n(y_0)|^2\right)$$

$$\leq \delta \cdot C^{1/2}$$

Hence the series

$$\tilde{K}(x,y_0) := \sum_{n=1}^{\infty} \lambda_n \varphi_n(x) \varphi_n(y_0)$$

converges uniformly in x. If $f \perp \varphi_n$ in $L^2(d\mu)$ for all n, then by the uniform convergence

$$\tilde{K}(\cdot, y_0) \perp f .$$

On the other hand, for any $n \geq 1$ we have

$$\int_T \tilde{K}(x,y_0)\, \varphi_n(x)\, d\mu(x) = \lambda_n \varphi_n(y_0)$$

$$= \int_T K(y_0,x)\, \varphi_n(x)\, d\mu(x) = \int_T K(x,y_0) \varphi_n(x)\, d\mu(x).$$

since φ_n is an eigenfunction, and K is symmetric. It follows that

(17) $$\tilde{K}(x,y_0) = K(x,y_0)$$

as elements in $L^2(d\mu)$, and since μ is nontrivial everywhere, continuity yields that (17) holds for every x (and y_0).

In particular, we have

(18) $$K(x,x) = \sum_{n=1}^{\infty} \lambda_n (\varphi_n(x))^2$$

By Dini's theorem (Exercise), the convergence in (18) is uniform, and we may choose N_0 so that

$$\sum_{n=N_0}^{\infty} \lambda_n |\varphi_n(x)|^2 < \delta \quad \text{for all } x \in T.$$

Then the argument at the end of page 94 verifies that convergence in (17) is uniform both in x and y, proving first part of the theorem. The last claim on the trace $\sum_{n=1}^{\infty} \lambda_n$ follows by integrating (18) over x. $\square$

Theorem (6.15) (Ito-Nishio). Assume that X and $X_1, X_2, \ldots$ are E-valued random variables, where E is a separable Banach space. Assume also that X_n:s are symmetric and almost surely

$$\lambda(X_n) \to \lambda(X) \quad \text{for any } \lambda \in E'.$$

Then $X_n \to X$ almost surely (in the norm-topology) in E.

Proof. This will be a guided exercise. $\square$

Theorem (6.16) (Karhunen-Loeve, version 2.) Assume that X is a centered Gaussian process on $\mathbb{T}$. Let $C = C_X \in C(\mathbb{T}^2)$ and (T, μ) be as in thm 6.14. Then

$$(19) \qquad X \sim \sum_{k=1}^{\infty} \lambda_k^{1/2} A_k e_k \qquad \text{(a.s. convergence in } L^2(d\mu))$$

where $A_k \sim N(0,1)$ are i.i.d and λ_k:s, e_k's are the eigenvalues and eigenfunctions of the covariance operator.

In case X has a.s. continuous realizations, the series (19) converges a.s. uniformly!

Proof. Recall that covariance operator is the operator A of Thm 6.14, i.e. it has the kernel A. According to Thm 6.14, A is nuclear, and then (19) and the convergence in $L^2(d\mu)$ follows from Corollary 5.3.

Assume then that X has almost surely continuous realizations. The convergence in (19) follows from Thm 6.15 as soon as we show that

$$(20) \quad \lim_{N\to\infty} \sum_{k=1}^{N} \lambda_k^{1/2} A_k \int_{T} e_k(t)\,\nu(dt) = \int_{T} X(t)\,\nu(dt)$$

almost surely, where ν is given finite Borel measure on T (recall that the dual of $C(T)$ is $M(T)$, the Banach space of finite Borel measures on T). Denote

$$\sum_{k=1}^{N} \lambda_k^{1/2} A_k e_k(t) =: X_N(t)$$

and estimate

$$\begin{aligned}
&\mathbb{E} \left| \int_T (X_N(t) - X(t))\,\nu(dt) \right| \\
\le\ & \mathbb{E} \int_T \mathbb{E}\,|X_N(t) - X(t)|\,|\nu|(dt) \\
\le\ & \int_T \left(\mathbb{E}\,|X_N(t) - X(t)|^2\right)^{1/2} |\nu|(dt) \\
=\ & \int_T \left(C(t,t) - \sum_{k=1}^{N} \lambda_k (e_k(t))^2 \right)^{1/2} |\nu|(dt) \\
&\xrightarrow[0]{} 0 \quad \text{as } N\to\infty
\end{aligned}$$

since by Thm 6.14 (Mercer) the integrand is dominated by $\sup_{t\in T} C(t,t)$, and converges to 0 for every $t \in T$. Hence we have convergence in probability in (20), and finally Thm 3.3 (applied in the case $E = \mathbb{R}$) upgrades it to convergence in probability. □

<u>Example</u> We have a.s. uniform convergence in the Karhunen-Loewe decomposition (3), p. 68, for the Brownian bridge, although this follows also from Thm. 4.18

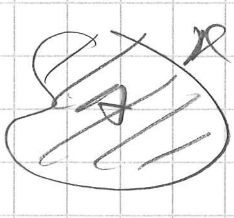
<u>Exercise</u> Try to find the Karhunen-Loeve expansion for the <u>Brownian sheet on $\mathbb{R}^d$</u> considered in Example 2 on p. 81 (there $d=2$). Thus, define B_t for $t \in [0,1]^d$ as the centered Gaussian field with the covariance structure

$$\mathbb{E}\, X_t X_s = \prod_{k=1}^{d} \min(t_k, s_k)$$

for $s, t \in [0,1]^d$.

7. GAUSSIAN ISOPERIMETRIC INEQUALITY AND LARGE DEVIATIONS

Recall the classical isoperimetric inequality in plane: if rectifiable Jordan curve γ encloses the area A, then



$$\big(\underbrace{\mathcal{H}^1}_{\text{1-Hausdorff measure (length)}}(\gamma)\big)^2 \geq 4\pi A$$

In $\mathbb{R}^n$ this inequality takes the form: if $A \subset \mathbb{R}^d$

$$|A|_{Leb(\mathbb{R}^d)}^{(d-1)/d} \leq \frac{1}{d\,\omega_d^{1/d}}\, \mathcal{H}^{d-1}(\partial A), \qquad \omega_d = \underbrace{|B(0,1)|}_{\text{unit ball in } \mathbb{R}^d}$$

and one may replace $\mathcal{H}^{d-1}$ on the right hand side by the $(d-1)$-dimensional Minkowski content.

One may also express the above in an integrated form: given $A \subset \mathbb{R}^d$ define

$$A_r = A + B(0,r) \qquad \text{(open ball)}$$

(the r-fattening of A). Then, if B is a ball with $|B| = |A|$, we have (A e.g. a Borel set)

$$|A_r| - |A| \geq |B_r| - |B|.$$

Or, equivalently, one has

$$|A_r| \geq \left(|A|^{1/d} + r\,\omega_d^{1/d}\right)^d$$

Our aim is to prove the corresponding result for Gaussian measure on $\mathbb{R}^d$.

Theorem 7.1 ([Borell], [Sudakov-Tsirelson])

Denote by γ_d the distribution of the Gaussian $N(0, I_d)$ variable in $\mathbb{R}^d$. Then for any Borel-set $A \subset \mathbb{R}^d$

$$\gamma(A_r) \geq \gamma(H_r),$$

where $A_r := A + B(0,r)$ and H is a half plane $H := \{x_1 > u\}$ with the same γ-measure as A.

An equivalent version of the statement is that

$$(1) \qquad \boxed{\gamma(A_r) \geq \Phi\left(\Phi^{-1}(\gamma(A)) + r\right)}$$

where

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-x^2/2}.$$

We also denote $\varphi(x) = \Phi'(x) = \frac{1}{2\pi} e^{-x^2/2}$.

For sets with rectifiable boundary the above may be stated in a more classical form:

$$\lim \frac{\gamma(A_r) - \gamma(A)}{r} = (2\pi)^{-\frac{d}{2}} \int_{\partial A} e^{-|x|^2/2} \, d\mathcal{H}^{d-1}(x)$$

$$\geq \varphi(\Phi^{-1}(\gamma(A))).$$

We shall present the short proof of Michael Ledoux that is a variant of Bakry's ingenious argument. For that end we need some easy properties of the Ornstein-Uhlenbeck semigroup.

Def. Let $f \in L^1(\gamma)$. Then we define for $t \geq 0$

$$P_t f(x) = \int_{\mathbb{R}^d} f(e^{-t}x + (1-e^{-2t})^{1/2} y)\, d\gamma(y)$$

(the Ornstein-Uhlenbeck

Lemma 7.2. (i) P_t is well-defined for $f \in L^1(\gamma)$.

(ii) $P_{t+s} f = P_t(P_s f)$, $t, s \geq 0$ (P_t is a semigroup)

(iii) $\|P_t f\|_{L^p(\gamma)} \leq \|f\|_{L^p(\gamma)}$ for any $f \in L^p(\gamma)$, $1 \leq p \leq \infty$

(iv) $P_t f \underset{t\to\infty}{\to} \int_{\mathbb{R}^d} f\, d\gamma$ uniformly (and the first derivatives also) if $f \in C_0^\infty(\mathbb{R}^d)$.

(v) $\int_{\mathbb{R}^d} (P_t f) g\, d\gamma = \int_{\mathbb{R}^d} f (P_t g)\, d\gamma$ (symmetry of the semigroup)

Proof. (i) if $t > 0$ then $|e^{-t}x + (1-e^{-2t})^{1/2} y| \leq |y|$ for $|y| \geq C(x)$, and using this one easily checks by the change of variable $e^{-t}x + (1-e^{-2t})^{1/2} y = z$ that $y \mapsto f(e^{-t}x + (1-e^{-2t})^{1/2} y)$ is in $L^1(\gamma)$.

(ii) Observe that if $Y_0 \sim N(0, \mathrm{Id}_d)$, we have

$$P_t f(x) = \mathbb{E}(e^{-t}x + (1-e^{-2t})^{\frac{1}{2}} Y_0)$$

If also $Y_1 \sim N(0, \mathrm{Id}_d)$ and $Y_1 \perp Y_0$ we obtain

$$P_t(P_s f)(x) = \mathbb{E}\, P_s f(e^{-t}x + (1-e^{-2t})^{\frac{1}{2}} Y_0)$$

$$= \mathbb{E}\left(f(e^{-s}(e^{-t}x + (1-e^{-2t})^{\frac{1}{2}} Y_0) + (1-e^{-2s})^{\frac{1}{2}} Y_1\right)$$

$$= \mathbb{E}\, f(e^{-(t+s)}x + Z),$$

where $Z \sim N(0, \sigma^2)$ with $\sigma^2 = (e^{-s}(1-e^{-2t})^{1/2})^2 + ((1-e^{-2s})^{1/2})^2 = 1 - e^{-2(s+t)}$, which proves the claim in view of ( ).

(iii) If also $X_0 \sim N(0, \mathrm{Id}_d)$ and $X_0 \perp Y_0$, we have:

$$\int |P_t f(x)|^p \, d\gamma(x) = \mathbb{E}\, |P_t(X_0)|^p$$

$$= \mathbb{E}\, \left| \mathbb{E}\big(f(e^{-t}X_0 + (1-e^{-2t})^{1/2} Y_0) \,\big|\, X_0\big) \right|^p$$

$$\leq \mathbb{E}\, |f(\underbrace{e^{-t}X_0 + (1-e^{-2t})^{1/2} Y_0}_{\sim N(0,1)})|^p = \int_{\mathbb{R}^d} |f(y)|^p \, d\gamma(y),$$

where we applied the L^p-contractivity of the conditional expectation. Case $p=\infty$ is trivial.

(iv) Exercise

(v)
$$\int_{\mathbb{R}^d} P_t f(x) g(x) \, d\gamma(y) = \mathbb{E}\, f(e^{-t} X_0 + (1-e^{-2t})^{1/2} Y_0) g(X_0)$$

and
$$\int_{\mathbb{R}^d} f(x) P_t g(x) \, d\gamma(y) = \mathbb{E}\, f(X_0) g(e^{-t} X_0 + (1-e^{-2t})^{1/2} Y_0).$$

It remains to observe that $(e^{-t}X_0 + (1-e^{-2t})^{1/2} Y_0, X_0)$ has the same covariance, and hence distribution, as $(X_0, e^{-t}X_0 + (1-e^{-2t})^{1/2} Y_0)$. $\square$

Lemma 7.3 Assume $f, g \in C_0^\infty(\mathbb{R}^d)$. Then

(i) $\frac{d}{dt} P_t f = P_t(Lf) = L(P_t f)$, where

$$Lg(x) := \Delta g(x) - x \cdot \nabla g(x)$$

(ii) $$-\int_{\mathbb{R}^d} f(x)\, Lg(x)\, d\gamma(x) = \int_{\mathbb{R}} \nabla f(x)\cdot \nabla g(x)\, d\gamma(x).$$

Proof. Let us first assume that $\boxed{f(x) = e^{i\,\xi\cdot x}}$
(f is smooth but not compactly supported)
Then
$$P_t f(x) = \mathbb{E}\, e^{i\,\xi\cdot e^{-t}x}\, e^{i\sqrt{1-e^{-2t}}\,Y_0\cdot\xi}$$
$$= \exp\left(e^{-t} i\,\xi\cdot x - \frac{1}{2}(1-e^{-2t})|\xi|^2\right)$$

so that $$\left[\frac{d}{dt} S_t f(x)\right]_{t=0} = \left(-i\,\xi\cdot x - |\xi|^2\right) e^{i\,\xi\cdot x}$$
$$= (\Delta - x\cdot\nabla)\, f(x).$$

If $f \in C_0^\infty$, we may write $f = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{i\xi\cdot x}\hat{f}(\xi)\,d\xi$,

where now $\hat{f} \in S(\mathbb{R}^d)$ is smooth with polynomially decaying derivatives. By linearity and easy reasoning we obtain
$$\left[\frac{d}{dt} S_t f(x)\right]_{t=0} = (2\pi)^{-d/2}\int_{\mathbb{R}^d} (\Delta_x - x\cdot\nabla_x)\, e^{i\xi\cdot x}\hat{f}(\xi)\,d\xi$$
$$= (\Delta - x\cdot\nabla)\, f.$$
The other claims are an exercise. □

We will then denote
$$U = \varphi\circ\Phi^{-1} : [0,1] \to [0,\infty)$$

Our aim is to prove the following striking result:

Theorem 7.4. (Bobkov) If $F: \mathbb{R}^d \to [0,1]$ is Lipschitz, then

$$U\Big(\int_{\mathbb{R}^d} F\,d\gamma\Big) \leq \int_{\mathbb{R}^d} \sqrt{U^2(F) + |\nabla F|^2}\,d\gamma$$

Proof of Thm 7.1 (Assuming Thm 7.4) We may assume by approximation that A is a finite union of open balls, whence $\gamma(\partial A) = 0$.
We apply Thm 7.4' to the function

$$F_r(x) := \Big(1 - \tfrac{1}{r} d(x, A)\Big)^+ \qquad \Big|\ x^+ = \max(0, x)$$

Then, as $r \to 0^+$,

- $F_r(x) \to \chi_A(x)$, especially $F_r \to \chi_A$ for a.e. x
- $U(F_r) \to 0$ for a.e. x since $U(0) = U(1) = 0$
- $\nabla F_r(x) = \begin{cases} 0, & x \in A \cup (\bar{A}_r)^c \\ 1/r, & x \in \operatorname{int}(A_r \setminus A) \end{cases}$

$$\Rightarrow \nabla F_r = \chi_{A \setminus A_r} \text{ for a.e. } x \quad (\text{as } |\partial A_r| = 0)$$

Thus Thm 7.4 implies that

$$(2)\quad \boxed{U(\gamma(A))} \leq \underbrace{\liminf_{r \to 0^+} \frac{1}{r}\big(\gamma(A_r) - \gamma(A)\big)}_{=: \gamma_s(\partial A)},$$

where for smooth enough sets $B \subset \mathbb{R}^d$ we set

$$\gamma_s(B) := \int_B (2\pi)^{-d/2} e^{-|x|^2/2}\, \mathcal{H}^{d-1}(dx).$$

($(d-1)$ dimensional)

(2) is already a version of the Gaussian isoperimetry. To get the integrated form we denote

$$h(r) := \Phi^{-1}(\gamma(A_r)) \quad , r \geq 0$$

and observe that

$$\frac{d}{dr} h(r) = \frac{\gamma_s(\partial A_r)}{\varphi \circ \Phi^{-1}(\gamma(A_r))} = \frac{\gamma_s(\partial A_r)}{U(\gamma(A_r))} \overset{(2)}{\geq} 1 \;,$$

so $h(r) - h(0) \geq r$, which is exactly the statement of Thm 7.1. $\square$

In order to prove Thm 7.4 we use our knowledge of the Ornstein-Uhlenbeck semigroup and the elementary observation

(3) $$U U'' = -1 ,$$

whose proof we leave as an exercise.

<u>Proof of Thm 7.4.</u> By simple approximation we may assume that $f \in C_0^\infty(\mathbb{R}^d)$. Consider the function $(t \geq 0)$

$$F(t) := \int_{\mathbb{R}^d} \sqrt{U^2(P_t f) + |\nabla P_t f|^2}\, dx .$$

We claim that

(4) $$F'(t) \leq 0 .$$

Once this is shown it follows that

$$\mho\Big(\int_{\mathbb{R}^d} F\,d\gamma\Big) \overset{\text{Lemma 7.2.(i)}}{=} \lim_{t\to\infty} F(t) \le F(0) = \int_{\mathbb{R}^d} \sqrt{\mho^2(f) + |\nabla f|^2}\, d\gamma,$$

i.e. the claim of Thm 7.4.

In order to prove (4), let us denote $g := P_t f$ and

$$K(g) := K(P_t f) = \sqrt{\mho^2(g) + |\nabla g|^2}$$

With these notations, we obtain (integrals below are always over $\mathbb{R}^d$, against the measure γ, and we also use the shorthand $\mho = \mho(g)$, $\mho' = \mho'(g)$) by Lemma 7.3

$$F'(t) = \int \frac{1}{\sqrt{K(g)}} \left[\mho\mho' Lg + \nabla g \cdot \nabla(Lg) \right]$$

Here

$$\int \frac{1}{\sqrt{K(g)}} \mho\mho' Lg \overset{\text{Lemma 7.3(ii)}}{=} - \int \nabla\Big(\frac{\mho\mho'}{\sqrt{K(g)}}\Big) \nabla g$$

$$= - \int \frac{|\nabla g|^2}{\sqrt{K(g)}} \left(\mho'^2 + \mho\mho''\right)$$

$$+ \int \frac{\mho\mho'}{(K(g))^{3/2}} \left(\mho\mho' |\nabla g|^2 + \nabla g\,(D^2 g\, \nabla g)\right)$$

Moreover, $\nabla g \cdot \nabla(Lg) = \nabla g \cdot L(\nabla g) - |\nabla g|^2$

Thus we may apply the integration by parts formula (Lemma 7.3(ii)) again to obtain

$$\int \frac{1}{\sqrt{K(g)}} \nabla g \cdot \nabla (Lg)$$

$$= - \int \frac{|\nabla g|^2}{\sqrt{K(g)}} + \int \frac{\nabla g}{\sqrt{K(g)}} L(\nabla g),$$

where

$$\int \frac{\nabla g}{\sqrt{K(g)}} L(\nabla g) = - \int \nabla \left(\frac{\nabla g}{\sqrt{K(g)}} \right) \cdot \overset{\text{in } \mathbb{R}^{d^2}}{\Downarrow} D^2 g$$

$$= - \int \frac{|D^2 g|^2}{\sqrt{K(g)}} + \int \frac{1}{(K(g))^{3/2}} \Big[v v' \nabla g \cdot (D^2 g \nabla g) +$$

$$+ (\nabla g \otimes D^2 g \nabla g) \cdot D^2 g \Big]$$

By observing that $(\nabla g \otimes D^2 g \nabla g) \cdot D^2 g = |D^2 g \nabla g|^2$ and using (3) we may collect everything together to obtain

$$F'(t) = - \int (K(g))^{-3/2} R, \qquad \text{where}$$

$$\begin{aligned} R := \; & |\nabla g|^2 (v^2 + |\nabla g|^2)(v'^2 - 1) \\ & - v^2 v'^2 |\nabla g|^2 - v v' \nabla g \cdot (D^2 g \nabla g) \\ & + |\nabla g|^2 (v^2 + |\nabla g|^2) \\ & + |D^2 g|^2 (v^2 + |\nabla g|^2) \\ & - v v' \nabla g \cdot (D^2 g \nabla g) \\ & - |D^2 g \nabla g|^2 . \end{aligned}$$

After cancelling terms we are left with

$$R = |\nabla g|^4 U'^2 - 2UU'\nabla g\cdot(D^2 g \nabla g) + |D^2 g|^2 U^2 + \underbrace{(|D^2 g|^2 |\nabla g|^2 - |D^2 g \nabla g|^2)}_{\geq 0}$$

$$\geq |\nabla g|^4 U'^2 - 2|U||U'| \|D^2 g\| |\nabla g|^2 + |D^2 g|^2 U^2$$

$\|D^2 g\|$: matrix norm $\leq |D^2 g|$ (= Hilbert-Schmidt norm)

$$\geq (|\nabla g|^2 U' - U|D^2 g|)^2 \geq 0 \qquad \square$$

Let $X \sim N(0,C)$ be a non-degenerate Gaussian on $\mathbb{R}^d$. Thus

$$X \sim C^{1/2} Y_0, \quad \text{where } Y_0 \sim N(0, \mathrm{Id}),$$

or, vice versa, $Y_0 \sim C^{-1/2} X$. If $A \subset \mathbb{R}^d$ is Borel and $r > 0$, by the Gaussian isoperimetry (Thm. 7.1) we thus have (we apply the isoperimetry on $C^{-1/2}A$):

$$\Phi^{-1}(\mathbb{P}(C^{-1/2}X \in (C^{-1/2}A + B(0,r))) \geq r + \Phi^{-1}(\mathbb{P}(C^{-1/2}X \in C^{-1/2}A))$$

or $$\Phi^{-1}(\mathbb{P}(X \in (A + C^{1/2}B(0,r)))) \geq r + \Phi^{-1}(\mathbb{P}(X \in A))$$

Now $C^{1/2}B(0,r) = r B_{H_X}$ ($r\cdot$ unit ball of the Cameron-Martin space by Thm. 5.4). We have shown

Lemma 7.5 For a finite-dimensional Gaussian variable

$$\Phi^{-1}(\mathbb{P}(X \in A + rB_{H_X})) \geq r + \Phi^{-1}(\mathbb{P}(X \in A)).$$

(the case C degenerate reduces to this, and the above clearly holds even if X takes values in E (a Banach space), and H_X is finite-dimensional)

Lemma 7.6 Let μ be a Gaussian measure on a separable Banach space E. Then

$$B_{H_\mu}$$

(the unit ball of the Cameron-Martin space) is compact (in the norm topology of E.

Proof Guided exercise. $\square$

Theorem 7.7 Let μ be a Gaussian measure on a separable Banach space E. Then, if $A \subset E$ is closed, $r > 0$, we have that

$$(4) \qquad \Phi^{-1}(\mu(A + r B_{H_\mu})) \geq r + \Phi^{-1}(\mu(A))$$

Proof Step 1. • Claim: $A + r B_{H_\mu}$ is closed so that LHS in (4) is well-defined.

To see this, let $x_k = y_k + z_k \to x_\infty \in E$ as $k \to \infty$. where $y_k \in A$, $z_k \in r B_{H_\mu}$. By moving to a subsequence, by Lemma 7.6 we may assume that $z_k \to z_\infty \in r B_{H_\mu}$. Then $y_k = x_k - z_k \to y_\infty \in A$ as A was closed. Hence $z_\infty = y_\infty + z_\infty \in A + r B_{H_\mu}$.

Step 2. Claim: may assume that A is compact.

This follows by choosing $K_\ell \subset A$ so that K_ℓ ($\ell \geq 1$) is compact, and $\mu(K_\ell) \nearrow \mu(A)$ as $\ell \to \infty$. Writing (4) for K_ℓ in place of A we obtain

$$\Phi^{-1}(\mu(A + r B_{H_\mu})) \geq \Phi^{-1}(\mu(K_\ell + r B_{H_\mu}$$

$$\geq r + \Phi^{-1}(\mu(K_\ell)).$$

Simply let $\ell\to\infty$.

<u>Step 3.</u> Case $E=\ell^2$ (a separable Hilbert space).

By Thm 5.4 we may assume that in suitable coordinates the covariance operator C_μ is diagonal:

$$C_\mu=\lceil\lambda_1,\dots,\lambda_n,\dots\rfloor,$$

(= diagonal operator with diagonal entries λ_k)

where $\lambda_k>0\ \forall k$, $\lambda_1\geq\lambda_2\geq\dots$, as we may assume that μ is not finite-dimensional. For fixed $N\geq1$ denote by $P_N:\ell^2\to\mathbb{R}^N$ the 'projection':

$$P_N(x_1,x_2,\dots)=(x_1,\dots,x_N),$$

and let μ_N be the measure on $\mathbb{R}^N$, where

$$\mu_N(A):=\mu(P_N^{-1}(A)),\quad A\in\mathcal{B}(\mathbb{R}^N).$$

Obviously μ_N is a Gaussian with the diagonal covariance

$$C_{\mu_N}:=\lceil\lambda_1,\dots,\lambda_N\rfloor$$

Moreover, $\quad B_{H_{\mu_N}}=\{(x_1,\dots,x_N):\sum_{k=1}^N\lambda_k^{-1}x_k^2\leq1\}.$

We claim that for any compact set $K\subset\ell^2$

(5) $$\mu(K)=\lim\mu_N(P_NK).$$

To see this, note that writing $U_N:=\{x\in\ell^2:x_j=0,\ 1\leq j\leq N\}$,

$$\lim_{N\to\infty}\mu_N(P_NK)=\lim_{N\to\infty}\mu(P_NK+U_N)$$
$$=\lim_{N\to\infty}\mu(K+U_N),$$

since $\quad P_N K + U_N = K + U_N$! Now

$$\lim_{N\to\infty} \mu(K+U_N) = \mu\Big(\bigcap_{N\ge1}^{\infty}(K+U_N)\Big) = \mu(K)$$

and (5) follows. Here we observed that $\bigcap_{N\ge1}^{\infty}(K+U_N) = K$, since if $x \in \bigcap_{N\ge1}^{\infty}(K+U_N)$, we may write $x = k_N + u_N$, where $k_N \in K \ \forall N\ge 1$, $u_N \in U_N$. By compactness we may assume that $k_N \to k \in K$ (moving to a subsequence if needed). Then $P_N x = P_N k_{N'}$ for all $N' > N$, especially letting $N' \to \infty$ we have $P_N x = P_N k \ \forall N \ge 1$, whence $x = k \in K$.

Clearly $P_N(r B_{H_\mu}) = r B_{H_{\mu_N}}$. Since $A + r B_{H_\mu}$ is compact, we may finally apply (5) to deduce

$$\phi^{-1}(\mu(A + rB_{H_\mu})) = \lim_{N\to\infty} \phi^{-1}(\mu_N(P_N(A + rB_{H_\mu})))$$

$$= \lim_{N\to\infty} \phi^{-1}(\mu_N(P_N A + rB_{H_{\mu_N}}))$$

$$\ge r + \lim_{N\to\infty} \phi^{-1}(\mu_N(P_N A)) = r + \phi^{-1}(\mu(A)).$$

Lemma 7.5

<u>Step 4.</u> • Case: $E = C(0,1)$.

The natural linear embedding $C(0,1) \subset L^2(0,1) \simeq \ell^2$ is continuous, and it is not difficult to check that $C(0,1)$ is a Borel subset of $L^2(0,1)$. Actually, one easily checks that any closed ball $B \subset B(0,1)$ is a Borel subset of $L^2(0,1)$, whence $\mathcal{B}(C(0,1)) = \mathcal{B}(L^2(0,1)) \cap C(0,1)$. The claim now follows by Theorem 4.28, as we also note that $A \subset C(0,1)$ compact implies that $A \subset L^2(0,1)$ compact by continuity of the embedding $C(0,1) \subset L^2(0,1)$.

Step 5. Case: general separable E.

By basic functional analysis, such an E embedds into $C([0,1])$ isometrically to a closed subspace, and the claim follows. $\square$

We shall apply the above on large deviations.

Definition Let μ be a Gaussian measure on a separable Banach space E. For any subset $A \subset E$ we set

$$J_\mu(A) := \begin{cases} \inf_{h \in A} \frac{\|h\|^2_{H_\mu}}{2} & \text{if } H_\mu \cap A \neq \emptyset \\ \infty & \text{if } H_\mu \cap A = \emptyset. \end{cases}$$

Theorem ~~7.~~ 7.8. (Gaussian large deviation principle)

For any open set $G \subset E$ we have

$$(5) \quad \liminf_{r \to \infty} \frac{\log \mu(rG)}{r^2} \geq -J_\mu(G)$$

and for any closed set $F \subset E$

$$(6) \quad \limsup_{r \to \infty} \frac{\log \mu(rF)}{r^2} \leq -J_\mu(F)$$

The set $A \subset E$ is regular if

$$J_\mu(\operatorname{int} A) = J_\mu(\overline{A})$$

A fortiori, for any regular set $A \subset E$ it holds that

$$\lim_{r \to \infty} \log \frac{\mu(rF)}{r^2} = -J(A).$$

Proof. Assume $G \subset E$ is open, and $\exists\, h \in H_\mu \cap G$. Pick $\delta > 0$ so small that $h + B(0,\delta) \subset G$. By Borell's Thm in Exercise 23 we have

$$\mu(rG) \geq \mu(rh + B(0,\delta r)) \geq \mu(B(0,\delta r))\, e^{-r^2 \|h\|^2_{H_\mu}/2}$$

Since $\mu(B(0,\delta r)) \nearrow 1$ as $r \to \infty$, taking logarithms we obtain in the limit $r \to \infty$ the inequality (5)

Assume next that $F \subset E$ is closed. It is enough to consider the case $J_\mu(F) > 0$. Choose arbitrary $\varrho > 0$ so that

$$\inf_{h \in F} \|h\|_{H_\mu} > \varrho > 0$$

(the infimum could take the value ∞). Then

$$\varrho B_{H_\mu} \cap F = \emptyset,$$

and by compactness of H_μ (an closedness of F) we may pick $\delta > 0$ so that

$$(\varrho B_{H_\mu} + \delta \overline{B(0,1)}) \cap F = \emptyset.$$

By Theorem 7.7 we obtain

$$\mu(rF) \leq \mu(E \setminus (r\varrho B_{H_\mu} + r\delta \overline{B(0,1)}))$$

$$= 1 - \mu(r\varrho B_{H_\mu} + \overline{B(0,r\delta)})$$

$$\leq 1 - \Phi(r\varrho + \Phi^{-1}(\overline{B(0,r\delta)}))$$

If r is large enough so that $\Phi^{-1}(\overline{B(0,r\delta)}) \geq 0$, this and Lemma 4.7 yield (again for large enough r)

$$\mu(rF) \le e^{-r^2 g^2/2}, \quad r \ge r_0(g).$$

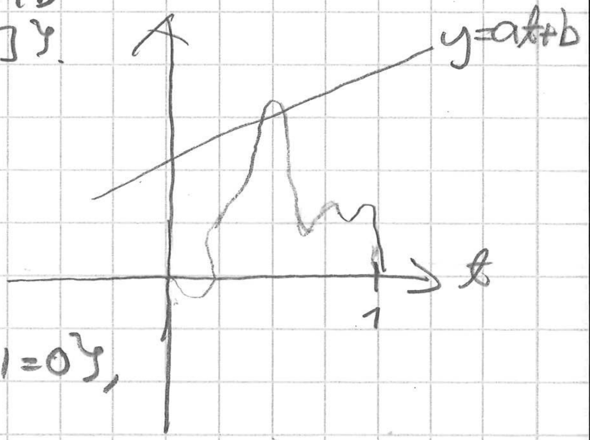
This yields that $\overline{\lim_{r\to\infty}} \frac{\log(\mu(rF))}{r^2} \le -g^2/2$,
and the (6) follows by taking sup over the g's allowed for.

Finally the claim for regular sets follows immediately from (5) and (6) $\square$

<u>Example</u> Let X be the Brownian bridge on $[0,1]$. Let $a,b>0$. We compute the asymptotics of

$$b_\lambda := \log \mathbb{P}\left(X(t) \ge \lambda(at+b) \text{ for some } t\in[0,1]\right)$$

Let $A = \{\varphi \in C(0,1) \mid \varphi(t) \ge at+b \text{ for some } t\in[0,1]\}$.



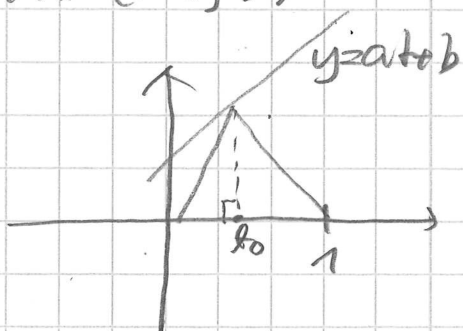
Then A is closed in $C(0,1)$.
If $h\in A \Rightarrow (1+\varepsilon)h \in \operatorname{int}(A)$,
which yields that A is regular (why?). We have

$$H_{\mu_X} = \{\varphi \in W^{1,2}(0,1) \mid \varphi(0)=\varphi(1)=0\},$$
$$\|h\|^2_{H_{\mu_X}} = \int_0^1 |h'(t)|^2\,dt.$$

To compute $J_{\mu_X}(A)$, we need to find (why?) $t_0\in(0,1)$ so that

$$\int_0^1 |\varphi'_{t_0}(t)|^2\,dt = \min,$$

where φ_{t_0} is as in the picture.



Since
$$\int_0^1 |\varphi'_{t_0}|^2 dt = t_0\left(\frac{at_0+b}{t_0}\right)^2 + (1-t_0)\left(\frac{at_0+b}{1-t_0}\right)^2$$
$$= \frac{(at_0+b)^2}{t_0(1-t_0)} = -a^2 + \frac{b^2+(2ab+a^2)t_0}{t_0(1-t_0)},$$
an easy computation shows that at the minimum
$$t_0 = \frac{a}{a+2b},$$
and
$$\log \mathbb{P}(X(t) \geq \lambda(at+b) \text{ for some } t \in [0,1])$$
$$\sim -2\lambda^2 b(a+b). \quad \square$$

<u>Remark.</u> The previous example is a special case of more general result obtained in the Exercises.

8. MULTIPLICATIVE CHAOS

In this last section of the lectures we take a look at the basics of multiplicative chaos. The basic idea is to construct random measures that have special scaling properties. Especially, heuristically the 'density' of the measure should possess almost independent factors which yield fluctuations at different dyadic scales.

The idea of chaos measures dates to Kolmogorov's work on turbulence, Mandelbrot in early 1970's proposed a good model (Mandelbrot cascades) He also proposed a continuous model, that actually appears in quantum field theory (continuous chaos). Kahane (and Peyrier) developed basic mathematical theory of cascades (mid 1970's). Kahane proved his famous inequality and developed nontrivial theory of continuous subcritical chaos in mid-1980's. The theory has become quite active during last 12 years, and it has connections to e.g.

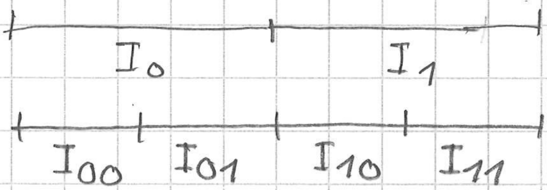
- scaling limits of 2-dim stat. physics models
- conformal weldings in SLE
- mathematical finance
- number theory (Riemann zeta)
- random matrices

⋮

Basic idea of multiplicative chaos is most easily explained through

Mandelbrot cascades

Fix a random variable $\tilde{U}$ with $\mathbb{E}\tilde{U} = \frac{1}{2}$. We aim to construct a random measure on $[0,1]$ with the help of U_0. For that end, consider dyadic decomposition of $[0,1]$:



$\vdots$

Thus the level $n \geq 1$ dyadic closed intervals are coded by I_σ , $\sigma \in \{0,1\}^n$, as in the picture. Set $S = \bigcup_{n=1}^{\infty} \{0,1\}^n$. We say that $\sigma \leq \sigma'$ if σ has length n, σ' length $m \geq n$ and the n-initial segment of σ' equals σ. E.g.

$$01 \leq 011 \quad \text{etc.}$$

- For each $\sigma \in S$ associate a random variable $U_\sigma \sim \tilde{U}$ so that U_σ:s are i.i.d.

- The n:th level approximation, μ_n, of the final cascade measure has constant density on level n intervals. Moreover we set

$$\boxed{\mu_n(I_\sigma) = \prod_{\sigma' \leq \sigma} U_{\sigma'}} \quad , \quad \sigma \in \{0,1\}^n$$

This defines μ_n uniquely.

Eg. $\mu(I_{011}) = U_0 U_{01} U_{011}$

Theorem 8.1. Almost surely,

$$\mu_n \xrightarrow[n\to\infty]{w} \mu ,$$

where μ is a random non-negative measure on $[0,1]$. Measure μ is called the multiplicative cascade measure (generated by $\tilde{U}$)

Proof. Let us denote $M_n := \mu_n([0,1])$ (the total mass of μ_n). We claim that

(1) $M_n \geq 0$ is a martingale with respect to the σ-algebra sequence $(F_n)_{n\geq 1}$, where $F_n := \sigma\left(U_\sigma : \sigma \in \bigcup_{k=1}^{n} \{0,1\}^k\right)$

To see that, observe that

$$\mathbb{E}(M_{n+1} \mid F_n) = \mathbb{E}\left(\sum_{\sigma\in\{0,1\}^n} \left(\mu_{n+1}(I_{\sigma 0}) + \mu_{n+1}(I_{\sigma 1})\right) \Big| F_n\right)$$

$$= \mathbb{E}\left(\sum_{\sigma\in\{0,1\}^n} \mu_n(I_\sigma)(U_{\sigma 0} + U_{\sigma 1})\right)$$

$$= \left(\sum_{\sigma\in\{0,1\}^n} \mu_n(I_\sigma)\, \mathbb{E}(U_{\sigma 0} + U_{\sigma 1})\right)$$

$$= \sum_{\sigma\in\{0,1\}^n} \mu_n(I_\sigma) = M_n$$

Hence (1) follows. By Doob's theorem, since non-negative martingales are bounded in L^1 automatically, they converge almost surely. Especially,

(2) $$M_n \xrightarrow[n\to\infty]{} M \quad \text{a.s.},$$

where $M \in [0,\infty)$ a.s.

Another, useful way to invoke martingales is to note that μ_n has density function

$$d\mu_n(x) = \rho_n(x)dx,$$

where for any $x \neq$ dyadic rational, we have

$$\rho_n(x) = 2^n U_{a_1} U_{a_1 a_2} \cdots U_{a_1 \ldots a_n}$$

for $x \in I_{a_1 \ldots a_n}$. Hence for almost every $x \in [0,1]$

$\rho_n(x)$ is a martingale w.r.t F_n

This immediately proves that for any $\varphi \in C([0,1])$

$$\int_0^1 \varphi(x)\mu_n(dx) = \int_0^1 \varphi(x)\rho_n(x)dx$$

is a martingale, which is again bounded in L^1:

$$\mathbb{E}\left|\int_0^1 \varphi(x)\mu_n(dx)\right| \leq \mathbb{E}\int_0^1 |\varphi(x)|\mu_n(dx)$$

$$\leq \|\varphi\|_\infty \mathbb{E} M_n = \|\varphi\|_\infty .$$

We deduce that $\int_0^1 \varphi(x)\mu_n(dx)$ converges for a countable dense set of φ:s in the unit ball of $C([0,1])$, and by simple approximation for all $\varphi \in C([0,1])$. By Prohorov's theorem we deduce that

$$\mu_n \xrightarrow{w^*} \mu \quad a.s.,$$

where μ is a (random) nonnegative finite Borel measure on $[0,1]$. $\square$

[<u>Lemma 8.2</u> $\quad \mathbb{P}(M \neq 0) \in \{0,1\}$

Proof. This is a immediate consequence of Kolmogorov's 0-1 thm since $\widetilde{U}>0$.

The construction of μ was quite easy. However, it is not so easy to get hold of basic properties of μ. A basic question:

- when is μ nontrivial?

Namely, an L^1-martingale may easily converge to 0 almost surely.

Exercise Construct martingale $(X_n)_{n\geq 1}$ so that $X_k \geq 0$ $\forall k$ and $X_k \to 0$ almost surely.

To get natural conditions on which $\mu \neq 0$ a.s., let us recall that if M_n would be uniformly integrable, we would obtain $M_n = E(M \mid F_n)$, since $M = \lim_{n\to\infty} M_n$ a.s., so that then $M \neq 0$.

On the other hand $(M_n)_{n\geq 1}$ is uniformly integrable if M is an L^p-martingale for some $p>1$, i.e

$$E|M_n|^p \leq C < \infty \quad \text{for all } n \geq 1$$

Lemma 8.3. (The 'smoothing equation')

(3) $\quad M \sim M'U' + M''U''$,

where M', U', M'', U'' are all independent and $M' \sim M'' \sim M_{1,\mu}$, $U \sim U' \sim \widetilde{U}$.

Similarly, (4) $M_{n+1} \sim M_n' U' + M_n'' U''$, where $M_n' \sim M_n'' \sim M_n$, and again $\{M_n', M_n'', U', U''\}$ is an independent set.

Proof. Both (3) and (4) are direct consequences of construction of the cascade measure. □

Corollary 8.4. Either $\mu = 0$ a.s or a.s. the measure μ is non-trivial on any open subset of $[0,1]$.

Proof. Exercise. □

Def. We say that μ is non-degenerate if $\mu \neq 0$ a.s.

The following result characterizes the existence of p:th moment for $p>1$.

Theorem 8.5 (Kahane) Let $p>1$. The random cascade measure is nondegenerate and satisfies

(4) $$\mathbb{E} M^p < \infty$$

if and only if $\mathbb{E}\tilde{U}^p < \frac{1}{2}$

Proof. Assume that M is nondegenerate and $\mathbb{E} M^p < \infty$. Then $0 < \mathbb{E} M^p < \infty$.
By using Lemma 8.3 and the inequality $(u+v)^p \geq u^p + v^p$ for $u, v \geq 0$, with equality iff $u=0$ or $v=0$, we obtain by independence

$$\mathbb{E} M^p = \mathbb{E}(\tilde{U}' M' + \tilde{U}'' M'')^p$$
$$\geq \mathbb{E}(\tilde{U}' M')^p + \mathbb{E}(\tilde{U}'' M'')^p = 2\mathbb{E} M^p \mathbb{E}\tilde{U}^p.$$

By positivity, there must be strict inequality above, whence $\mathbb{E}\tilde{U}^p < 1/2$.

Assume then that $\mathbb{E}\tilde{U}^p < 1$. For simplicity, we will only treat the case $p \in (1,2]$

APPENDIX

A. The Bochner integral

Standing assumption: E a Banach space, (Ω, F, μ) a probability space

Rem. All results and proofs (with trivial modifications) in this section work if μ is σ-finite.

Def. • $X : \Omega \to E$ is a simple function if
$$X(\omega) = \sum_{k=1}^{n} a_k \chi_{A_k} \quad ,$$
where $n \geq 1$, $a_k \in E$, $A_k \in F$ for all $k \leq n$.

• $X : \Omega \to E$ is strongly measurable if
$$X(\omega) = \lim_{n\to\infty} X_n(\omega) \quad \forall \omega \in \Omega \, ,$$
where X_n:s are E-valued simple functions

• $X : \Omega \to E$ is weakly measurable if $e'(X(\omega))$ is a measurable, scalar valued function for each $e' \in E$.

Lemma A1 X is strongly measurable if and only if ($X \to E$ is measurable in the usual sense (i.e. $X^{-1}(B) \in F$ for every open $B \Leftrightarrow X^{-1}(B) \in F$ for every Borel set $B \in \mathcal{B}(E)$) and $X(\Omega)$ is separable subset of E)

Proof. Assume first that X is strongly measurable. Then $X(\Omega) \in \overline{\bigcup_{k\geq 1} X_k(\Omega)}$, which is a separable subset of E. If $V \subset E$ is open, denote

$$V_k := \{x \in V,\ d(x, V^c) > \tfrac{1}{k}\} \quad , k = 1, 2 \ldots$$

Then V_k:s are open, $V = \bigcup_{k=1}^{\infty} V_k$. By assumption,

$$X^{-1}(V) = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcap_{m=j}^{\infty} X_m^{-1}(V_k) \qquad \text{(why?)},$$

which shows that $X^{-1}(V) \in F$ as X_m:s are clearly measurable.

Assume then that $X(\Omega)$ is separable and $X : (\Omega, F) \to (E, \mathcal{B}(E))$ is measurable. Let $X(\Omega) \subset \overline{\{a_n\}_{n\geq 1}}$ ($a_n \in E$, $n \geq 1$). Denote

$$\begin{cases} U_k^n := \{y \in E : \|y - a_k\| \leq \min(\|y - a_j\|, 1 \leq j \leq n)\} \\ B_k^n := X^{-1}(U_n^k), \\ D_n^n := B_n^n \setminus \left(\bigcup_{j=1}^{k-1} B_j^n\right), \quad D_1^n := B_1^n \end{cases}$$

and set

$$X_n(\omega) = \sum_{k=1}^{n} a_k \chi_{D_k^n}(\omega).$$

(a "best" approximation of X using $a_1, \ldots, a_n$).

Clearly $X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)$ for every $\omega \in \Omega$. □

Theorem A 2. (Pettis measurability thm)

$X : \Omega \to E$ is strongly measurable if and only if it is weakly measurable and $X(\Omega)$ is separable.

Proof. If X is strongly measurable, we have $X = \lim_{n\to\infty} X_n$ where X_n:s are simple. Simple functions are separably valued and weakly measurable, and these properties carry to X as a pointwise limit.

Assume then that X is weakly measurable and separably valued. We may assume that E is separable. Recall, that then we may pick $\{e_k'\}_{k=1}^{\infty}$ with $e_k' \in E'$, $\|e_k'\| = 1$ $\forall\, k \geq 1$ and such that

$$\|e\| = \sup_{k\geq 1} |e'_k(e)| \qquad \text{for all } e \in E.$$

Then $\quad X^{-1}(\overline{B(0,r)}) = \bigcap_{k=1}^{\infty} \{|e'_k(X(\omega))| \leq r\} \in F$

for any $r>0$. Especially $X^{-1}(B) \in F$ for any open ball $B \subset E$ with center at 0. This carries to any open ball B by considering $X(\omega)+a$ $(a \in E)$, and hence to any open set B since E is separable. Thus X is measurable in the usual sense, and the rest follows from Lemma A.1. $\square$

<u>Definition</u> A strongly measurable function $X: \Omega \to E$ is <u>Bochner-integrable</u> if there exist simple functions $X_n(\omega) \to X(\omega)$ for every $\omega \in \Omega$ and

$$\int_\Omega \|X_n(\omega) - X_m(\omega)\| \, d\mu(\omega) \xrightarrow[n,m\to\infty]{} 0. \tag{A.1}$$

Then we set $\quad \int_\Omega X(\omega) d\mu(\omega) := \lim_{n\to\infty} \int_\Omega X_n(\omega) d\mu(\omega)$,

where for the simple function $X_n(\omega) := \sum_{k=1}^{m(n)} a_{n,k} \chi_{A_{n,k}}$ one defines

$$\int_\Omega X_n(\omega) d\mu(\omega) := \sum_{k=1}^{m(n)} a_{n,k} \mu(A_{n,k}).$$

<u>Theorem A.3.</u> (i) The Bochner integral is well-defined.

(ii) A strongly measurable $X: \Omega \to E$ is Bochner-integrable if and only if

$$\int_\Omega \|X(\omega)\| \mu(d\omega) < \infty$$

(iii) If X is Bochner integrable, then

$$e'\left(\int_\Omega X(\omega) d\mu(\omega)\right) = \int_\Omega e'(X(\omega)) d\mu(\omega) \qquad \forall\, e' \in E'$$

$$\left\| \int_\Omega X(\omega)\, d\mu(\omega) \right\|_E \leq \int_\Omega \|X(\omega)\|\, d\mu(\omega).$$

<u>Proof</u> (i) The integral of a simple function is well-defined since if $\sum_{k=1}^{m_1} a_k \chi_{A_k}(\omega) = \sum_{j=1}^{m_2} b_k \chi_{\tilde{A}_j}(\omega) \quad \forall\, \omega \in \Omega$

$\begin{cases} A_k, \tilde{A}_j \in F \\ a_k, b_j \in E \end{cases}$, then $\sum_{k=1}^{m_1} e'(a_k) \chi_{A_k}(\omega) = \sum_{j=1}^{m_2} e'(b_k) \chi_{\tilde{A}_j}(\omega)$

For all e'. Standard integration yields

$$\sum_{k=1}^{m_1} e'(a_k)\, \mu(A_k) = \sum_{j=1}^{m_2} e'(b_j)\, \mu(\tilde{A}_j)$$

or
$$e'\Big(\sum_{k=1}^{m_1} a_k \mu(A_k)\Big) = e'\Big(\sum_{j=1}^{m_2} b_j \mu(\tilde{A}_j)\Big).$$

Since this holds for all e', we deduce that

$$\sum_{k=1}^{m_1} a_k \mu(A_k) = \sum_{j=1}^{m_2} b_j \mu(\tilde{A}_j).$$

If $Y(\omega) = \sum_{k=1}^{m} a_k \chi_{A_k}$ is simple (may assume that A_k's are disjoint) we obtain

(A2) $$\left\| \int_\Omega Y(\omega)\mu(d\omega) \right\| = \left\| \sum_{k=1}^{m} \mu(A_k) a_k \right\| \leq \sum_{k=1}^{m} \mu(A_k) \|a_k\| = \int_\Omega \|Y(\omega)\| \mu(d\omega).$$

Hence, if $\int_\Omega \|X_n(\omega) - X_m(\omega)\|\, d\mu(\omega) \xrightarrow[n,m\to\infty]{} 0$

for simple functions $X_n : \Omega \to E$, we have that

$$\left\| \int_\Omega X_n(\omega)\, d\mu(\omega) - \int X_m(\omega)\, d\mu(\omega) \right\|$$
$$\leq \int_\Omega \|X_n(\omega) - X_m(\omega)\|\, d\mu(\omega) \xrightarrow[m,n\to\infty]{} 0.$$

Thus, $\int_\Omega X(\omega)\,d\mu(\omega) := \lim_{n\to\infty} \int_\Omega X_n(\omega)\,d\mu(\omega)$ exists,

and it is well-defined as soon as we show that it does not depend on the choice of the sequence (X_n). Assume thus that

$$X(\omega) = \lim_{n\to\infty} X_n(\omega) = \lim_{n\to\infty} Y_n(\omega) \quad \forall\, \omega\in\Omega$$

where both sequences (X_n) and (Y_n) satisfy (A1). It follows that (we use (A2))

$$\left\| \int_\Omega X_n\,d\mu - \int_\Omega Y_m\,d\mu \right\| \le \int_\Omega \|X_m - X\|\,d\mu + \int_\Omega \|Y_m - X\|\,d\mu$$

$$\le \liminf_{k\to\infty} \int_\Omega \|X_n - X_k\|\,d\mu + \liminf_{k\to\infty} \int_\Omega \|Y_m - Y_k\|\,d\mu$$

$$\le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

as soon as $n \ge n_0(\varepsilon)$, $m \ge n_0(\varepsilon)$ (by Fatou's lemma).
This obviously yields that (X_n) and (Y_n) yield the same value for $\int_\Omega X\,d\mu$.

(ii) Let us assume that X is Bochner-integrable and (X_n) is a sequence defining $\int X\,d\mu$ (X_n's simple). By definition $\sup_{n\ge 1} \int_\Omega \|X_n\|\,d\mu \le \int_\Omega \|X_1\|\,d\mu + \sup_{n\ge 1} \int_\Omega \|X_n - X_1\|\,d\mu < \infty$. Hence by Fatou's lemma

$$\int_\Omega \|X\|\,d\mu = \int_\Omega \liminf_{n\to\infty} \|X_n\|\,d\mu \le \liminf_{n\to\infty} \int_\Omega \|X_n\|\,d\mu < \infty.$$

Conversely, assume that X is strongly measurable and

$$\text{(A.3)} \qquad \int_\Omega \|X\|\,d\mu < \infty.$$

We may pick simple functions $X_n(\omega)$ such that $X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)$ for every $\omega\in\Omega$. Define a new sequence $\tilde{X}_n(\omega)$ of simple functions by setting

$$\tilde{X}_n(\omega) = \begin{cases} X_n(\omega) & \text{if } \|X_n(\omega)\| \le 2\|X(\omega)\| \\ 0 & \text{otherwise.} \end{cases}$$

Then $\tilde{X}_n(\omega) \underset{n\to\infty}{\to} X(\omega)$ at every ω and

$$\|\tilde{X}_n(\omega) - \tilde{X}_m(\omega)\| \le 4\|X(\omega)\| \quad \forall\, \omega \in \Omega.$$

We may thus use dominated convergence to obtain

$$\lim_{n,m\to\infty} \int_\Omega \|\tilde{X}_n(\omega) - \tilde{X}_m(\omega)\|\, d\mu(\omega) = 0,$$

and hence X is Bochner integrable.

(iii) Exercise. □

Theorem A.4. (Dominated convergence thm for the Bochner integral) Assume that X_n, X ($n \ge 1$) are strongly measurable with

$$\|X(\omega)\|, \|X_n(\omega)\| \le h(\omega) \quad \forall n,$$

where $h : \Omega \to \mathbb{R}$ is measurable and $\int_\Omega |h(\omega)|\, \mu(d\omega) < \infty$.
Then, if $X_n(\omega) \underset{n\to\infty}{\to} X(\omega)$

for a.e $\omega \in \Omega$, one has $\int_\Omega X(\omega)\, d\mu(\omega) = \lim_{n\to\infty} \int_\Omega X_n(\omega)\, d\mu(\omega)$

Proof. By Thm A.3 (ii), X and X_n are integrable and the statement follows by applying the scalar-valued dominated convergence thm on the inequality

$$\left\| \int_\Omega X_n\, d\mu - \int_\Omega X\, d\mu \right\| \le \int_\Omega \|X - X_n\|\, d\mu. \quad \square$$

Remark
- In many statements it is enough to assume the statements hold only a.s. (just like in the standard integration theory).

- Often one writes $X \in L^1_X(d\mu)$ if X is Bochner integrable. More generally,

$F \in L^p_X(d\mu)$ if F is strongly measurable and

$$\int_\Omega \|F\|^p \mu < \infty .$$