17.1

1. Some probability Facts

$X : \Omega \to \mathbb{R}$ is a real random variable, if X is measurable.

Here $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space.

If E is a topological space, then

$X : \Omega \to E$ is a E-valued random variable

if $X^{-1}(G) \in \mathcal{F}$ for every open set $G \subset E$.
Then X defines a measure μ_X on E by

$$\mu_X(B) = \mathbb{P}(X^{-1}(B)) \quad \text{for } B \in \mathcal{B}(E),$$

Borel σ-algebra.

Clearly μ_X is a prob. measure, the 'law' of X.

We say that random variables

$X_1, \dots, X_n$ (with values in E) are

independent if the σ-algebras

$$\sigma(X_1), \dots, \sigma(X_n)$$

are independent. Here

$$\sigma(X) = \{ X^{-1}(B) \mid B \in \mathcal{B}(E) \},$$

and independence of $\sigma(X_1), \dots, \sigma(X_n)$ mean that

$$\mathbb{P}(B_1 \cap B_2 \cap \dots \cap B_n) = \prod_{k=1}^{n} \mathbb{P}(B_k) \quad \text{if } B_k \in \sigma(X_k),\ k=1,\dots,n.$$

Recall that (if X is real or $\mathbb{C}$-valued)

$$\mathbb{E}X = \int_\Omega X(\omega)\,\mathbb{P}(d\omega). \qquad (\text{ass. } \mathbb{E}|X|<\infty \iff X \text{ integrable})$$

Then $V(X) = \mathbb{E}(|X-\mathbb{E}X|^2) = \mathbb{E}X^2 - (\mathbb{E}X)^2$.
Remember that

$$\mathbb{E}X_1X_2 = (\mathbb{E}X_1)(\mathbb{E}X_2)$$

if X_1 and X_2 are independent.

If $\mu_X = \mu_{X'}$, we say that

X and X' are identically distributed (or similar)
This is denoted by $X \sim X'$.
If $X_1, X_2, \dots$ are independent and identically distributed, we say $X_1, X_2, \dots$ are i.i.d.

Given probability spaces $(\Omega_k, \mathcal{A}_k, \mathbb{P}_k)$, one may construct a unique product measure

$$\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2 \times \dots = \prod_{k=1}^{\infty} \mathbb{P}_k,$$

which is defined in the σ-algebra

$$\prod_{k=1}^{\infty} \mathcal{A}_k = \sigma\{A_1 \times \dots \times A_\ell \times \Omega_{\ell+1} \times \Omega_{\ell+2} \times \dots \mid A_i \in \mathcal{A}_i \ \forall i \le \ell,\ \ell \ge 1\}$$

[Recall that if $\mathcal{C} \subset P(\Omega)$, $\sigma(\mathcal{C})$ is the smallest σ-algebra containing all sets $\mathcal{C}$]
and satisfies

$$\mathbb{P}(A_1 \times \dots \times A_\ell \times \Omega_{\ell+1} \times \Omega_{\ell+2} \times \dots) = \mathbb{P}_1(A_1)\dots\mathbb{P}_\ell(A_\ell) \quad \text{for all such } A_i\text{'s.}$$

An easy way to generate i.i.d sequence $X_1, X_2, \dots$ given $X_0 : \Omega_0 \to \mathbb{R}$ is to let for all $k \geq 1$ $(\Omega_k, \mathcal{F}_k, P_k) = (\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$, and set

$$(\Omega, \mathcal{F}, \mathbb{P}) = \prod_{k=1}^{\infty} (\Omega_k, \mathcal{F}_k, \mathbb{P}_k),$$

$$X_k(\omega) = X_0(\omega_k) \qquad , k \geq 1.$$

A Rademacher sequence $(\varepsilon_k)_{k=1}^{\infty}$, is an i.i.d sequence such that

$$\mathbb{P}(\varepsilon_k = 1) = \tfrac{1}{2} \quad , \quad \mathbb{P}(\varepsilon_k = -1) = \tfrac{1}{2}.$$

Let $A_1, \dots, A_n, \dots$ be events. Then the event

$$\overline{\lim_{n\to\infty}} A_n = \bigcap_{k\geq 1} \bigcup_{n\geq k} A_k$$

takes place if infinitely many of the A_n's hold.

Lemma 1.1. (Borel-Cantelli)

(i) If $\sum_{k=1}^{\infty} \mathbb{P}(A_k) < \infty$, then $\mathbb{P}(\overline{\lim_{n\to\infty}} A_n) = 0$

(2) If $\sum_{k=1}^{\infty} \mathbb{P}(A_k) = \infty$ and the A_k's are independent, then $P(\overline{\lim_{n\to\infty}} A_n) = 1$.

Proof. (i) Now $\int (\sum_{k=1}^{\infty} \chi_{A_k}) \mathbb{P}(d\omega) = \sum_{k=1}^{\infty} \mathbb{P}(A_k) < \infty$, whence $\sum_{k=1}^{\infty} \chi_{A_k} < \infty$ a.s.

(ii) Fixing $N \geq 1$ we have $\mathbb{E} \prod_{N=1}^{n} (1 - \chi_{A_N})$ $(n > N)$

$$= \prod_{N=1}^{n} \mathbb{E}(1 - \chi_{A_N}) = \prod_{N=1}^{n} (1 - \mathbb{P}(A_k)).$$

Recall that $1-x < e^{-x}$ so that

$$\lim_{n\to\infty} \prod_{k=N}^{n} (1-\mathbb{P}(A_k)) \leq \lim_{n\to\infty} \exp\left(-\sum_{k=N}^{n} \mathbb{P}(A_k)\right) = 0$$

By dominated convergence we obtain
$\mathbb{E}\left(\prod_{k=N}^{\infty}(1-\chi_{A_k})\right) = 0$, whence $\prod_{k=N}^{\infty}(1-\chi_{A_k}) = 0$ a.s,
or $\mathbb{P}\left(\bigcup_{k=N}^{\infty} A_k\right) =$. As this holds for any $N \geq 1$, the claim follows. $\square$

Let $\mathcal{A}_1, \mathcal{A}_2, \dots$ be σ-algebras. Recall that $\sigma(\mathcal{A}_1, \mathcal{A}_2, \dots)$ is the σ-algebra generated by elements in the $\mathcal{A}_k$:s. The <u>tail σ-algebra</u> of the $\mathcal{A}_k$:s is

$$\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, X_{n+2}, \dots).$$

<u>Theorem 1.2.</u> (Kolmogorov's 0-1 law)
Assume that $\mathcal{A}_1, \mathcal{A}_2, \dots$ are independent σ-algebras and $\mathcal{T}$ is the corresponding tail σ-algebra. Then,

$$\mathbb{P}(A) = 0 \text{ or } \mathbb{P}(A) = 1 \quad \text{for all } A \in \mathcal{T}.$$

<u>Proof.</u> For any $N \geq 1$, the σ-algebras

$$\mathcal{B}_N := \sigma(\mathcal{A}_1, \dots, \mathcal{A}_N) \quad \text{and} \quad \mathcal{T} \text{ are independent.}$$

Put $\mathcal{B} = \sigma(\mathcal{A}_1, \dots \mathcal{A}_N, \dots) = \sigma(\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots)$.

Then $\bigcup \mathcal{B}_n$ is a π-system (see below), whose all members are independent of $\mathcal{T}$. By Lemma 1.4 below it follows that

$\mathcal{B}$ is independent of $\mathcal{T}$. But since $\mathcal{T} \subset \mathcal{B}$ we obtain that $\mathcal{T}$ is independent of itself! Especially,

$$\mathbb{P}(A) = \mathbb{P}(A \cap A) = (\mathbb{P}(A))^2$$

$$\Rightarrow \quad \mathbb{P}(A) \in \{0,1\}. \quad \square$$

A subset $\mathcal{I} \subset \mathcal{F}$ is a π-system if

$$A, B \in \mathcal{I} \Rightarrow A \cap B \in \mathcal{I}.$$

Lemma 1.3. Let μ_1 and μ_2 be prob. measures on $(\Omega, \mathcal{F})$ and assume that $\mu_1(A) = \mu_2(A)$ for every $A \in \mathcal{T}$, where $\mathcal{T}$ is a π-system with $\sigma(\mathcal{T}) = \mathcal{F}$. Then $\mu_1 = \mu_2$.

Proof. Most basic textbooks on probability. $\square$

Lemma 1.4. Assume that $\mathcal{A}$ and $\mathcal{B}$ are subsigma-algebras so that $\mathcal{A} = \sigma(\mathcal{I})$, $\mathcal{B} = \sigma(\mathcal{J})$, where $\mathcal{I}$ and $\mathcal{J}$ are π-systems. Then, if $\mathcal{I}$ and $\mathcal{J}$ are independent, then also $\mathcal{A}$ and $\mathcal{B}$ are independent.

Proof. Fix $A \in \mathcal{I}$. The measures

$$B \mapsto \mathbb{P}(A \cap B) \quad \text{and} \quad B \to \mathbb{P}(A)\mathbb{P}(B)$$

have the same total mass and they agree on $\mathcal{J}$.

Hence they agree on $\mathcal{B}$ (by Lemma 1.3). Finally,
Fix $B \in \sigma(\mathcal{J}) = \mathcal{B}$ and do the same trick with the measures

$$A \mapsto \mathbb{P}(A \cap B) \quad \text{and} \quad A \mapsto \mathbb{P}(A)\mathbb{P}(B). \qquad \square$$

We finally study when the sum of positive random variables converges.

Theorem 1.5 (Beppo-Levi) Assume $X_n \in L^1(\Omega)$, $X_n \geq 0$ and $\sum_{n=1}^{\infty} \mathbb{E}X_n < \infty$. Then $\sum_{n=1}^{\infty} X_n < \infty$ a.s.

Proof. Simply compute (by monotone convergence)

$$\mathbb{E}\left(\sum_{n=1}^{\infty} X_n\right) = \sum_{n=1}^{\infty} \mathbb{E}X_n < \infty.$$

There is a converse (a full characterization!) assuming that the X_n are independent.

Theorem 1.6. Assume that the random variable $X_n \geq 0$ $\forall n$, and they are independent. Denote $X_n' = \min(1, X_n)$.

(i) If $\sum \mathbb{E}X_n' < \infty$, then a.s. $\sum_{n=1}^{\infty} X_n < \infty$

(ii) if $\sum \mathbb{E}X_n' = \infty$, then a.s. $\sum_{n=1}^{\infty} X_n = \infty$.

Proof. (i) By thm 1.5 $\sum_{n=1}^{\infty} X_n' < \infty$ a.s.

Moreover, $\mathbb{P}(X_n' \neq X_n) = \mathbb{P}(X_n > 1) \leq \mathbb{P}(X_n' = 1)$
$\leq \dfrac{\mathbb{E}X_n'}{1} = \mathbb{E}X_n'$. Hence $\sum_{n=1}^{\infty} \mathbb{P}(X_n' \neq X_n) < \infty$,
and Borel-Cantelli yields that a.s. $X_n = X_n'$ for large values of n, which suffices.

(ii) Observe that $\sum_{n=1}^{\infty} X_n'$ and $\sum_{n=1}^{\infty} X_n$ converge simultaneously. Thus we may assume that $0 \le X_n \le 1$ and the series $\sum_{n=1}^{\infty} \mathbb{E} X_n$ diverges. For $0 \le y \le 1$ one has

$$e^{-y} \le 1 - \tfrac{1}{2} y .$$

Hence $\mathbb{E}\, e^{-(X_1 + \dots + X_N)} = \prod_{k=1}^{N} \mathbb{E}(e^{-X_k})$

$$\le \prod_{k=1}^{N} \mathbb{E}(1 - \tfrac{1}{2} X_k) = \prod_{k=1}^{N} (1 - \tfrac{1}{2} \mathbb{E} X_k) \xrightarrow[N \to \infty]{} 0$$

Thus (by dominated convergence) $\mathbb{E}(\exp(-\sum_{k=1}^{\infty} X_k)) = 0$.
This yields $\sum_{k=1}^{\infty} X_k = \infty$ a.s. □

<u>Corollary 1.7.</u> Let $X_1, X_2, \dots$ be non-negative, i.i.d and integrable. Then, if $a_k \ge 0 \ \forall k$

$$\sum_{k=1}^{\infty} a_k X_k = \begin{cases} < \infty & \text{a.s if } \sum_{k=1}^{\infty} a_k < \infty \\ = \infty & \text{a.s. if } \sum_{k=1}^{\infty} a_k = \infty . \end{cases}$$

(assume $X_1 \ne 0$ a.s.)

<u>Proof.</u> If $\sum a_k < \infty$, then $\sum a_k \mathbb{E} X_k = \sum a_k \mathbb{E} X_1 < \infty$. If $\sum a_k = \infty$, then

$$\sum a_k \underbrace{\mathbb{E} \min(1, X_k)}_{>0 \text{ (does not depend on } k)} = \infty \Rightarrow \sum a_k \min(1, X_k) = \infty \text{ a.s.}$$

$$\Rightarrow \sum a_k X_k = \infty \text{ a.s. } \square$$

<u>Remark</u> Simple examples (exercise) show that Thm 1.6 (ii) is not true if $\mathbb{E} X_k'$ is replaced by $\mathbb{E} X_k$.

18.1

2. Random variables in Banach spaces

In these lectures we often consider random variables with values in a Banach space E. Also distribution valued ones appear, but they will be reduced to this case (actually to the case when is Hilbert).

In this section E is a (real) Banach space and $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. We equip E with $\mathcal{B}(E)$, i.e the Borel σ-algebra generated by all open subsets of E.

Def. A probability measure on E is tight if for any $\varepsilon > 0$ there is $K \subset E$, K compact with
$$\mu(E \setminus K) < \varepsilon.$$

Remark. μ on E is tight $\Leftrightarrow$ μ is Radon (regular) measure, i.e for any Borel set $A \subset E$ there are compact K_ε and open O_ε with $K_\varepsilon \subset A \subset O_\varepsilon$ and $\mu(O_\varepsilon \setminus K_\varepsilon) < \varepsilon$, for all $\varepsilon > 0$.

Theorem 2.1. μ on E is tight if and only if there is a separable closed subspace $M \subset E$ such that $\mu(E) = 1$.

Proof. If μ is tight, pick compact sets $K_j \subset E$ with $\mu(K_j) \geq 1 - 2^{-j}$. Take
$$M = \overline{\operatorname{span}\Big(\bigcup_{j=1}^{\infty} K_j\Big)},$$
which is clearly separable.

In order to prove the other direction, take a dense set $\{x_i\}_{i=1}^{\infty}$ from M. Let $\varepsilon>0$. Then, if $n\geq 1$ we have $M=\bigcup_{k=1}^{\infty}\overline{B}(x_k,2^{-n})$, whence we may pick $N(n)$ so that

closed ball

$$\mu\Big(\bigcup_{k=1}^{N(n)}\overline{B}(x_k,2^{-n})\Big)>1-\varepsilon 2^{-n}$$

Then $K_\varepsilon := \bigcap_{n=1}^{\infty}\bigcup_{k=1}^{N(n)}\overline{B}(x_k,2^{-n})$ is compact (why?) and $\mu(K_\varepsilon)\geq 1-\varepsilon$. $\square$

Def. A random variable $X:\Omega\to E$ is separable ($\Leftarrow$ radon $\Leftarrow$ tight) if μ_X on E is tight, i.e. there is a closed subspace $M\subset E$ such that $X(\omega)\in M$ a.s.

Obs. We used Thm 2.1. in the definition. Note also, that for separable random variable X one has (exercise)

$$\mathbb{P}(X\in A)=\sup\{\mathbb{P}(X\in K): K\subset A, K \text{ compact}\}$$

for all $A\subset\mathcal{B}(E)$.

Remark We shall mainly consider separable spaces E (or at least separable random variables). Namely, in the nonseparable case it is not even known if a sum of two random variables is measurable ?!

Def. Let E' be the dual of E. The cylindrical σ-algebra of E is the one generated by elements of E', i.e. by the sets

$$\{x: x'(x)<\alpha \mid x'\in E'\}.$$

Def. A cylindrical subset of E is of the form
$$\{x \in E : (y_1'(y), \dots, y_n'(y)) \in \Gamma\},$$
where $\Gamma \subset \mathcal{B}(\mathbb{R}^n)$, $y_1, \dots, y_n \in \mathbb{R}^n$, $n \geq 1$.

Thm 2.2. Let E be separable. Then cylindrical σ-algebra coincides with $\mathcal{B}(E)$.

Proof. Since E is separable, one may pick elements $x_k' \in E'$ with $\|x_k'\| = 1$ and
$$\|x\| = \sup_{k \geq 1} |x_k'(x)| \quad \forall\, x \in E.$$
Hence any closed ball $\bar{B}(a, r)$ $(a \in E, r > 0)$ can be written as
$$\bar{B}(a,r) = \bigcap_{k=1}^{\infty} \{|x_k'(x-a)| \leq 1\}.$$
Thus, by separability, any open set is in the cylindrical σ-algebra, whence it coincides with $\mathcal{B}(E)$. $\square$

From now on we will consider only, namely separable Banach spaces E.

Remark. It would be natural to consider also nonseparable spaces, like ℓ^∞ since one would then look at realizations of random processes as single vectors. E.g., one would like to consider the Rademacher sequence $(\varepsilon_k)_{k=1}^{\infty}$ as an ℓ^∞-valued random variable. The following example demonstrates the difficulties one can have.

Example Write each number $\omega \in [0,1] \setminus \mathbb{Q} = \Omega$ in binary expansion

$$\omega = \sum_{k=1}^{\infty} \omega_k 2^{-k} ,$$

and set $X_k(\omega) = 2\omega_k - 1$. Then (with $F = \mathcal{B}(\Omega)$, $\mathbb{P}$ = the Lebesgue measure)

$$(X_k)_{k=1}^{\infty} \sim (\varepsilon_k)_{k=1}^{\infty} ,$$

thus we have a natural realization of the Rademacher variables. Let

$$X : \Omega \to \ell^{\infty} ,$$

$$X(\omega) = (X_1(\omega), X_2(\omega), \dots) .$$

However, X is not measurable

Take a non-Lebesgue-measurable set $F \subset \Omega$, denote $X(F) = H \subset \ell^{\infty}$. Now elements of H are separated by distance ≥ 2, so

$$F = X^{-1}\Big(\underbrace{\bigcup_{f \in F} \overset{\uparrow \text{ open ball in } \ell^{\infty}}{B(f, 1/2)}}_{\text{open set}} \Big) . \qquad \square$$

23.1

Def Let μ be a prob. measure on E. The Fourier transform of μ is the function

$$x' \longrightarrow \int_E e^{i x'(x)} \mu(dx) := \hat{\mu}(x')$$

Similarly, the characteristic function of an E-valued random variable X is

$$\varphi_X(x') = \mathbb{E} \exp(i x'(X(\omega))) = \hat{\mu}_X(x').$$

Theorem 2.3. Let E be separable. Then $\hat{\mu}$ determines μ uniquely.

Proof. By the n-dimensional result for any $x_1', \dots x_n' \in E'$ $\hat{\mu}$ determines the distribution μ_g of the vector

$$(x_1'(x), \dots, x_n'(x)) =: g(x)$$

(i.e. the push forward of μ onto $\mathbb{R}^n$ by g)

since
$$\hat{\mu_g}(\lambda) = \hat{\mu}(\lambda_1 x_1' + \dots + \lambda_n x_n')$$
$$\Vert$$
$$\int_{\mathbb{R}^n} e^{\lambda \cdot y} \mu_g(dy)$$

Hence $\hat{\mu}$ determines μ on cylindrical sets of the form

$$\{x_1'(x) < a_1, \dots, x_n'(x) < a_n\}.$$

These have property π and generate the cylindrical sets. The claim now follows from Lemma 1.3 and Thm 2.2. □

Especially, φ_X determines the law of X. One can also consider weak convergence of measures and characterize it in terms of $\hat{\mu}_X$:s, but we will do it later if needed. Same with Bochner integral.

3. Random series in a Banach space

In this section we consider series

$$\sum_1^\infty X_n ,$$

where $(X_n)_{n\geq 1}$ consists of independent random variables with values in a Banach space E. Especially, we consider the case

$$\sum_{n=1}^\infty \varepsilon_n u_n \quad \left(\sim \sum_{n=1}^\infty \pm u_n \right),$$

where ε_n is a Rademacher sequence, or ε_n:s are independent standard Gaussians.

Assume that (X_n) are independent with values in E. Write

$$Y_m(\omega) = \sum_{n=1}^m X_n(\omega),$$

and $Y(\omega) = \sum_{n=1}^\infty X_n(\omega)$ if this series converges.

Lemma 3.1. (Levy) Assume that $\sum_{n=1}^\infty X_n$ (as above) converges almost surely and the independent variables X_n are symmetric.*)

Denote $M(\omega) = \sup_m \|Y_m(\omega)\|$. Then

$$\mathbb{P}(M(\omega) > r) \leq 2\,\mathbb{P}(\|Y(\omega)\| > r)$$

Proof. We may assume that $Y(\omega)$ is defined everywhere. Denote

---

*) i.e. $\boxed{X_n \sim -X_n}$

$$A = \{ M(\omega) > r \} \quad , \quad B = \{ \|Y(\omega)\| > r \}.$$

Write $\quad A = \bigcup_{k=1}^{\infty} A_k \quad$ (disjoint union),

where $\quad A_1 = \{ \|Y_1\| > r \}$, and for $k \geq 2$

$$A_k = \{ \|Y_1\|, \dots \|Y_{k-1}\| \leq r \; , \|Y_k\| > r \}.$$

If $\omega \in A_k$ then $\max(\|Y(\omega)\|, \|Y'(\omega)\|) > r$,

where $\quad Y'(\omega) = Y_k(\omega) - (X_{k+1}(\omega) + X_{k+2}(\omega) + \dots)$.

By the symmetry of X_n's we have

$$Y' \sim Y ,$$

whence $\quad \mathbb{P}(A_k \cap \{\|Y'(\omega)\| > r\}) = \mathbb{P}(A_k \cap \{\|Y(\omega)\| > r\})$.
Especially, it follows that

$$\mathbb{P}(B \cap A_k) \geq \frac{1}{2} \mathbb{P}(A_k).$$

Summing over k we obtain $\mathbb{P}(B) \geq \frac{1}{2} \mathbb{P}(A)$. □

We will finally prove a fairly keen tail behaviour for the sum

$$S(\omega) = \sum_{k=1}^{\infty} \varepsilon_k u_k \quad , \quad u_k \in E$$

just assuming the almost everywhere convergence.

First some results of independent interest. Recall important observation from Thm 1.2.

<u>Lemma 3.2.</u> $\sum_{k=1}^{\infty} X_k$ either converges a.s. or diverges a.s.

Lemma 3.2 (Kahane) Assume that $\overbrace{\sum_{n=1}^{\infty} \varepsilon_n u_n}^{=Y(\omega)}$ $(u_n \in E)$ Rademachers

converges a.s. and satisfies

$$\mathbb{P}(\|Y(\omega)\| > r) < \alpha \quad \text{for some } \alpha, r > 0.$$

Then $\mathbb{P}(\|Y(\omega)\| > 2r) < (2\alpha)^2$.

<u>Proof.</u> Denote $M(\omega) = \sup \|Y_m(\omega)\|$, and set

$$A = \{M > r\}$$
$$B = \{\|Y\| > r\} \qquad \Big| \; Y_m = \sum_{j=1}^{m} \varepsilon_j u_j$$
$$C = \{\|Y\| > 2r\}.$$

Let $A_m = \{\|Y_1\| \le r, \ldots, \|Y_{m-1}\| \le r, \|Y_m\| > r\}$

$$C_m = \{\|\sum_{m}^{\infty} \varepsilon_m u_n\| > r\}.$$

Observe: C_m depends only on

$(\varepsilon_m \varepsilon_{m+1}, \varepsilon_m \varepsilon_{m+2}, \ldots)$

$\sim (\varepsilon_{m+1}, \varepsilon_{m+2}, \ldots)$

regardless of the value of ε_m. Thus

$A_m \perp C_m$ (i.e. A_m and C_m are independent).

We obtain

$$\mathbb{P}(A_m \cap C_m) = \mathbb{P}(A_m)\mathbb{P}(C_m).$$

Since $A_m \cap C \subset C_m$, we have

$$P(A_m \cap C) \le P(A_m \cap C_m)$$

Adding these with $m=1,2\ldots$ we get

$$\mathbb{P}(C) \le \mathbb{P}(A) \sup_m \mathbb{P}(C_m)$$

Let $(\bar\varepsilon_1,\ldots,\bar\varepsilon_{m-1})$ be a ~~fixed~~ ± 1 sequence.
As in the proof of the Levy inequalities, we infer

$$\mathbb{P}\big(C_m \mid (\varepsilon_1,\ldots,\varepsilon_{m-1}) = (\bar\varepsilon_1,\ldots,\bar\varepsilon_{m-1})\big)$$
$$\le 2\,\mathbb{P}\big(B \mid (\varepsilon_1,\ldots,\varepsilon_{m-1}) = (\bar\varepsilon_1,\ldots,\bar\varepsilon_{m-1})\big)$$

Summing up over $(\bar\varepsilon_1,\ldots,\bar\varepsilon_{m-1})$ we obtain

$$\mathbb{P}(C_m) \le 2\,\mathbb{P}(B) \quad \text{so that}$$

$$(*) \quad \mathbb{P}(C) \le 2\,\mathbb{P}(A)\,\mathbb{P}(B) .$$

Observe finally that Levy yields that

$$\mathbb{P}(B) = \mathbb{P}(M > r) \le 2\,\mathbb{P}(S > r)$$
$$= 2\,\mathbb{P}(A) .$$

Substitution in $(*)$ yields the claim. $\square$

Let us give applications of our last results.

Theorem 3.3. Let $X_k : \Omega \to E$ be independent r.v.s, $k \ge 1$. If $\sum_{k=1}^{\infty} X_k$ converges in probability, then it converges almost surely.

<u>Proof.</u> Convergence in probability states that given $\varepsilon > 0$ there is $j = j(\varepsilon)$ such that

$$\mathbb{P}\left(\left\| \sum_{k=j}^{m} X_k \right\| > \varepsilon \right) < \varepsilon \quad \text{For all } m > j.$$

Pick an increasing sequence $j_1, j_2, j_3, \dots$ with

$$\mathbb{P}\left(\left\| \sum_{k=j_\ell}^{m} X_k \right\| > 2^{-\ell} \right) < 2^{-\ell-1}, \quad m \geq j_\ell$$

Lemma 3.1 applies and verifies that

(1) $$\mathbb{P}\left(\sup_{j_\ell \leq r} \left\| \sum_{k=j_\ell}^{r} X_k \right\| > 2^{-\ell} \right) \leq 2^{-\ell}$$

Hence, by Borel-Cantelli almost surely

$$\sup_{r \geq j_\ell} \left\| \sum_{k=j_\ell}^{r} X_k \right\| < 2^{-\ell}, \quad \ell \geq \ell_0(\omega),$$

which yield the convergence by the Cauchy-criterion. $\square$

Next an application on Rademacher series.

<u>Theorem 3.4</u> (Kahane). Let $u_k \in E \; \forall k \geq 1$. Then, if the series $Y(\omega) := \sum_{k=1}^{\infty} \varepsilon_k u_k$ converges almost surely, one has $\|Y\| \in L^p(\Omega) \quad \forall p \geq 1$.

Actually, $\mathbb{E}\, e^{\kappa \|Y\|} < \infty$ for small enough $\kappa > 0$.

<u>Proof</u> Denote $p(t) = \mathbb{P}(\|Y(\omega)\| > t)$, $t > 0$.

Choose $t_0 > 0$ with $p(t_0) \leq \frac{1}{8}$

Then Lemma 3.2. yields inductively

$$p(2t_0) \le 4(p(t_0))^2 \le \frac{1}{16}$$
$$p(4t_0) \le 4(p(2t_0))^2 \le 1/64 \, , \dots$$

In general, $\quad p(2^k t_0) \le 2^{-2^k - 2} \quad (*)$

so that $\quad \boxed{p(t) \le C e^{-\kappa t}}$ For some $\kappa, C > 0$.

We get

$$\mathbb{E} e^{\frac{c\|X\|}{2}} = 1 + \kappa \int_0^\infty e^{ct/2} p(t)\,dt < 1 + \kappa C \int_0^\infty e^{-ct/2}\,dt$$
$$< \infty.$$

This implies naturally that $\mathbb{E}\|X\|^p < \infty \quad \forall\, p \ge 1.$ $\square$

Theorem 3.5 (Khintchine-Kahane inequality)

Let $u_k \in E$ as before, and assume that $S(\omega) := \sum_{k=1}^\infty \varepsilon_k u_k$ converges a.s. Then, for any $0 < p < q < \infty$ one has

$$(\mathbb{E}\|S\|^p)^{1/p} \le C_{p,q} (\mathbb{E}\|S\|^q)^{1/q} ,$$

where $C_{p,q}$ depends only on p, q.

Proof. May assume $\mathbb{E}\|X\|^p = 1$. Let $t_0 = 8^{1/p}$ so that

$$\mathbb{P}(\|X\| \ge t_0) \le \frac{\mathbb{E}\|X\|^p}{t_0^p} = \frac{1}{8}.$$

Then $(*)$ above (on this page) yields

$$\mathbb{E}\|X\|^q \le q \int_0^\infty t^{q-1} p(t) \le q C \int_0^\infty t^{q-1} e^{-\kappa t}\,dt < \infty,$$

where the right hand side depends only on p, q. $\square$

Lemma 3.6. (Kahane's principle of contraction)
Let $Y(\omega) = \sum_{k=1}^{\infty} \varepsilon_k u_k \quad (u_k \in E)$ converge a.s.
Then for any bounded scalar sequence (λ_k) with $|\lambda_k| \le 1$ the series $Y'(\omega) := \sum_{k=1}^{\infty} \lambda_k \varepsilon_k u_k$
converges a.s and satisfies

$$(*) \qquad \mathbb{E}\|Y'(\omega)\|^p \le \mathbb{E}\|Y(\omega)\|^p \qquad \forall\, p \ge 1.$$

Proof. The map $(\lambda_1, \lambda_2, .., \lambda_\ell) \mapsto \left(\mathbb{E}\|\sum_{k=1}^{\ell} \lambda_k \varepsilon_k u_k\|^p\right)^{1/p}$
is convex ($\ell \ge 1$ arbitrary), takes the constant value $(\mathbb{E}\|\sum_{k=1}^{\ell} \varepsilon_k u_k\|^p)^{1/p}$ on $(\lambda_1, .., \lambda_\ell) \in \{-1,1\}^\ell$. This gives $(*)$ for this finite sum. Denote $M(\omega) := \sup_{n \ge 1} \|\sum_{k=1}^{n} \varepsilon_k u_k\|$, and $S(\omega) = \sum_{k=1}^{\infty} \varepsilon_k u_k$. By Lemmata 3.2. and 3.4 we deduce that $\|S(\omega)\| \in L^p$ and $M(\omega) \in L^p$. By dominated convergence we may pick subsequence $\ell_1, \ell_2, \dots$ so that

$$\mathbb{E}\|\sum_{k=\ell_r}^{j} \varepsilon_k u_k\|^p \le 2^{-r} \quad \forall\, r \ge 1,\ j \ge \ell_r.$$

$$\Rightarrow \quad \mathbb{E}\|\sum_{k=\ell_r}^{j} \lambda_k \varepsilon_k u_k\|^p \le 2^{-r} \qquad - \,''\, -$$

This implies that $\sum_{k=1}^{\infty} \lambda_k \varepsilon_k u_k$ converges in probability, hence a.s. by Th. 3.3. Finally $(*)$ for the full sequence follows by dominated convergence (details exercise.). $\square$

Lemma 3.7. (Pisier) Assume that $X_1, \dots, X_n, \dots$ are independent and symmetric real r.v.:s with
$$\inf_{k\geq 1} \mathbb{E}|X_k| \geq \kappa_0 > 0.$$
Then, if $u_1, u_2, \dots \in E$ are such that $(p>1)$ $\mathbb{E}\|\sum_{k=1}^{\infty} X_k u_k\|^p < \infty$, we have
$$\mathbb{E}\|\sum_{k=1}^{\infty}\varepsilon_k u_k\|^p \leq \kappa_0^{-p}\,\mathbb{E}\|\sum_{k=1}^{\infty} X_k u_k\|^p.$$

Proof. Enough to consider finite sums. To simplify thinking, set

$\Omega = \Omega_1 \times \Omega_1$, $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2$ and ε_k:s depend only on $\omega_1 \in \Omega_1$ and X_k:s on $\omega_2 \in \Omega_2$. Denote the corresponding expectations by $\mathbb{E}_1, \mathbb{E}_2$, $\mathbb{E} = \mathbb{E}_1\mathbb{E}_2$. By the contraction principle
$$\kappa_0^p\,\mathbb{E}_1\|\sum \varepsilon_n u_n\|^p \leq \mathbb{E}_1\|\sum \mathbb{E}_2|X_n|\,\varepsilon_n u_n\|^p$$
$$= \mathbb{E}_1\|\mathbb{E}_2(\sum \varepsilon_n |X_n| u_n)\|^p \leq \mathbb{E}_1(\mathbb{E}_2\|\sum \varepsilon_n |X_n| u_n\|)^p$$
$$\leq \mathbb{E}_1(\mathbb{E}_2\|\sum \varepsilon_n |X_n| u_n\|^p) \qquad \text{(Jensen)}$$
$$= \mathbb{E}\|\sum \varepsilon_n |X_n| u_n\|^p = \mathbb{E}\|\sum X_n u_n\|^p,$$
since the series $\sum \varepsilon_n |X_n| u_n$ is similar to $\sum X_n u_n$. $\square$

We are now ready for Kwapien's theorem.

Theorem 3.8. (Kwapien) If $\sum_{k=1}^{\infty} \varepsilon_k u_k =: S(\omega)$ is almost surely convergent. Then
$$\mathbb{E}\exp(\alpha^2\|S(\omega)\|^2) < \infty \quad \forall \alpha > 0.$$

$$S(\omega) := \sum_{k=1}^{\infty} \varepsilon_k u_k$$

Proof. Let us first recall that the proof of Thm 3.4 yields that (same notation)

$$\mathbb{E}\, e^{\delta \|S(\omega)\|} = \delta \int_0^\infty e^{\delta t} p(t)\,dt$$
$$\le c + \sum_{k=0}^{\infty} e^{2^{k+1} t_0 \delta}\, 2^{-2^k - 1} < \infty,$$

assuming that $t_0 \delta < 1/2$. However, by using this estimate for the remainder $\sum_{k=m}^{\infty} \varepsilon_k u_m$ with large enough m we obtain t_0 as small as we wish. Hence

$$\mathbb{E}\, e^{\lambda \|S(\omega)\|} < \infty \quad \forall \lambda > 0.$$

In order to still improve, consider for $j = 1, 2, \dots$ independent copies of the series $S(\omega)$:

$$S_j(\omega) = \sum_{k=1}^{\infty} \varepsilon_{kj} u_k$$

Set for $m = 2, 3, \dots$

$$\frac{1}{\sqrt{m}}(S_1 + \dots + S_m) = \sum_{k=1}^{\infty} \frac{\varepsilon_{k1} + \dots + \varepsilon_{km}}{\sqrt{m}}\, u_k$$
$$=: \sum_{k=1}^{\infty} X_{km} u_k.$$

By the central limit theorem, $\mathbb{E}|X_{km}| \ge c_0 > 0$ for all $k, m \ge 1$. By applying lemma 3.7 we get

$$\mathbb{E}\,\|S\|^m \le c_0^{-p} m^{-m/2}\, \mathbb{E}\,\|S_1 + \dots + S_m\|^m$$
$$\le c_0^{-m} m^{-m/2} \lambda^{-m} m!\, \mathbb{E} \exp(\lambda \|S_1 + \dots + S_m\|)$$
$$\le c_0^{-m} m^{-m/2} \lambda^{-m} m!\, (\mathbb{E} \exp \|\lambda S\|)^m$$

$\Big|\; u^m \le \lambda^{-m} m!\, e^{\lambda u}$

$\lambda > 0$

Especially, by replacing m by $2m$, denoting

$$A = \frac{\mathbb{E} \exp \|\lambda S\|}{\lambda c_0},$$ and

by observing that $\frac{(2m)!}{(2m)^m} \leq m!$ we get

$\mathbb{E}\,\|S\|^{2m} \leq m!\,A^{2m}$.

Thus $\mathbb{E}\,\exp(\alpha^2\|S\|^2) = \sum_{m=0}^{\infty} \frac{\alpha^{2m}}{m!}\,\mathbb{E}\|S\|^{2m} < \infty$

assuming that $\alpha A < 1$. To cover all values of α one first chooses λ and m_0 so large that

$$\alpha \frac{\mathbb{E}\exp\left(\|\lambda \sum_{k=m_0}^{\infty} \varepsilon_k u_k\|\right)}{\lambda r_0} < 1,$$

and it follows that $\mathbb{E}\,\exp\left(\alpha^2\|\sum_{k=m_0}^{\infty}\varepsilon_k u_k\|^2\right) < \infty$, which easily implies the claim.
(details for large α are an exercise!) $\square$.

4. Gaussian measures in Banach spaces

Let us recall quickly Gaussian measures (or random variables with values in $\mathbb{R}^d$.
A probability measure μ on $\mathbb{R}^d$ is Gaussian if

(1) $$\hat{\mu}(\ell) = \exp\left(i\ell\cdot m - \frac{1}{2}(Q\ell)\cdot\ell\right), \quad \ell \in \mathbb{R}^d,$$

where $m \in \mathbb{R}^d$ (the mean) and $Q \in \mathbb{R}^{d\times d}$ (covariance) is a symmetric and nonnegative definite matrix (the covariance). Thus

$$Q \geq 0 \iff \langle Q\ell, \ell\rangle_{\mathbb{R}^d} = (Q\ell)\cdot\ell \geq 0 \quad \forall \ell \in \mathbb{R}^d.$$

A r.v. $X : \Omega \to \mathbb{R}^d$ is Gaussian if μ_X is Gaussian.
In case μ is not supported on a hyperplane $\neq \mathbb{R}^d$, then μ has the density

$$\mu(dx) = \frac{1}{(2\pi)^{d/2}\sqrt{\det Q}} \exp\left(-\frac{1}{2}(Q^{-1}x)\cdot x\right) dx$$

If $\hat{\mu}_X$ is as in (1), we write $X \sim N(m, Q)$.
Observe that

$$m = \mathbb{E}X$$
$$Q = \mathbb{E}(X-m)(X-m)^T,$$

where expectations are taken componentwise.

Def. Let E be a Banach space. A probability measure μ on $(E, \mathcal{B}(E))$ is Gaussian, if for every $x' \in E'$ (the dual space of E) the map

$$x' : \underbrace{(E, \mathcal{B}(E), \mu)}_{\text{a prob. space!}} \to \mathbb{R}$$

defines a Gaussian random variable.

Def. Let X be an E-valued r.v. X is Gaussian if μ_X is Gaussian. Equivalently, $x' \circ X$ is Gaussian $\forall x' \in X'$.

We need to recall some fundamental properties of Gaussian r.v. on $\mathbb{R}^d$. We say that Gaussian r.v. $X : \Omega \to \mathbb{R}^d$ is centered if $\mathbb{E} X = 0$. A collection random variables $X_\alpha : \Omega \to \mathbb{R}^{d_\alpha}$, $\{X_\alpha\}_{\alpha \in \mathcal{A}}$, is (jointly) Gaussian if

$$(X_{\alpha_1}, \dots, X_{\alpha_k}) : \Omega \to \mathbb{R}^{d_{\alpha_1} + \dots + d_{\alpha_k}}$$

is Gaussian for any $\alpha_1, \dots, \alpha_k \in \mathcal{A}$.

If X is Gaussian, we can also think of X as an element of $L^2(\Omega) := L^2(\Omega, \mathcal{F}, \mathbb{P})$ since $\mathbb{E}|X|^2 = \operatorname{tr} Q + |m|^2$ if $X \sim N(m, Q)$.

Lemma 4.1. Let (X, Y) be centered Gaussian, where $X : \Omega \to \mathbb{R}^{d_1}$ & $Y : \Omega \to \mathbb{R}^{d_2}$. Then the following are equivalent

(i) $X \perp Y$ (i.e. X and Y are independent)

(ii) $X_k \perp Y_j$ $\forall k \in \{1, \dots, d_1\}$ and $j \in \{1, \dots, d_2\}$

(iii) $\mathbb{E} X_k Y_j = 0$ — " — — " —

Proof. (i) $\Rightarrow$ (ii) and (ii) $\Rightarrow$ (iii) are trivial.

Assume (iii). Write $Z = (X, Y)$ so that Z is Gaussian, $Z \sim N(0, Q)$, $Q \in \mathbb{R}^{d_1 + d_2}$.
Recall that

$$Q_{ij} = \mathbb{E} Z_i Z_j \qquad 1 \le i, j \le d_1 + d_2$$

(25.1)

By assumption (iii) $\mathcal{Q}$ is of block type:

$$\mathcal{Q} = \begin{bmatrix} \mathcal{Q}_1 & 0 \\ 0 & \mathcal{Q}_2 \end{bmatrix}, \quad \mathcal{Q}_1 \in \mathbb{R}^{d_1 \times d_1}, \mathcal{Q}_2 \in \mathbb{R}^{d_2 \times d_2}.$$

Thus $\quad \hat{\mu}_Z((t_1, t_2)) = e^{-\frac{1}{2}(\mathcal{Q}_1 t_1)\cdot t_1} e^{-\frac{1}{2}(\mathcal{Q}_2 t_2)\cdot t_2},$

for $(t_1, t_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Especially

$$\mu_Z = \mu_X \times \mu_Y,$$

from which the stated independence follows. □

Lemma 4.2. Let $X : \Omega \to \mathbb{R}^d$ be a r.v. Then X is Gaussian if and only if

$$a \cdot X \text{ is Gaussian } \forall a \in \mathbb{R}^d.$$

Proof. By considering $X - \mathbb{E}X$ we may assume $\mathbb{E}X = 0$. One direction is clear. If $a \cdot X$ is Gaussian $\forall a \in \mathbb{R}^d$, we may write for $\lambda \in \mathbb{R}$ for the characteristic functions

$$\varphi_{a\cdot X}(\lambda) = e^{-\frac{1}{2} q_a \lambda^2},$$

where $\quad q_a = \mathbb{E}(a\cdot X)^2 = \sum_{i,j=1}^{d} a_i a_j \mathbb{E} X_i X_j = (\mathcal{Q}a)\cdot a.$

Thus $\quad \varphi_X(t) = \varphi_{t\cdot X}(1) = e^{-\frac{1}{2}\mathcal{Q}t\cdot t}$. □

Corollary 4.3. Let X be an E-valued Gaussian random variable (E Banach). Then for any $x_1', x_2', \ldots, x_\ell' \in E'$ $(\ell \geq 1)$ the vector $(x_1'(X), \ldots, x_\ell'(X))$ is Gaussian.

Proof. Simply observe that for any $\lambda_1, \dots, \lambda_\ell \in \mathbb{R}$
$$\sum_{j=1}^{\ell} \lambda_j x_j'(X) = \Big(\sum_{j=1}^{\ell} \lambda_j x_j'\Big)(X)$$
is Gaussian. □

Def. μ is symmetric if $\mu(A) = \mu(\underbrace{-A}_{=\{-x : x \in A\}})$ $\forall A \in \mathcal{B}(E)$.
$X : \Omega \to E$ is symmetric if $-X \sim X$.
Equivalently μ_X is symmetric.

Exercise: X is symmetric $\Leftrightarrow$ $x'(X)$ is symmetric for every $x' \in E'$.

Lemma 4.4. Let μ be a Radon (equiv., tight) probability measure on $(E, \mathcal{B}(E))$. Then there is a random variable $X : \Omega \to E$ with $\mu_X = \mu$.

Proof. Since μ is Radon we may assume that E is separable. Let $\{a_n\}_{n \geq 1}$ be dense in E.

We define inductively for any $\ell \geq 1$ and $(k_1, \dots, k_\ell) \in \mathbb{N}^\ell$ the sets
$$A_{k_1, \dots, k_\ell} \subset E .$$
If $\ell = 1$ we set $A_1 = B(a_1, 1)$,
$$A_m = B(a_{k_1}, 1) \setminus \bigcup_{1 \leq j < m} A_j \quad , \quad m \geq 2.$$

In general, set by induction assuming that the sets $A_{k_1, \dots, k_\ell}$ are all already defined:

$$A_{k_1 k_2 \dots k_\ell, 1} = B(a_1, \tfrac{1}{j}) \cap A_{k_1, \dots, k_\ell},$$

$$A_{k_1 k_2 \dots k_\ell, m} = B(a_m, \tfrac{1}{j}) \cap A_{k_1, \dots, k_\ell} \setminus \bigcup_{1 \le j < m} A_{k_1, \dots, k_\ell j}, \quad m \ge 2.$$

Then $A_{k_1 \dots k_{\ell+1}} \subset A_{k_1 \dots k_\ell}$ always and

$$E = \bigcup_{k_1, \dots, k_\ell = 1}^{\infty} A_{k_1 \dots k_\ell} \qquad \forall \ell$$

(disjoint union) Obviously we may pick subsets $C_{k_1 \dots k_\ell}$ $(k_1, \dots, k_\ell \in \mathbb{N}, \ell \ge 1)$ of $[0,1]$ so that

$$C_{k_1, \dots k_{\ell+1}} \subset C_{k_1 \dots k_\ell} \quad \text{always}$$

and $$[0,1] = \bigcup_{k_1 \dots k_\ell = 1}^{\infty} C_{k_1 \dots k_\ell} \quad \text{(disjoint union)}$$

and $$\mu(A_{k_1 \dots k_\ell}) = |C_{k_1 \dots k_\ell}| \quad \forall k_1 \dots k_\ell, \ell \ge 1.$$

Pick $y_{k_1 \dots k_\ell} \in A_{k_1 \dots k_\ell}$ for each index set and define for $\ell \ge 1$ the random variable

$$X_\ell : [0,1] \to E$$

by setting $X_\ell(\omega) = y_{k_1 \dots k_\ell}$ if $\omega \in A_{k_1 \dots k_\ell}$

Then $\|X_{\ell'} - X_\ell\| \le \frac{1}{\ell}$ if $\ell' \ge \ell$ by construction. Cauchy's criterion shows that $\exists\, X = \lim_{\ell \to \infty} X_\ell$ for all ℓ.

It is an easy exercise to verify that now $$\mu_X = \mu. \qquad \square$$

The following result is fundamental

Theorem 4.5 (Fernique) Let μ be a symmetric (Radon) Gaussian measure on a Banach space E. Then there is $\delta > 0$ s.t.

$$\int_E e^{\delta \|x\|^2} \mu(dx) < \infty.$$

Proof. It is clearly enough to show the following: if $\lambda > 0$ and $r > 0$ satisfy

$$(2) \quad \log\left(\frac{1-\mu(\bar{B}(0,r))}{\mu(\bar{B}(0,r))}\right) + 32\lambda r^2 + 1 \leq 0$$

then

$$(3) \quad \int_E e^{\lambda \|x\|^2} \mu(dx) \leq e^{16\lambda r^2} + \frac{e^2}{e^2-1}.$$

Since μ is Radon, we may apply lemma 6.4 to pick E-valued and independent r.v.:s with

$$\mu_X = \mu_Y = \mu.$$

We claim that if

$$\tilde{X} = \frac{1}{\sqrt{2}}(X+Y) \quad \text{and} \quad \tilde{Y} = \frac{1}{\sqrt{2}}(X-Y),$$

then likewise $\tilde{X} \perp \tilde{Y}$ and $\mu_{\tilde{X}} = \mu_{\tilde{Y}} = \mu$.
If we show that for any

$$a_1', \ldots, a_\ell' \in E' \quad \text{and} \quad b_1', \ldots, b_\ell' \in E'$$

the vectors

$$(a_1'(\tilde{X}), \ldots, a_\ell'(\tilde{X})) \quad \text{and} \quad (b_1'(\tilde{Y}), \ldots, b_\ell'(\tilde{Y}))$$

are independent, then the independence of $\tilde{X}$ and $\tilde{Y}$ follows from Lemma 1.4.

In turn, by Lemma 4.3 it is enough to verify for each $1 \leq j, k \leq \ell$ that

$$(4) \qquad \mathbb{E}\, a_j'(\tilde{X})\, b_k'(\tilde{Y}) = 0$$

since the independence of $\tilde{X}$ and $\tilde{Y}$ (together with corollary 6.3) yields that the joint distribution of the vectors is a Gaussian on $\mathbb{R}^{2\ell}$. Since $X \perp Y$ and $\mu_X = \mu_Y$ we get

$$\mathbb{E}\, a_j'(\tilde{X})\, b_k'(\tilde{Y}) = \tfrac{1}{4}\big(\mathbb{E}\, a_j'(X)\, b_k'(X)$$

$$+ \mathbb{E}\, a_j'(Y)\, b_k'(X) - \mathbb{E}\, a_j'(X)\, b_k'(Y)$$

$$- \mathbb{E}\, a_j'(Y)\, b_k'(Y)\big)$$

$$= \tfrac{1}{4}\big(\mathbb{E}\, a_j'(Y)\,\mathbb{E}\, b_k'(X) - \mathbb{E}\, a_j'(X)\,\mathbb{E}(b_k'(Y))\big) = 0.$$

Moreover, if $x' \in E'$ we have

$$\widehat{\mu}_{\tilde{X}}(x') = \mathbb{E} \exp\Big(\frac{i x'(X+Y)}{\sqrt{2}}\Big)$$

$$= \mathbb{E} \exp\Big(\frac{i x'(X)}{\sqrt{2}}\Big)\, \mathbb{E}\Big(\frac{i x'(Y)}{\sqrt{2}}\Big)$$

$$= \Big(\exp - \tfrac{1}{4}\, \mathbb{E}|x'(X)|^2\Big)^2$$

$$= \widehat{\mu}_X(x').$$

Hence $\tilde{X} \sim X$ and similarly $\tilde{Y} \sim X$.

Let us then assume that $t \geq s \geq 0$. We obtain

$$\mathbb{P}(\|X\| \leq s)\, \mathbb{P}(\|Y\| > t)$$

$$= \mathbb{P}\Big(\frac{\|X-Y\|}{\sqrt{2}} \leq s \text{ and } \frac{\|X+Y\|}{\sqrt{2}} > t\Big)$$

$$\leq \mathbb{P}\Big(\|X\| > \frac{t-s}{\sqrt{2}} \text{ and } \|Y\| > \frac{t-s}{\sqrt{2}}\Big)$$

$$\|X\| \geq \frac{\|X+Y\| - \|X-Y\|}{2}$$

$$\|Y\| \geq \frac{\|X+Y\| - \|X-Y\|}{2}$$

$$= \mathbb{P}\left(\|X\| > \frac{t-s}{\sqrt{2}}\right)^2$$

Since $P_X = P_Y$ we deduce

$$(5) \quad \mathbb{P}(\|X\| > t) \leq \frac{\mathbb{P}\left(\|X\| > \frac{t-s}{\sqrt{2}}\right)^2}{\mathbb{P}(\|X\| \leq s)}.$$

Let $t_0 = r$ ($r > 0$ fixatta), $t_{n+1} = r + \sqrt{2}\, t_n$.

Denote $\quad a_n := \dfrac{\mathbb{P}(\|X\| > t_n)}{\mathbb{P}(\|X\| \leq r)}$, $n = 0, 1, \ldots$

By applying (5) it follows that

$$a_{n+1} = \frac{\mathbb{P}(\|X\| > t_{n+1})}{\mathbb{P}(\|X\| \leq r)} = \frac{\mathbb{P}(\|X\| > r + \sqrt{2}\, t_n)}{\mathbb{P}(\|X\| \leq r)}$$

$$\leq \left(\frac{\mathbb{P}(\|X\| > t_n)}{\mathbb{P}(\|X\| \leq r)}\right)^2 = a_n^2$$

Thus $\quad a_n \leq (a_0)^{2^n}, \quad n = 0, 1, \ldots$

One checks by induction that $\quad 2^{2+n/2} r > t_n$.
Thus

$$\mathbb{P}(\|X\| > 2^{2+n/2} r) \leq \mathbb{P}(\|X\| > t_n)$$

$$\leq a_n = (a_0)^{2^n}.$$

Finally $\quad \displaystyle\int_{\|X\| > 4r} e^{\lambda \|X\|^2} \leq \sum_{n=0}^{\infty} \mathbb{P}(\|X\| > 2^{\frac{n}{2}+2} r)\, e^{\lambda r^2 2^{n+5}}$

$$\leq \sum_{n=0}^{\infty} e^{2^n (\log a_0 + 32 \lambda r^2)}.$$

This yields (2) & (3) as $\log a_0 + 32 \lambda r^2 \leq -1$ by (2) $\square$

Lemma 4.6. Let Y_n $(n\geq 1)$, Y be real-valued r.v.s s.t. each Y_n is Gaussian and $Y_n \xrightarrow[n\to\infty]{} Y$ a.s. Then Y is Gaussian as well.

Proof. Exercise (use characteristic functions). $\square$

Lemma 4.7. Let $Y \sim N(0,1)$. Then for $\lambda > 0$

$$\frac{1}{\sqrt{2\pi}}\left(\frac{1}{\lambda} - \frac{1}{\lambda^3}\right) e^{-\lambda^2/2} \leq \mathbb{P}(Y > \lambda) \leq \frac{1}{\sqrt{2\pi}\,\lambda} e^{-\lambda^2/2}$$

Especially,
$$\lim_{\lambda\to\infty} \frac{\mathbb{P}(Y>\lambda)}{\frac{1}{\sqrt{2\pi}\,\lambda} e^{-\lambda^2/2}} = 1.$$

Proof.
$$\int_\lambda^\infty e^{-t^2/2}\,dt = \int_\lambda^\infty \left(t e^{-t^2/2}\right)\frac{dt}{t} = \frac{e^{-\lambda^2/2}}{\lambda} - \int_\lambda^\infty e^{-t^2/2}\frac{dt}{t^2}.$$

Moreover,
$$\int_\lambda^\infty \left(e^{-t^2/2} t\right)\frac{dt}{t^3} \leq \lambda^{-3}\int_\lambda^\infty e^{-t^2/2} t\,dt = \lambda^{-3} e^{-\lambda^2/2}. \square$$

Example. Let $E = c_0$. Consider

$$(6) \qquad X := (X_1, X_2, \ldots),$$

where the X_n are independent, $X_n \sim N(0, \sigma_n^2)$. If $\sigma_n \to 0$ quickly, clearly X yields a c_0-valued random variable. 'How quickly' is answered precisely by

Theorem 4.8 Formula (6) defines a c_0-valued r.v. if and only if $\sum_{n=1}^\infty \exp(-\varepsilon \sigma_n^{-2}) < \infty$ $\forall \varepsilon > 0$.

Proof. We may clearly assume $\sigma_n \xrightarrow[n\to\infty]{} 0$. Then, by Lemma 4.7

$$\sum_{n=1}^{\infty} \mathbb{P}(|X_k| > \delta) < \infty \iff \sum_{n=1}^{\infty} \frac{\sigma_n}{\delta} \exp\left(\frac{-\delta^2}{2\sigma_n^2}\right) < \infty$$

By Borel-Cantelli (and independence) the left hand side is equivalent to

$$\mathbb{P}(|X_k| \leq \delta \text{ eventually}) = 1$$

This is necessary for X being c_0-valued, taking $\delta = 1/j$ $(j = 1,2,3,\dots)$ we see that it is also sufficient. Finally, one may forget the factor σ_n/δ since $e^{-\eta/x^2} \geq \frac{1}{x}$ for $0 < x \leq x_0(\eta)$, $\eta > 0$.

Finally, assume that $X(\omega) \in c_0$ a.s. In order to check (weak $\Rightarrow$ Borel-) measurability and Gaussianity, observe that for any $h \in c_0' = \ell_1$ one has a.s.

$$h(X) = \lim_{n\to\infty} \underbrace{\sum_{k=1}^{n} X_k h_k}_{\text{Gaussian for every } n}.$$

Just apply Lemma 4.6. □

Corollary 4.9. (of Fernique's Thm 4.5) If X is a Gaussian Radon variable (values in E), then (even if X is not symmetric) $\exists\ \delta > 0$ s.t.

$$\mathbb{E} \exp(\delta \|X\|^2) < \infty.$$

Proof. Let $Y \sim X$, $Y = Y(\omega')$, $X = X(\omega)$, as usual take $\bar\Omega = \Omega \times \Omega$, $\bar{\mathbb{P}} = \mathbb{P} \times \mathbb{P}$. Then $Y \perp X$

and applying Fernique on $X - Y$ we obtain

$$\int_\Omega \left(\int_\Omega e^{\delta \| X(\omega) - Y(\omega') \|^2} \mathbb{P}(d\omega) \right) \mathbb{P}(d\omega') < \infty$$

for some ω'. Hence for some ω', by denoting $Y(\omega') = a \in E$ we have

$$\mathbb{E}\, e^{\delta \| X(\omega) - a \|^2} < \infty .$$

Use $\frac{\delta}{2} \| X(\omega) - a \|^2 \le \delta \left(\| X(\omega) \|^2 + \| a \|^2 \right)$. $\square$

!! **Convention** From now on, in this section E is a separable Banach space. Especially, every probability measure on E is automatically tight ($\Leftrightarrow$ Radon)

Def. Let μ be a Gaussian measure on E. Set*) $m_\mu := \int_E x \, d\mu(x) \in E$

Define* the operator $R_\mu : E' \to E$ by

$$R_\mu g' = \int_E \left(g'(x) - g'(m_\mu) \right) (x - m_\mu)\, \mu(dx)$$

Observe that if μ is symmetric we have

$$R_\mu g' = \int_E x g'(x)\, \mu(dx) .$$

We need compute the Fourier-transform in terms of the quantities above

* By Bochner-integral (Fernique yields integrability)

(m_μ is the <u>mean</u> of μ (or X if $\mu = \mu_X \ldots$),
the quadratic form

$$g' \mapsto g' R_\mu g' := g'(R_\mu(g'))$$

is the covariance of μ)

<u>Lemma 4.10</u> (i) If μ is Gaussian on E, then

$$\hat{\mu}(g') = \exp\left(i g'(m_\mu) - \tfrac{1}{2} g' R_\mu g'\right).$$

(ii) Conversely, if μ is a prob. measure on E and $\hat{\mu}(g') = \exp(i M(g') - \frac{1}{2} R(g';g'))$,

where $M: E' \to \mathbb{R}$ is linear*) and
R is a bilinear*) form on E', then μ is Gaussian
and $M(g') = g'$ and $R(g';g') = g' R_\mu g'$
$\forall g' \in E'$.

<u>Proof.</u> (i) Observe that $(\Omega = (E, \mathcal{B}(E), \mu))$

$$\mathbb{E}\, g'(x) = g'(m_\mu),$$

$$\mathbb{E}\,(g'(x) - g'(m_\mu))^2 = g' R_\mu g'.$$

(ii) It follows that $\mathbb{E}\, e^{i\lambda g'(x)} =$ $(\lambda \in \mathbb{R})$

$$= \exp\left(i\lambda M(g') - \tfrac{1}{2}\lambda^2 R(g';g')\right),$$

whence $R(g';g') \geq 0$ and $g'(x)$, is Gaussian.
Rest is clear. $\square$

<u>Corollary 4.11</u> A Gaussian measure μ on E is symmetric iff $m_\mu = 0$

<u>Proof</u> Exercise.

*) no continuity assumed here!

We now define a central concept;

The Cameron-Martin space

Def. If $h \in E$ let

$$\|h\|_{H_\mu} := \sup\{g'(h) \mid g' \in E',\ g' R_\mu g' \leq 1\}$$

Then the Cameron-Martin space of μ is

$$H_\mu := \{h \in E : \|h\|_{H_\mu} < \infty\}.$$

Observe that $H_\mu = H_{\mu+a}$ for $a \in E$, where $(\mu+a)(A) = \mu(A-a)$. Thus, for simplicity we will consider often just symmetric measures.

Def. $$E'_\mu := \overline{\{g' - g'(m_\mu) \mid g' \in E'\}}^{L^2(d\mu)}$$

Theorem 4.12 R_μ extends to an injective operator

$$R_\mu : E'_\mu \to E.$$

In fact, $R_\mu(E'_\mu) = H_\mu$ and the map

$$\boxed{R_\mu : E'_\mu \to H_\mu}$$

is an isometric isomorphism. Especially, H_μ is a Hilbert space!

Proof. Assume μ symmetric. The formula

$$(9) \qquad R_\mu \varphi = \int_E x\,\varphi(x)\,\mu(dx)$$

automatically extends R_μ to a bounded linear map

$$(10) \qquad \widetilde{R_\mu} : L^2(d\mu) \to E$$

since by Cauchy-Schwarz we have

$$\|\widetilde{R_\mu}\varphi\|_E \leq \int_E \|x\| |\varphi(x)| \mu(dx)$$
$$\leq C \left(\int_E |\varphi(x)|^2 \mu(dx)\right)^{1/2},$$

where $C := \left(\int_E \|x\|^2 \mu(dx)\right)^{1/2} < \infty$ by Fernigues Thm.
If $\lambda \in E'$, we have

$$(11) \quad \|\lambda\|^2_{L^2(d\mu)} = \int_E |\lambda(x)|^2 \mu(dx) \leq \|\lambda\|^2_{E'} \int_E \|x\|^2 \mu(dx)$$
$$\leq C^2 \|\lambda'\|^2_{E'},$$

which shows that $E' \subset L^2(d\mu)$, and E_μ' is well-defined. Especially, we obtain for any $g' \in E'$

$$g'(R_\mu g') \leq \|g'\|_{E'} \|R_\mu g'\|_E$$
$$\leq \|g'\|_{E'} \|R_\mu\|_{L^2(d\mu)\to E} \|g'\|_{L^2(d\mu)}$$
$$\leq C^2 \|g'\|^2_{E'}$$

Hence, if $\|g'\|_{E'} \leq C^{-1}$ we have $g'(R_\mu g') \leq 1$ and for any $h \in E$

$$(12) \quad \|h\|_{H_\mu} \geq \sup\{g'(h) \mid \|g'\|_{E'} \leq C^{-1}\} = C^{-1}\|h\|_E$$

By this inequality and a standard reasoning* one checks that $H_\mu \subset E$ is a Banach space.

---

*) This is an exercise. Hint: enough to check that H_μ is complete which is not difficult by (12) and (7).

We observe that for any $g' \in E'$ one has

(13) $$g' R_\mu g' = g'\left(\int_E x\, g'(x)\,\mu(dx)\right) = \int_E (g'(x))^2 \mu(dx) = \|g'\|^2_{L^2(d\mu)}$$

Hence, if $\lambda \in E'$ we may estimate the $\| \ \|_{H_\mu}$-norm of $R_\mu \lambda$ by computing

$$\|R_\mu \lambda\|_{H_\mu} = \sup_{g' R_\mu g' \leq 1} |g'(R_\mu \lambda)|$$
$$= \sup_{\|g'\|_{L^2(d\mu)} \leq 1} \int_E g'(x)\lambda(x)\mu(dx)$$
$$= \sup_{\|g'\|_{L^2(d\mu)} \leq 1} \langle g', \lambda \rangle_{L^2(d\mu)}$$
$$= \|\lambda\|_{L^2(d\mu)}$$
$$= \|\lambda\|_{E'_\mu}$$

It follows that $R_\mu : E' \to H_\mu$ is an isometry on a dense subspace $E' \subset E'_\mu$, whence see that

(14) $R_\mu : E'_\mu \to H_\mu$ is an isometry

by an easy application of (12).

It remains to check that R_μ is onto. Assume that $h \in H_\mu$. Then by the definition of $\|\cdot\|_{H_\mu}$ we see that

$$|g'(h)| \leq \|g'\|_{E'_\mu} \|h\|_{H_\mu} \quad , g' \in E'$$

Hence the map $g' \mapsto g'(h)$ defines a bounded linear functional on the dense subset $E' \subset E'_\mu$.

By continuity, this is a restriction of an element $\eta \in E'_\mu$, whence by the Riesz representation thm there is $\varphi \in E'_\mu$ so that for all $g' \in E'$

$$g'(h) = \eta(g') = \langle \varphi, g' \rangle_{E'_\mu} = \langle \varphi, g' \rangle_{L^2(d\mu)}$$
$$= \int_E g'(x)\varphi'(x)\,\mu(dx)$$
$$= g'\Big(\int_E x\varphi'(x)\mu(dx)\Big)$$
$$= g'(R_\mu \varphi)$$

Thus $h = R_\mu \varphi$, and we are done. □

<u>Remark.</u> If $\mu \sim X$, where $X : \Omega \to E$ is Gaussian with distribution μ, we have
$$m_\mu = \mathbb{E} X .$$
We denote $R_\mu = R_X$, $E'_\mu = E'_X$, $H_\mu = H_X$, if convenient. For example,
$$R_X \varphi = \mathbb{E}\, X \varphi(X),$$
$$g' R_X g' = \mathbb{E}\,(g'(X))^2 \quad \text{etc.}$$

<u>Example</u> Let X be the c_0-valued r.v. defined on p. 31. Thus,

$$X = (X_1, X_2, \dots),$$

where X_n:s are independent, $X_n \sim N(0, \sigma_n^2)$, where $\sigma_n \to 0$ quickly enough. Let us determine R_X, E'_X and H_X.

Now $E = c_0$, $E' = \ell_1$. If $g' \in \ell^1 = (c_0)'$, we have

$$R_X g' = \mathbb{E}\, X g'(X)$$
$$= \left(\mathbb{E}\, X_1 \left(\sum_{n=1}^{\infty} g_n' X_n\right), \mathbb{E}\, X_2 \left(\sum_{n=1}^{\infty} g_n' X_n\right), \dots\right)$$
$$= \left(g_1' \mathbb{E} X_1^2, g_2' \mathbb{E} X_2^2, \dots\right)$$
$$= \left(g_1' \sigma_1^2, g_2' \sigma_2^2, \dots\right)$$

Thus $\quad g' R_X g' \leq 1 \iff \sum_{n=1}^{\infty} \sigma_n^2 g_n'^2 \leq 1$

and for $h \in c_0$

$$\|h\|_{H_X} = \sup_{\left\{\substack{\sum_{n=1}^{\infty} \sigma_n^2 g_n'^2 \leq 1, \\ g' \in \ell^1}\right.} \sum_{n=1}^{\infty} h_n g_n' = \sup_{\sum_{n=1}^{\infty} y_n^2 \leq 1} \sum_{n=1}^{\infty} h_n \sigma_n^{-1} y_n$$

plays no role!

$$= \left(\sum_{n=1}^{\infty} h_n^2 \sigma_n^{-2}\right)^{1/2}.$$

Hence $\quad H_X = H_{\mu_X} = \left\{h \in \ell^1 : \|h\|_{H_X} = \left(\sum_{n=1}^{\infty} \sigma_n^{-2} h_n^2\right)^{1/2} < \infty\right\}$

<u>Exercise:</u> Show that $E_X' = \left\{(h_n)_{n=1} : \sum_{n=1}^{\infty} \sigma_n^2 h_n^2 < \infty\right\}$.

<u>Exercise:</u> Assume that $\sum_{n=1}^{\infty} \sigma_n < \infty$. Show that then the above X is also ℓ^2-valued. Compute again H_μ. Is the answer different?

Our next goal is to show that H_X enables us to obtain representations of the form

$$X \sim \sum_{n=1}^{\infty} X_n u_n, \quad u_n \in E, \quad X_n \sim N(0,1) \text{ i.i.d.}$$

For that end we need couple of auxiliary results.

Def. Let μ_α ($\alpha \in A$) be an E valued prob. measure. We say that the family $\{\mu_\alpha\}_{\alpha\in A}$ is tight if for any $\varepsilon > 0$ there is $K \subset E$ compact so that

$$\mu_\alpha(K) \geq 1-\varepsilon \quad \forall\, \alpha \in A.$$

Rem. Family $\{X_\alpha\}_{\alpha\in A}$ of E-valued r.v:s is tight if $\{\mu_{X_\alpha}\}_{\alpha\in A}$ is tight.

(Prokhorov)

Thm 4.13 If $\{X_n\}_{n\geq 1}$ is a tight sequence of E-valued r.v.:s, then there is a subsequence $\{X_{n_k}\}_{k\geq 1}$ and an E-valued r.v. X so that

$$X_{n_k} \xrightarrow[k\to\infty]{d} X.$$

Proof. (Sketch) Denote $\mu_{X_n} =: \mu_n$. For $\ell \geq 1$, choose K_ℓ so that

$$\mu_n(K_\ell) \geq 1-\ell^{-1} \quad \forall\, n\geq 1,\ \ell \geq 1.$$

By functional analysis (or by an easier version of Prokhorov), the set of probability measures on compact seprarable metric space is compact with respect to weak convergence (integration against continuous functions). Apply this on $\mu_n|K_\ell$ for any fixed ℓ.

Get subsequene so that $\mu_{n_k}|K_1$ converges weakly,

a further subsequence so restrictions to K_2 converge weakly, and so on. Pick a diagonal subsequence. □

Lemma 4.14 Let X_n, $n \geq 1$, be a tight sequence of E-valued r.v.s such that

$$\hat{\mu}_{X_n}(g') \xrightarrow[n\to\infty]{} 1 \quad \text{for all } g' \in E'.$$

Then $X_n \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$.

Proof. As in the case of scalar-valued r.v.s, it is enough to show that

$$(15) \qquad X_n \xrightarrow[n\to\infty]{d} 0$$

since testing against the continuous function

$$\varphi(x) := \max\left(0, \min\left(1, 2(1 - \varepsilon^{-1}\|x\|)\right)\right)$$

shows then that $\mathbb{P}(\|X_n\| \leq \varepsilon) \xrightarrow[n\to\infty]{} 1$ for any $\varepsilon > 0$.

If (15) would not be true, we could find $\varphi \in C(E)$, φ bounded, so that

$$\mathbb{E}\,\varphi(X_n) \underset{n\to\infty}{\not\to} \varphi(0).$$

By moving to a subsequence we might assume that e.g. $(16) \quad \mathbb{E}\,\varphi(X_n) \leq \varphi(0) - \delta_0 \quad \forall\, n \geq 1,$

where $\delta_0 > 0$. By Lemma 4.13, moving to a further subsequence, we may also assume that

$$X_n \xrightarrow[n\to\infty]{w} X.$$

However, then $\hat{\mu}_{X_n}(g') \to \hat{\mu}_X(g') \quad \forall\, g' \in E'$,

whence $\hat{\mu}_X(g') = 1 \quad \forall\, g'$. Thm 2.3 implies that $\mu_X = \delta_0$ (Dirac delta at the origin). This contradicts (16). $\square$

Lemma 4.15 Let $X, Y : \Omega \to E$ be E-valued r.v:s. If $X \perp Y$, we have for all $A \in \mathcal{B}(E)$

$$(16) \qquad \mu_{X+Y}(A) = \int_E \mu_X(A-y)\,\mu_Y(dy)$$

Proof. Clearly $\mu_{X+Y}(A) = \int_{E\times E} \chi_A(x+y)\,\mu_X(dx)\,\mu_Y(dy)$

$$= \int_E \left(\int \chi_A(x+y)\,\mu_X(dx) \right) \mu_Y(dy) = \int_E \mu_X(A-y)\,\mu_Y(dy),$$

where we applied vector-valued Fubini $\square$

Lemma 4.16 Assume that $\boxed{X = X_n + Y_n}$, $n \geq 1$, where X, X_n, Y_n are E-valued r.v:s and X_n, Y_n are symmetric for each $n \geq 1$, and $\boxed{X_n \perp Y_n}$. Then both families $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ are tight.

Proof. Denote $\mu_X = \mu$, $\mu_{X_n} = \mu_n$, $\mu_{Y_n} = \nu_n$.

As E is separable, μ is tight and we may pick $K \subset E$ compact so that $\mu(K) \geq 1-\varepsilon$.

By replacing K by $K \cup (-K)$ we may assume that K is symmetric. By applying (16) with $A = K$ and $X = X_n$, $Y = Y_n$, it follows that

$$\mu(K) = \int_E \mu_n(A-y)\,\nu_n(dy)$$

and hence we may pick $y_n \in E$ so that

$$\mu_n(K - y_n) \geq 1-\varepsilon.$$

Since μ_n and K are symmetric, also $\mu(K+y_n) \geq 1-\varepsilon$, whence

$$\mu_n((K-y_n)\cap(K+y_n)) \geq 1-2\varepsilon$$

Especially, $(K-y_n)\cap(K+y_n) \neq \emptyset$ so that for some $k_n, k_n' \in K$ $\quad k_n - y_n = k_n' + y_n$, or

$$y_n = \frac{k_n - k_n'}{2}.$$

This yields that $y_n \in \frac{1}{2}K + \frac{1}{2}K$ and, finally

$$K + y_n \subset K + \tfrac{1}{2}K + \tfrac{1}{2}K =: \widetilde{K}.$$

Also $\widetilde{K}$ is compact (why) and $\mu_n(\widetilde{K}) \geq 1-\varepsilon \;\forall n \geq 1$, and we obtain the tightness of $\{\mu_n\}_{n\geq 1}$. Same proof works for $\{\nu_n\}_{n\geq 1}$. $\square$

Lemma 4.17. (i) Let μ be a Borel-measure on a complete separable metric space. Then $L^2(d\mu)$ is separable.

(ii) H_μ is separable if μ is Gaussian on a separable E.

Proof. (i) A guided exercise.

(ii) Follows from (i) since $L^2(d\mu)$ is separable, $E_\mu' \subset L^2(d\mu)$ is a closed subspace, and H_μ is isomorphic to E_μ' by Thm. 4.12. $\square$

Theorem 4.18 Let $X : \Omega \to E$ be Gaussian, where E is a separable Banach space. Then

$$X \sim m_X + \sum_{n=1}^{\infty} X_n h_n ,$$

where $X_n \sim N(0,1)$ are i.i.d and $\{h_n\}_{n\geq 1}$ is an orthonormal basis of H_X.

Proof. May assume that $\mu = \mu_X$ is symmetric ($m_\mu = 0$). Case $\dim(H_\mu) < \infty$ easy, assume that $\dim(H_\mu) = \infty$. In any case E'_μ and H_μ are separable.

Apply the density of E' in E'_μ and Gramm-Schmidt algorithm to pick a sequence $(g'_n)_{n\geq 1}$ from E' so that

(17) $(g'_n)_{n\geq 1}$ is an orthonormal basis of E'_μ

Then, if $h_n := R_\mu g'_n$, $n \geq 1$, the sequence $(h_n)_{n\geq 1}$ is an orthonormal basis of H_μ.

We now consider $(E, \mathcal{B}(E), \mu)$ as a probability space. Then (by definition of Gaussianity of μ) g'_n:s are (jointly Gaussian). In fact,

(18) $g'_n \sim N(0,1)$, $n \geq 1$, and g_n:s are i.i.d.

To check this we note that

$$E\, g'_n(x)\, g'_m(x) = \int g'_n(x) g'_m(x)\, d\mu$$

$$= \langle g'_n, g'_m \rangle_{L^2(d\mu)} = \langle g'_n, g'_m \rangle_{E'_\mu} = \delta_{n,m}.$$

It is thus enough to show that

(19) $$x = \lim_{N\to\infty} \sum_{n=1}^{N} g'_n(x)\, h_n \quad \text{for } \mu\text{-a.e } x \in E$$

Denote $$S_N(x) = \sum_{n=1}^{N} g'_n(x) h_n$$

and $$R_N(x) = x - S_N(x).$$

We claim that as E-valued random variables

(20) $\qquad S_N(x) \perp R_N(x) \quad , \quad \forall N \geq 1.$

By Thm 2.2 it is enough to show that

$$(s_1'(S_n), \dots, s_\ell'(S_n)) \perp (r_1'(R_n), \dots, r_\ell'(R_n))$$

For any $r_1', \dots, r_\ell', s_1', \dots, s_\ell' \in E'$. Since the above vectors are jointly Gaussian, it is enough to check that

$$r'(R_n) \perp s'(S_n) \quad \text{For all } r', s' \in E'.$$

But we may directly compute:

$$\mathbb{E}\, r'(R_n)\, s'(S_n) = \mathbb{E}\left(r'(x) - \sum_{n=1}^N g_n'(x)\, r'(h_n)\right) \times \left(\sum_{n=1}^N g_n'(x)\, s'(h_n)\right)$$

$$= \sum_{n=1}^N \langle r', g_n' \rangle_{E_\mu'} s'(h_n) - \sum_{n=1}^N \sum_{k=1}^N \underbrace{\langle g_n', g_k' \rangle_{E_\mu'}}_{=\delta_{n,k}} r'(h_n)\, s'(h_n)$$

$$= \sum_{n=1}^N s'(h_n)\left(\langle r', g_n' \rangle_{E_\mu'} - r'(h_n)\right) = 0$$

since $r'(h_n) = r' R_\mu g_n' = \langle r', g_n' \rangle_{E_\mu'}$.
Thus, (20) is true.

From the identity $S_N(x) + R_N(x) = x$, (20), and Lemma 4.16 it follows that

(21) $\qquad \{R_N\}_{N \geq 1}$ is tight.

Since R_N is Gaussian, we have

(22) $\qquad \hat{\mu}_{R_N}(g') = \exp\left(-\frac{1}{2}\mathbb{E}\,(g'(R_N))^2\right).$

Given $g' = \sum_{n=1}^{\infty} \alpha_n g'_n$ in E'_μ ($\Leftrightarrow$ in $L^2(d\mu)$),

we have $\alpha_n = \langle g', g'_n \rangle_{E'_\mu}$ and

$$\mathbb{E}\,|g'(R_N)|^2 = \mathbb{E}\Big|\sum_{n=1}^{\infty} \alpha_n g'_n(x) - \sum_{n=1}^{\infty} \alpha_n \sum_{j=1}^{N} g'_j(x)\underbrace{g'_n(h_j)}_{=\delta_{n,j}}\Big|^2$$

$$= \mathbb{E}\Big|\sum_{n=N+1}^{\infty} \alpha_n g'_n(x)\Big|^2 = \sum_{n=N+1}^{\infty} \alpha_n^2$$

$$\xrightarrow[N\to\infty]{} 0 .$$

Thus, by (22) $\hat{\mu}_{R_N}(g') \xrightarrow[N\to\infty]{} 1 \quad \forall\, g' \in E'$.

At this stage, this fact together with (21) and Lemma 4.14 implies that

$$R_N(x) \xrightarrow[N\to\infty]{} 0 \quad \text{in probability},$$

or, in other words

$$S_N(x) \xrightarrow[N\to\infty]{} x \quad \text{in probability}.$$

This is then upgraded to a.s. convergence by Thm 3.3 □

<u>Remark</u>. Observe that the representation in Thm 4.18 is not unique different choices for the ON-basis of H_μ can be used, some being better for given application at hand.

<u>Corollary 4.19</u> For a Gaussian measure μ on E it holds that

$$\mu(\overline{H_\mu}^E) = 1$$
□

closure of H_μ in $\|\cdot\|_E$

Lemma 4.20 Let μ_1, μ_2 be probability measures on E. Then μ_1 and μ_2 are mutually singular if and only if $\|\mu_1 - \mu_2\|_{TV} = 2$

↑ total variation norm

Proof. Easy exercise. □

If μ is a finite measure on $(E, \mathcal{B}(E))$ and

$$f: E \to X \qquad (X \text{ topological space})$$

is a continuous map, the push forward $f^*(\mu)$ is the measure on $(X, \mathcal{B}(X))$ defined by

$$f^*(\mu)(A) = \mu(f^{-1}(A)) \quad , A \in \mathcal{B}(X)$$

Lemma 4.21 (i) Let $f: E \to X$ (X topol. space) is continuous and μ_1, μ_2 are prob. measures on E, then

$$\|f^*(\mu_1) - f^*(\mu_2)\|_{TV} \leq \|\mu_1 - \mu_2\|_{TV}$$

(ii) If f is as above, μ is a finite Borel measure on E, and $h \in L^1(d f^*(\mu))$, then

$$\int_X h(x)\, f^*(\mu)(dx) = \int_E h \circ f(y)\, \mu(dy)$$

Proof. (ii) This is clear for $h = \chi_A$, $A \in \mathcal{B}(X)$, hence for simple functions, and then by standard reasoning for integrable functions.

(i) Denote $\lambda = \mu_1 - \mu_2$ and choose a Borel-measurable $b: X \to \{-1, 1\}$ so that

$$\|f^*(\lambda)\|_{TV} = \int_X b(x)\, f^*(\lambda)(dx)$$

Then $$\|f^*(\lambda)\|_{TV} \overset{(ii)}{=} \int_E (b \circ f)(y)\, \lambda(dy) \leq \|\lambda\|_{TV} \qquad \square$$

If μ is Gaussian, $\mu \sim N(0, C)$ in $\mathbb{R}^d$, where the covariance matrix C is invertible, the density of μ is

$$\mu(dx) = (2\pi)^{-\frac{d}{2}} |C|^{-\frac{1}{2}} \exp(-\tfrac{1}{2} x C^{-1} x)$$

Let us shift μ by vector $h \in \mathbb{R}^d$ – define $\mu_h(A) = \mu(A-h)$. Then μ_h has the density

$$\mu_h(dx) = (2\pi)^{-\frac{d}{2}} |C|^{-\frac{1}{2}} \exp(-\tfrac{1}{2} x C^{-1} x)\, \rho_h(x),$$

where

$$\rho_h(x) := \exp(C^{-1}(h) \cdot x - \tfrac{1}{2} h C^{-1} h)$$

In the infinite-dimensional dimension there arise some subleties: a shifted Gaussian may be singular w.r. the original Gaussian!
We will sort this out in the following

Theorem 4.22 Let μ be a Gaussian measure on a separable Banach space E. For given $h \in E$ denote $\mu_h(A) = \mu(A-h)$ for $A \in \mathcal{B}(E)$.

(the transation of μ by the vector h).

(i) if $h \notin H_\mu$, then μ and μ_h are mutually singular.

(ii) if $h \in H_\mu$, then μ and μ_h are absolutely continuous w.r.t. each others. The Radon-Nikodym derivative is given by

$$\frac{d\mu_h}{d\mu} = \rho_h(x), \text{ where}$$

$$\rho_h(x) := \exp(R_\mu^{-1}(h)(x) - \tfrac{1}{2} \|h\|^2_{H_\mu})$$

Note that this means that

$$\mu_n(A) = \int_A \rho_n(x)\mu(dx) \quad , A \in \mathcal{B}(E).$$

Before proving the Thm, let us observe

Lemma 4.23 Let μ be a centered Gaussian on, (separable) E and let $\varphi \in E'_\mu$. Then φ is Gaussian as a random variable on

$$(E, \mathcal{B}(E), \mu)$$

and (23) $\displaystyle\int_E \exp(a|\varphi(x)|)\mu(dx) < \infty \quad \forall a \in \mathbb{R}.$

Proof. Let $g_k \in E'$ with $\|g_k - \varphi\|_{E'_\mu} \xrightarrow[k\to\infty]{} 0$

Thus $\mathbb{E}|g_k - \varphi|^2 = \displaystyle\int_E |g_k(x) - \varphi(x)|^2 \mu(dx)$

$$= \|g_k - \varphi\|^2_{E'_\mu} \to 0.$$

Thus $g_k \xrightarrow{\mathbb{P}} \varphi$ as $k \to \infty$, and the Gaussianity of φ is a consequence of Lemma 4.6. A scalar valued Gaussian has all exponential moments finite:

$$\mathbb{E}\exp(a|\varphi|) < \infty \quad \forall a \in \mathbb{R},$$

which yields 23. □

Proof of Thm 4.22 We may clearly assume that $m_\mu = 0$ (i.e μ is centered).

(i) Since $h \notin H_\mu$, we may pick $g'_k \in E'$

so that $\mathbb{E}|g_k(x)|^2 = g'_k R_\mu g'_k = 1$ and

$$g'_k(h) \geq k \quad \text{for } k \geq 1.$$

Denote $\lambda_k := (g_k')^*(\mu)$, i.e λ_k is the distribution of the Gaussian r.v g_k'. Thus λ_n is a probability measure on $\mathbb{R}$. Similarly, denote $\tilde{\lambda}_k := g_k'^*(\mu_h)$.

Clearly g_k' is centered and $\mathbb{E}|\lambda_k|^2 = g_k' R_\mu g_k = 1$.

Thus $\lambda_k \sim N(0,1)$.

Analogously $\tilde{\lambda}_k \sim \lambda_k + g_k'(h)$, thus

$$\tilde{\lambda}_k \sim N(g_k'(h), 1)$$

As $g_k'(h) \to \infty$ we clearly have

$$\|\lambda_k - \tilde{\lambda}_k\|_{TV} \xrightarrow[k\to\infty]{} 2$$

We thus obtain by Lemma 4.21

$$\begin{aligned}\|\mu_h - \mu\|_{TV} &\geq \|(g_k')^*(\mu_h) - (g_k')^*(\mu)\|_{TV} \\ &= \|\lambda_k - \tilde{\lambda}_k\|_{TV} \to 2 \quad \text{as } k\to\infty.\end{aligned}$$

Hence $\|\mu_h - \mu\|_{TV} = 2$ and we may conclude by Lemma 4.20.

(ii) Now $h \in H_\mu$, and we denote $\varphi = R_\mu^{-1} h \in E_\mu'$.

By Lemma 4.10 we have

(24) $$\widehat{\mu_h}(g') = \exp\left(i g'(h) - \tfrac{1}{2} g' R_\mu g'\right), \quad g' \in E'$$

It is thus enough to show that

(25) $$\widehat{S_h(x)\mu}(g') = \widehat{\mu_h}(g') \quad \forall g' \in E',$$

since the Fourier transform determines a measure on E uniquely.

Here $g_h(x)\mu(dx)$ defines a measure since now

$$g_h(x) = \exp(\varphi(x) - \tfrac{1}{2}\|h\|^2_{H_\mu}) \in L^1(d\mu)$$

by Lemma 4.23.

Let us fix elements $g', g'_k \in E'$ and $z \in \mathbb{R}$. We may compute

$$(26)\quad \mathbb{E}\exp(i(g'(x) + z g'_k(x)))$$
$$= \exp(-\tfrac{1}{2}\mathbb{E}|(g' + z g'_k)(x)|^2)$$
$$= \exp(-\tfrac{1}{2}g' R_\mu g' - z g' R_\mu g'_k - \tfrac{z^2}{2} g'_k R_\mu g'_k)$$

We next specify the g'_k:s so that

$$(27)\quad \|g'_k - \varphi\|_{E'_\mu} = \|g'_k - \varphi\|_{L^2(d\mu)} \xrightarrow[k\to\infty]{} 0.$$

Then

$$(28)\quad g'_k R_\mu g'_k = \|g'_k\|^2_{E'_\mu} \to \|\varphi\|^2_{E'_\mu} = \|h\|^2_{H_\mu}$$

and

$$(29)\quad g' R_\mu g'_k \xrightarrow[k\to\infty]{} g'(h)$$

since $R_\mu g'_k \to h$ in H_μ, and hence in E.

From 27, by moving to a subsequence, we may assume that

$$(30)\quad g'_k(x) \to \varphi(x) \quad \text{for } \mu\text{-a.e } x \in E.$$

By combining (28), (29) and (30) and applying the dominated convergence thm on (26) as $k \to \infty$ we obtain the formula

$$(31) \qquad \int \exp(i g'(x) + iz\varphi(x))\,\mu(dx)$$
$$= \exp\left(-\tfrac{1}{2} g' R_\mu g' - z g'(h) - \tfrac{z^2}{2}\|h\|^2_{H_\mu}\right)$$

for all $g' \in E'$. By Lemma 4.23 it is clear that the integral in (31) is analytic in $|z| \leq 2$, say, and by analytic continuation we may substitute $z = -i$ in (31) and divide by $\exp(\frac{1}{2}\|h\|^2_{H_\mu})$ both sides. By comparing to (24) this exactly gives (25), and the proof is complete. □

__Corollary 4.24__ $H_\mu = \{h \in E \mid \mu \text{ equivalent to } \mu_h\}$ □

__Theorem 4.25__ Let μ be a centered Gaussian measure on (separable) E. Then

(i) $H_\mu = \bigcap_{M \in \mathcal{V}} M$, where

$$\mathcal{V} = \{M : M \in \mathcal{B}(E),\ M \text{ linear subspace with } \mu(M) = 1\}.$$

(ii) if $\dim(H_\mu) = \infty$, then $\mu(H_\mu) = 0$.

__Proof.__ (i) Assume that $M \subset E$, $M \in \mathcal{B}(E)$, is a subspace with $\mu(M) = 1$. Let $h \in H_\mu$. Then $\mu_h(M) = 1$ by Cor. 4.24, especially
$$\mu((M-h) \cap M) = 1,$$
whence $(M-h) \cap M \neq \emptyset$, which implies that $h \in M$. We deduce that
$$H_\mu \subset \bigcap_{M \in \mathcal{V}} M.$$

For the other direction, assume that $h \notin H_\mu$. We may pick $g_k' \in E'$ so that for all $k \geq 1$

(a) $\mathbb{E}\, |g_k'(x)|^2 = g_k' R_\mu g_k' = 1$ and

(b) $g_k'(h) \geq k$.

Consider the linear space

$$M_0 := \{ x \in E : \text{the series } \sum_{k=1}^{\infty} \frac{g_k'(x)}{k^2} \text{ converges} \}$$

Since g_k':s are continuous, one easily checks (exercise) that $L \in B(E)$. By condition (b) $h \notin M_0$. On the other hand,

$$\int_E \Big(\sum_{k=1}^{\infty} k^{-2} |g_k'(x)| \Big) \mu(dx)$$

$$= \sum_{k=1}^{\infty} k^{-2} \int_E |g_k'(x)| \mu(dx) \leq \sum_{k=1}^{\infty} k^{2} \Big(\int_E |g(x)|^2 \mu(dx) \Big)^{1/2}$$

$$\overset{(a)}{=} \sum_{k=1}^{\infty} k^{-2} < \infty,$$

whence the series $\sum_{k=1}^{\infty} \frac{g_k'(x)}{k^2}$ converges

absolutely for μ-a.e. $x \in E$, and $\mu(M_0) = 1$. Hence $\bigcap_{M \in \mathcal{V}} M \subset H_\mu$.

(ii) Let $(g_n')_{n \geq 1}$ be an ON-basis of E_μ'. Then, it holds that

$$(32) \qquad h \in H_\mu \Rightarrow \sum_{k=1}^{\infty} |g_k'(h)|^2 < \infty$$

since by setting $g' := (R_\mu)^{-1} h \in E_\mu'$ we have

$\langle g_k', g\rangle_{E_\mu} = g_k'(R_\mu g') = g_k'(h)$, whence

$$\sum_{k=1}^{\infty} |g_k'(h)|^2 = \|g'\|_{E_\mu}^2 < \infty.$$

On the other hand, as random variables on $(E, \mathcal{B}(E), \mu)$, we have that g_k':s are i.i.d. and $g_k' \sim N(0,1)$. Corollary 1.7 verifies that

$$\sum_{k=1}^{\infty} |g_k'(x)|^2 = \infty \quad \text{for } \mu\text{-a.e. } x \in E.$$

This together with (32) yields that $\mu(H_\mu) = 0$. $\square$

Lemma 4.26. If μ is a probability measure on a separable Banach space E, there is a unique smallest closed set $S \subset E$ so that $\mu(S) = 1$ (called the topological support of μ).

Proof. Let $A \subset E$ be countable and dense. Let

$$U := \bigcup \{ \underbrace{B(a, \tfrac{1}{n})}_{\text{open ball}} \mid a \in A,\ n \in \mathbb{N},\ \mu(B(a, \tfrac{1}{n})) = 0 \}.$$

Then U is open, and clearly $\mu(U) = 0$.
Set $S := E \setminus U$, so that $\mu(S) = 1$.
Assume that $S' \subset E$ is closed with $\mu(S') = 1$. We want to show that $S \subset S'$. If not, there is $x_0 \in S$ so that $\varepsilon_0 = d(x_0, S') > 0$. Pick $n_0 \in \mathbb{N}$ so that $1/n_0 < \varepsilon_0/2$ and $a_0 \in A$ with $|a_0 - x_0| < \frac{1}{2 n_0} < \frac{\varepsilon_0}{4}$. Then

$$x_0 \in B(a_0, \tfrac{1}{n_0}) \quad \text{and} \quad B(a_0, \tfrac{1}{n_0}) \cap S' = \emptyset,$$

whence $\mu(B(a_0, \frac{1}{n_0})) = 0$. Thus $x_0 \notin S$, which is a contradiction. $\square$

Exercise: In the case of a general topological space E what could go wrong if one simply defines the support as the set

$$S := \bigcap_{\substack{A \subset E \text{ closed} \\ \mu(A)=1}} A \qquad ?$$

Theorem 4.27 Let μ be a Gaussian measure on a separable metric space E. Then the topological support of μ is the affine subspace

$$m_\mu + \overline{H_\mu}^E.$$

Proof. May assume $m_\mu = 0$. Denote $L = \overline{H_\mu}^E$. By Corollary 4.19 $\mu(L) = 1$, so it is enough to show that for any $x_0 \in L$ and $\varepsilon_0 > 0$ one has

$$\mu(\underbrace{B(x_0, \varepsilon_0)}_{\text{open ball}}) > 0.$$

Pick $h \in H_\mu$ with $|h - x_0| < \varepsilon_0/4$.
Then $\mu(B(x_0 - h, \varepsilon_0)) = 0$ by Thm 4.22 (ii)
Since $B(0, \varepsilon_0/2) \subset B(x_0 - h, \varepsilon_0)$ we also have

$$\mu(B(0, \varepsilon_0/2)) = 0.$$

Choose a dense and countable set $\{a_k\}_{k\geq 1}$ from E_μ (dense in the norm of E). Then

$$\mu(B(a_k, \varepsilon_0/2)) = 0 \quad \forall k$$

and $\quad L \subset \bigcup_{k=1}^{\infty} B(a_k, \varepsilon_0/2) \quad$ so that $\mu(L) = 0$,

which is a contradiction.

Our last general result on the Cameron-Martin space notes that it is essentially independent of the Banach space E.

Theorem 4.28 Let $E \subset F$ with continuous embedding (both E and F are separable Banach spaces), so that $\|x\|_F \leq c\|x\|_E$ for $x \in E$. If E is a Borel subset of F, then for every Gaussian measure μ on E the Cameron-Martin space of μ w.r.t. E is the same as w.r.t. F.

(Sketch)

Proof. We may assume that E is dense in F (otherwise we may replace F by $\overline{E}^F$). Then $F' \subset E'$ in a natural way. Also we note that μ is thus a Gaussian measure also on F, as the imbedding $E \subset F$ is continuous. As $F' \subset E'$, we see that $F'_\mu \subset E'_\mu$ isometrically (the norm in $L^2(d\mu)$ does not change if E is replaced by F). Hence (H_μ w.r. to F) $\subset$ (H_μ w.r.t. E) isometrically. Finally, (H_μ w.r. to E) $\subset$ (H_μ w.r.t F) by Corollary 4.24. $\square$

Brownian motion provides us an interesting and important example. Recall that a standard Brownian motion $(B_t)_{t\in[0,1]}$ on $[0,1]$ is the stochastic process with Gaussian (joint distributions) and

$$\begin{cases} \mathbb{E}\, B_t = 0 & (\text{centered BM}) \\ B_0 \equiv 0 & \\ \mathbb{E}\, B_t B_s = \min(t,s) & \text{for } t,s \in [0,1]. \\ t \mapsto B_t(\omega) \text{ is continuous for each } \omega & \end{cases}$$

The last condition enables us to consider (B_t) as a $C([0,1])$-valued random variable.

Lemma 4.29. Let $(B_t)_{t\in[0,1]}$ be a standard BM. Then the map $\omega \mapsto B_{(\cdot)}$ yields a $C[0,1]$ valued Gaussian random variable

Proof. We just need to check that for each $\nu \in M([0,1])$ (finite Borel measure) the r.v.

$$\omega \mapsto \int_0^1 B_t(\omega)\,\nu(dt)$$

is Gaussian (This also yields weak measurability → Borel-measurability since $C([0,1])$ is separable). Choose (exercise) $\nu_n \xrightarrow{w^*} \nu$ with $\|\nu_n\|_{TV} \le C$ and each ν_n finite sum of deltas. The claim holds for ν_n in place of ν, apply Lemma 6.6. $\square$

Theorem 4.30. Let B be a standard Brownian motion on $[0,1]$ considered as a $C[0,1]$-valued r.v. Then the Cameron-Martin space of B equals:

$$H_B = \{ f \in C([0,1]) : f(0)=0 \text{ and } \|f\|_{H_B}^2 = \int_0^1 |\tfrac{d}{dt} f(t)|^2\,dt < \infty$$

weak derivative

<u>Proof.</u> Let us first identify the covariance[*)] R_B. For that end, choose $\lambda \in \mathcal{M}[0,1] = (C[0,1])'$ and compute

$$\lambda R_B \lambda = \mathbb{E} \Big| \int_0^1 B(t)\,\lambda(dt) \Big|^2$$
$$= \mathbb{E} \int_0^1 \int_0^1 B(t) B(u)\,\lambda(dt)\lambda(du)$$
$$= \int_0^1 \int_0^1 \min(t,s)\,\lambda(dt)\lambda(ds).$$

Since $C^\infty[0,1]$ is weak*-dense in $\mathcal{M}(0,1)$, we deduce easily that for any $f \in C[0,1]$

$$\|f\|_{H_B} = \sup \Big\{ \int_0^1 f(t) g(t)\,dt \;\Big|\; g \in C^\infty[0,1] \text{ with } \int_0^1\int_0^1 \min(t,s) g(t) g(s)\,dt\,ds \le 1 \Big\}$$

Let us denote for $g \in C^\infty[0,1]$: $G(x) = \int_x^1 g(t)\,dt$.
Then
$$\int_0^1 G(x)^2\,dx = \int_0^1 \Big(\int_x^1 g(t)\,dt \int_x^1 g(u)\,du \Big) dx$$
$$= \int_0^1 \int_0^1 g(t) g(u) \int_0^{\min(t,u)} x\,dx = \int_0^1 g(t) g(u) \min(t,u)\,dt\,du$$
$$= g R_B g.$$

Hence
$$\|f\|_{H_B} = \sup \Big\{ \int_0^1 f(t)\, D G(t)\,dt \;\Big|\; G \in C^\infty[0,1],\ G(1) = 0 \text{ and } \|G\|^2_{L^2[0,1]} \le 1 \Big\}. \tag{19}$$

---

*) Actually, covariance is $\lambda R_B \lambda$, sometimes R_B is called the covariance operator

Especially, if $\|f\|_{H_B} < \infty$, we obtain

$$(20) \quad \infty > \sup \left\{ \int_0^1 F(t)\, DG(t)\, dt \;\middle|\; G \in C_0^\infty(0,1),\ \|G\|^2_{L^2(0,1)} \leq 1 \right\}.$$

If we denote by DF the distributional derivative of F, we have

$$\int_0^1 F(t)\, DG(t)\, dt = - \int_0^1 DF(t)\, G(t)\, dt.$$

Hence Riesz representation theorem together with (20) shows that $DF \in L^2[0,1]$, whence $F \in W^{1,2}[0,1]$. We may thus integrate by parts in (19) in order to write there

$$\int_0^1 F(t)\, DG(t) = - \int_0^1 DF(t)\, G(t)\, dt - G(0)\, F(0).$$

Since $\sup \{ G(0) \mid G \in C^\infty[0,1],\ G(1) = 0,\ \|G\|^2_{L^2[0,1]} \leq 1 \} = \infty$, we deduce that $F(0) = 0$ and (under this necessary condition)

$$(21) \quad \|F\|_{H_B} = \|DF\|_{L^2[0,1]}.$$

Conversely, if $DF \in L^2[0,1]$ and $f(0) = 0$ we obtain $F \in H_B$ with (xxx) from the above computations. $\square$

<u>Example</u> The distribution of Brownian motion on $C[0,1]$ is called the Wiener measure. In an interesting way $H_B = W^{1,2}[0,1] \cap \{f(0) = 0\}$ by the above theorem. Let us e.g. consider

$$\tilde{B}(t) = B(t) + at \qquad (\text{BM with drift}).$$

It is absolutely continuous w.r. to $B(t)$, and, in order to determine the Radon-Nikodym derivative w.r. to the Wiener measure, we observe that

$$\delta_t R_B \delta_1 = \int_{[0,1]^2} \min(u,s)\, \delta_t(du)\, \delta_1(ds)$$
$$= \min(t,1) = t$$

In other words, $R_B\delta_1(t) = t \quad \forall\, t \in [0,1]$.
Thus $R_B(a\delta_1)(t) = at$, whence Thm 4.22(i) yields that

$$\boxed{\frac{d\tilde{B}}{dB}} = \exp\left((a\delta_1)(B(\omega)) - \frac{1}{2}\|D(at)\|^2_{L^2[0,1]}\right)$$
$$= \boxed{e^{aB(1) - \frac{1}{2}a^2}\, dB}$$

(Girsanov Formula for the drift).

Remark: Basically all the results of this section are valid for Radon Gaussian measures on locally convex topological vector spaces with some good properties. Especially, if the topological vector space is Souslin, Gaussian measures are automatically Radon and weak measurability implies strong one. For 'applications' one may note that the space of distributions on $\mathbb{R}^d$ (or the Schwartz distributions) are Souslin:

$$\mathcal{D}'(\Omega)\text{- or } \mathcal{S}'(\mathbb{R}^d)\text{-}$$

valued Gaussians have good properties.

5. Gaussian measures on Hilbert spaces

In this section we consider the extra bonus that Hilbert space structure provides for us. The basic difference is that we can now characterize covariances in an easy manner: Gaussian covariances are exactly the positive definite nuclear operators! For later purposes we give two proofs of the necessity, the first one elementary, and a short one that relies on Fernique's theorem.

Recall that the bounded linear operator $A: H \to H$ (H real separable Hilbert space) is <u>nuclear</u> ($\Leftrightarrow$ $\boxed{A \in N(H)}$) if $(A^*A)^{1/2}$ (defined through spectral calculus) is compact and its eigen values $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \to 0$ satisfy

$$\|A\|_N := \sum_{n=1}^{\infty} \lambda_n((A^*A)^{1/2}) < \infty .$$

This makes the nuclear operators a Banach space, and one has for $B, C \in L(H)$

$$\|BAC\|_N \leq \|B\| \|A\|_N \|C\|$$

If $A \in N$, $A \geq 0$ and $A^* = A$ (A symmetric), then $(A^*A)^{1/2} = A$ and we may pick a sequence of eigenvectors g_n of A so that $Ag_n = \lambda_n g_n$, and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \to 0$ as $n \to \infty$,

$$A = \sum_{n \geq 1}^{\infty} \lambda_n g_n \otimes g_n ,$$

$$(1) \qquad \|A\|_N = \sum_{n=1}^{\infty} \lambda_n .$$

(conversely, if (g_n) is ON, $\lambda_n \geq 1$, then (1) is true).

Assume that H is a separable real Hilbert space.
In this situation it is customary to write

$$C = C_X := R_X$$

for the covariance operator of a H-valued Gaussian X with distribution $\mu = \mu_X$.
Via the standard identification $H' = H$, we have

$$Cy = \mathbb{E}\langle y, X - \mathbb{E}X\rangle_H \langle X - \mathbb{E}X\rangle$$

or

$$\langle z, Cy\rangle_H = \mathbb{E}\langle y, X - \mathbb{E}X\rangle_H \langle z, X - \mathbb{E}X\rangle \qquad | \; z, y \in H$$

In the following result we describe the basic properties of covariance operators in the Hilbert space setting.

Theorem 5.1. The bounded linear operator $C : H \to H$ is the covariance operator of a Gaussian H-valued random variable if and only if

C is symmetric, $C \geq 0$, and $C \in N(H)$

Proof. Assume first that $C = C_X$ for a Gaussian H-valued X. May assume that $m_X = 0$. Then, if $y, z \in H$, we obtain by symmetry

$$\langle z, Cy\rangle_H = \mathbb{E}\langle y, X\rangle_H \langle z, X\rangle_H = \langle Cz, y\rangle_H$$

(recall that everything is well-defined and, especially, C is bounded by Fernique's Thm). Moreover

$$\langle z, Cz\rangle_H = \mathbb{E}(\langle z, X\rangle)^2 \geq 0 .$$

Thus C is symmetric and $C \geq 0$. For such operators it is known that

$$\|C\|_N = \sum_{k=1}^{\infty} \langle e_k, Ce_k\rangle \quad \text{for any ON-basis}$$

$(e_k)_{k\geq 1}$ of H. Thus, fixing such a basis

$$\|C\|_N = \sum_{k=1}^{\infty} \langle e_k, Ce_k\rangle = \sum_{k=1}^{\infty} \mathbb{E}(\langle e_k, X\rangle)^2$$
$$= \mathbb{E}\sum_{k=1}^{\infty} (\langle e_k, X\rangle)^2 = \mathbb{E}\|X\|_H^2 < \infty$$

by Fernique's Thm.

Conversely, let $C \geq 0$, $C^* = C$, $\|C\|_N < \infty$. Then we may apply the spectral theory of compact operators to write

$$(1) \qquad Cy = \sum_{k=1}^{\infty} \sigma_k^2 \langle y, h_k\rangle h_k ,$$

where $\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_n^2 \underset{n\to\infty}{\longrightarrow} 0$ are the eigenvalues of C, and h_k:s are corresponding eigenvectors so that $(h_k)_{k\geq 1}$ is an ON-basis of H. Moreover,

$$\sum_{k=1}^{\infty} \sigma_k^2 = \|C\|_N < \infty$$

Let $(X_k)_{k\geq 1}$ be an i.i.d. sequence of $N(0,1)$-variables and set

$$X := \sum_{k=1}^{\infty} \sigma_k X_k h_k .$$

Now
$$\mathbb{E}\|X\|_H^2 = \mathbb{E}\sum_{k=1}^{\infty} \sigma_k^2 X_k^2 = \sum_{k=1}^{\infty} \sigma_k^2 < \infty,$$

so that X is H-valued. Clearly X is Gaussian and (do it!) direct computation shows that $C = C_X$. $\square$

<u>Corollary 5.2</u> Gaussian distributions on H are in 1-1-correspondence to pairs

$$(C_\mu, m_\mu),$$

where $m_\mu \in H$, C_μ satisfies the conditions of

the previous Thm. One says that
$$\mu \sim N(m_\mu, C_\mu).$$

<u>Corollary 5.3</u> (Karhunen-Loewe, version 1)

Let X be a Gaussian random variable with covariance C_X, eigenvalues σ_n^2 of C_X and ON-sequence (h_n) of eigen vectors of C_X. Then

(2) $$X \sim \mu_X + \sum_{k=1}^{\infty} \sigma_k X_k h_k,$$

where $(X_k)_{k\geq 1}$ is an i.i.d-sequence of $N(0,1)$-variables.

<u>Remark.</u> One might wonder what makes Karhunen-Loewe so special as we have our Thm 4.18. The point is that in (1) the vectors h_k are <u>simultaneously</u> orthogonal in H and in H_μ! This is seen by the next result:

<u>Thm 5.4.</u> Let $\mu \sim N(m_\mu, C_\mu)$ be a Gaussian measure on H. Then $H_\mu = C^{1/2}H$ and
$$C^{1/2}: (\operatorname{Ker} C)^\perp \to H_\mu$$
is an isometry. If C has the spectral decomposition (1), then
$$x \in H_\mu \Leftrightarrow x = \sum_{k=1}^{\infty} a_k h_k \text{ with}$$
$$\|x\|_{H_\mu}^2 = \sum_{k=1}^{\infty} \sigma_k^{-2} a_k^2 < \infty$$
(note that $a_k = 0$ for indices k with $\sigma_k = 0$).

Proof. $\langle x, Cx\rangle = \sum_{k=1}^{\infty} x_k^2 \sigma_k^2$ if $x = \sum_{k=1}^{\infty} x_k h_k$.

Hence $\|y\|_{H_\mu}^2 = \sup\{|\langle y, x\rangle| \mid \langle x, Cx\rangle \leq 1\}$,

which easily yields (9). Since

$C^{1/2} x = \sum_{k=1}^{\infty} \sigma_k x_k h_k$, we obtain the

first claim by noting that $(\operatorname{Ker} C)^{\perp} = \overline{\operatorname{span}_{\sigma_k > 0}\{h_k\}}$. $\square$

Remark. All the results of Section 6 for H_μ remain valid.

We shall consider Brownian bridge on $[0,1]$. Brownian bridge is standard Brownian motion $(B_t)_{t\in[0,1]}$ conditioned to have $B_1 = 0$. As $\mathbb{P}(B_1 = 0) = 0$, one has to take care to choose the right definition. In the case of Brownian bridge a convincing argument can be found e.g. by a limiting process for a conditioned random walk. Or, one may observe the following:

Fix for a moment $0 < t_1 < t_2 < \ldots < t_n < 1$. Write

$$\tilde{B} = (B(t_1), \ldots, B(t_n)).$$

We have $\tilde{B} = \tilde{B}_b + \tilde{B}_e$,

where $\tilde{B}_b = (B(t_1) - t_1 B(1), \ldots, B(t_n) - t_n B(1))$

$\tilde{B}_e = (t_1 B(1), \ldots, t_n B(1))$.

Direct computation shows that components of B_b and B_e are orthogonal in $L^2(\Omega)$. Since the vectors are jointly Gaussian, it follows that $\tilde{B}_b$ and $\tilde{B}_e$ are independent.
Also $\sigma(B_e) = \sigma(B(1))$; especially B_b is independent of $B(1)$ and B_e is independent of $B(1)$.
Finally, if $B(1) = 0$ we have $B_e = 0$. Thus it is natural to declare $\boxed{\tilde{B}_b = \text{the Brownian bridge}}$ centered,

Thus $B_b := \tilde{B}_b$ is Gaussian, and we may compute

$$\mathbb{E}\, B_b(t)\, B_b(u) = \mathbb{E}\,(B(t) - tB(1))(B(u) - uB(1))$$
$$= \min(t,u) - tu$$

By the regularity of the covariance one verifies in the same manner as with BM that one may assume realizations continuous on $[0,1]$ (or see next section).

Def. Standard Brownian Bridge B_b is the Gaussian process on $[0,1]$ with $B_b(0) = B_b(1) = 0$ and $\mathbb{E}\, B_b(t)\, B_b(u) = \min(t,u) - tu$, and with continuous realizations.

Of course one may consider B_b as $C([0,1])$-valued random Gaussian.

The "white noise decomposition" (2) (or the one given by Thm 4.12) is often called a Loeve-Karhunen decomposition of X.
Let us determine it for the Brownian bridge.

Let us consider B_b as a $L^2[0,1]$-valued random Gaussian. The covariance operator takes the form

$$CF(t) = \int_0^1 c(t,u) F(u) du,$$

where $c(t,u) := \min(t,u) - tu$. We need to find the eigen-functions and -values of C. Assume $F \in L^2(0,1)$ satisfies

$$CF = \lambda F.$$

We actually know that H_{B_b} (the Cameron-Martin) space consist of elements is $C[0,1]$, so $F \in C[0,1]$ (or check this directly). Then

$$CF(t) = \int_0^t u F(u) du + t \int_t^1 F(u) - t \int_0^1 u F(u)$$

$$= (1-t) \int_0^t u F(u) du + t \int_t^1 (1-u) F(u) du$$

Especially, $CF \in C^1[0,1]$ and

$$\frac{d}{dt} CF(t) = - \int_0^t u F(u) du + (1-t) t F(t) + \int_t^1 (1-u) F(u) du - t(1-t) F(t)$$

$$= \int_t^1 F(u) - \int_0^1 u F(u) du.$$

Thus $CF \in C^2[0,1]$ and

$$\left(\frac{d}{dt}\right)^2 CF(t) = - F(t).$$

We have shown that (observe that always $\boxed{CF(0) = CF(1) = 0}$, and we did not yet use the condition that F is an eigenfunction)

[<u>Lemma 5.5.</u> The covariance operator of B_b is the Greens function for $-(\frac{d}{dx})^2$ on $[0,1]$

<u>Remark:</u> In other words, B_b is the 1-dimensional GFF on $[0,1]$ (compare section 9)!

Thus, if $CF(t) = 0$ for all $t \in [0,1]$ we deduce that $F \equiv 0$, so $0 \notin \sigma(C)$. If $\lambda \neq 0$, we obtain $F \in C^2[0,1]$ and

$$-\lambda \left(\frac{d}{dt}\right)^2 F(t) = f(t) .$$

This has solutions with $f(0) = f(1) = 0$ if and only if $\boxed{\lambda = (n\pi)^{-2}}$, $n \in \mathbb{N}$, and the corresponding (normalized) eigenfunction is

$$f_n(t) = \sqrt{2}\, \sin(n\pi t)$$

Thus, the covariance has the decomposition

$$(3) \quad c(t,u) = \frac{2}{\pi^2} \sum_{n=1}^{\infty} \frac{\sin(n\pi t)\sin(n\pi u)}{n^2}$$

Moreover, the Karhunen-Loeve decomposition for the Brownian bridge is (using (8))

$$(4) \quad B_b(t) = \frac{\sqrt{2}}{\pi} \sum_{n=1}^{\infty} \frac{X_n}{n} \sin(n\pi t) ,$$

where the X_n's are independent $N(0,1)$-variables. By Theorem 6.17 series (4) a.s. converges uniformly.

There are many results on Gaussians on Hilbert spaces that we do not touch here. Let us mention without proof the result which completely characterizes mutual absolute continuity of two Gaussians on H.

Theorem 5.6. (Feldman-Hajek). Let

$$\mu_1 = N(m_1, C_1) \quad \text{and} \quad \mu_2 = N(m_2, C_2)$$

be two Gaussian measures on the Hilbert space H. Then

(1) μ_1 and μ_2 are either singular or equivalent

(2) They are equivalent if and only if the following three conditions hold:

(i) $H_{\mu_1} = H_{\mu_2}$

(ii) $m_1 - m_2 \in H_{\mu_1}$

(iii) $(C_1^{-1/2} C_2^{1/2})(C_1^{-1/2} C_2^{1/2})^* - Id$

is a Hilbert-Schmidt operator on $\overline{H_{\mu_1}}^H$.

Note that case $C_1 = C_2$ follows from our Thm 4.22. It also implies the necessity of (i) and (ii) in (2).

6. Regularity of pointwise-defined Gaussian Fields

In this section we shall briefly consider the most important results on regularity of Gaussian Fields, that are defined pointwise, i.e. as a collection of (jointly Gaussian, centered (for simplicity)) variables

$$(X_t)_{t \in T},$$

where T is a topological space. We shall always assume that X is as above, and

(1) T is a compact metric space $T \equiv (T,d)$

Typically $T = [0,1]$, $[0,1]^n$ or the torus. In the literature one often handles possible measurability problems by assuming e.g. a separability condition (see Doob: 'Stochastic processes'). We do not do it, since below in this section we are interested basically in the situation where X_t can be modified to have a.s. continuous realizations, and the problem does not occur since a countable dense set of t's in T determines the process.

We shall denote

$$C(t,u) = \mathbb{E}\, X(t) X(u), \quad t,u \in T.$$

A standing assumption is:

(2) $\quad C \in C(T \times T)$

We shall denote by d_X the (pseudo) metric induced by

$$d_X(s,t) = \left(\mathbb{E}\,|X(t)-X(s)|^2\right)^{1/2}$$
$$= \left(C(t,t) + C(s,s) - 2C(s,t)\right)^{1/2}$$

If needed, one may assume that $d_X(s,t) \neq 0$ for $s \neq t$ - if this were not the case one could identify the points s,t for which $d_X(s,t)=0$.

<u>Example</u> If one identifies the points $0,1$ on $[0,1]$, one obtains d_X is a metric for B_b (the Brownian bridge)

<u>Remark.</u> In case $d_X(s,t) > 0$ for $s \neq t$ the compactness of T, (2) and a simple argument show that d_X and d yield the same topology on T. Hence, existence of version of $(X_t)_{t \in T}$ with continuous realizations does not care whether we use d_X or d!

Next our aim is to prove Dudley's theorem (a la Talagrand) that gives a very sharp sufficient condition for the continuity of our field $(X_t)_{t \in T}$. This condition is also necessary in the stationary case.

closed ball

Def. Let $B_{d_X}(t,a) = \{ s \in T : d_X(s,t) \leq a \}$.

Let $N(T, d_X, \varepsilon) = N(\varepsilon)$ be the smallest number of balls of d_X-radius ε that cover T. Then N is the entropy function and

$$H := \log N$$

is the log-entropy function of X (or (X,T)).

Example Let $B^a(t)$ be the Fractional Brownian motion (FBM) on $[0,1]$ with Hurst-index a. Then

$$d_X(s,t) = c|s-t|^a \quad ,$$

so d_X is a 'snowflake' of the standard metric and we obtain

$$H(\varepsilon) = \log N(\varepsilon) \sim \log(\varepsilon^{-\frac{1}{a}}) \sim a^{-1}\log(\tfrac{1}{\varepsilon}) .$$

The following two Theorems are essentially due to Dudley. Fix a countable $D \subset T$ with $\overline{D} = T$, and fix $t_0 \in D$.

Theorem 6.1. $\mathbb{E} \sup_{t \in D} |X_t - X_{t_0}| \leq c \int_0^{\mathrm{diam}(T)} \sqrt{H(\varepsilon)}\, d\varepsilon$

Define the modulus of continuity:

$$\omega_X(\delta) := \sup_{\substack{d_X(t,u) \leq \delta \\ t,u \in D}} |X(t) - X(u)|$$

Theorem 6.2. There exist random $\eta : \Omega \to (0,\infty)$ so that a.s. $\omega_X(\delta) \leq c \int_0^{\delta} \sqrt{H(\varepsilon)}\, d\varepsilon$ for $\delta < \eta$

the constant c above is universal

Corollary 6.3. If $\int_0^{\mathrm{diam}(T)} \sqrt{H(\varepsilon)}\,d\varepsilon$ converges, then there is a modification of X with continuous realizations.

Proof. Thm 6.2 $\Rightarrow$ $w_X(\delta) \underset{\delta\to 0}{\longrightarrow} 0$ a.s.

Thus a.s. $X|_D$ is uniformly continuous. Extend by this to an a.s. continuous process $\bar{X}$ on T. This has still the right covariance structure as is seen by continuity of $t \to \bar{X}$ and of the covariance C. $\square$

Proof of Thm 6.1 Fix $t_0 \in T$. We shall estimate $\sup_{t\in D} |X_t - X_{t_0}|$

The claim is dilation invariant (observe that scaling in d_X is just multiplying each X_t by a constant!). Then, we may assume $\mathrm{diam}_{d_X}(T) = 1$. Set $N_0 = 1$ and

$$N_j = N(2^{-j}a) = \exp(H(2^{-j}a)), \quad j = 1, 2, \ldots$$

Here $a \in (0,1)$ is arbitrary. Choose for each $j \geq 1$ a finite subset $A_j \subset D$ so that

$$\inf_{s\in A_j} d_X(s,t) \leq 2^{-j} \quad \forall t \in T.$$

$D \subset T$ countable, $\bar{D} = T$

We may assume that $\#(A_j) \leq N_j$.
Set $A_0 = \{t_0\}$. Define

$\pi_j : T \to A_j$ inductively so that for $j \geq 0$

$$\sup d_X(t, \pi_j(t)) \leq 2^{-j} \quad \forall t \in T.$$

We claim that almost surely for a <u>fixed</u> $t \in T$

$$(1) \quad X_t - X_{t_0} = \sum_{j=1}^{\infty} X_{\pi_j(t)} - X_{\pi_{j-1}(t)}$$ *)

Namely, denote as usual $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-x^2/2} dx$, recall Lemma 4.7, and observe that

$$(2) \qquad d_X(\pi_j(t), \pi_{j-1}(t)) \le 2^{1-j} 2^{-j} \le 2^{2-j},$$

whence

$$\mathbb{P}(|X_{\pi_j(t)} - X_{\pi_{j-1}(t)}| \ge 2^{-j/2})$$

$$\le 2\Phi\left(\frac{2^{-j/2}}{2^{2-j}}\right) = 2\,\Phi(2^{j/2-2}) \le 4 e^{-2^{j-5}}$$

As $\sum_{j=1}^{\infty} \exp(e^{-2^{j-5}}) < \infty$, Borel-Cantelli shows that the series in (1) converges a.s. Moreover,

$$\mathbb{E}\left|\sum_{j=1}^{n} (X_{\pi_j(t)} - X_{\pi_{j-1}(t)}) - (X(t) - X(t_0))\right|^2$$

$$= \mathbb{E}|X_{\pi_n(t)} - X_t|^2 = |d_X(\pi_n(t), t)|^2 \xrightarrow[n\to\infty]{} 0.$$

Hence the a.s. limit of the series in (1) must coincide with $X_t - X_{t_0}$.

Denote next $\qquad M_j = N_j N_{j-1}$,

$$a_j := 2^{5/2-j} \sqrt{\log(2^j M_j)} \qquad (\ge 2^{2-j}),$$

and $\qquad S := \sum_{j=1}^{\infty} a_j.$

Assume $S < \infty$.
Since $M_j = \# \{(a, a') \mid a \in A_j, a' \in A_{j-1}\}$, we obtain the crude estimate for $u \ge 1$
(there are at most M_j different random variables in the set $X_{\pi_j(t)} - X_{\pi_{j-1}(t)}$ and use (2))

---

*) Thus, (1) holds a.s. for any $t \in D$!

$$\mathbb{P}(\| X_{\pi_j(t)} - X_{\pi_{j-1}(t)} \|_{\ell^\infty(D)} > u a_j)$$

$$\leq M_j \; 2 \, \phi\left(\frac{u a_j}{2^{2-j}} \right)$$

$$\leq 2 M_j \frac{2^{2-j}}{u a_j \sqrt{2\pi}} \exp\left(\frac{-u^2 a_j^2}{32 \; 2^{-2j}} \right)$$

$$\leq \quad M_j (2^j M_j)^{-u^2} \leq 2^{-j u^2} \qquad | \, u \geq 1$$

By summing up over $j \geq 1$ we obtain

(3) $$\mathbb{P}(\| X_t - X_{t_0} \|_{\ell^\infty(D)} > u S) \leq \sum_{j=1}^{\infty} 2^{-j u^2} \leq 2 \cdot 2^{-u^2}$$

Hence

$$\mathbb{E} \sup_{t \in D} | X_t - X_{t_0} | = \int_0^\infty \mathbb{P}(\sup_{t \in D} | X_t - X_{t_0} | \geq u) \, du$$

$$\leq S \int_0^\infty \mathbb{P}(\sup_{t \in D} | X_t - X_{t_0} | > S u') du' \leq S + S \int_1^\infty 2^{1-u'^2} du'$$

$$\leq 3 S .$$

Thus it remains to estimate S in terms of the entropy. Since $\sqrt{ab} \leq \sqrt{a} + \sqrt{b}$ we obtain

$$S \leq 2^{5/2} \sum_{j=j_0}^{\infty} 2^{-j} \sqrt{j + 2 \log N_j}$$

$\operatorname{diam}_{d_X}(T') = 1 \Rightarrow N_j \geq 2$ for $j \geq 2$?

$$\leq 10 \Big(\underbrace{\sum_{j=1}^{\infty} 2^{-j} \sqrt{j}}_{\leq 2} + \underbrace{\sum_{j=1}^{\infty} 2^{-j} \sqrt{\log N_j}}_{\geq \sqrt{\log 2}} \Big)$$

$$\leq 40 \sum_{j=1}^{\infty} 2^{-j} \sqrt{\log N_j}$$

$$\leq 80 \sum_{j=1}^{\infty} \int_{2^{-j-1}}^{2^{-j}} \sqrt{H(2^{-j} a)} \, d\varepsilon$$

$$\leq 80 \int_0^\infty \sqrt{H(\varepsilon a)} \, d\varepsilon = \frac{80}{a} \int_0^\infty \sqrt{H(\varepsilon)} \, d\varepsilon$$

As $a \le 1$ is arbitrary, we obtain $S \le 80 \int_0^\infty \sqrt{H(\varepsilon)}\,d\varepsilon$ □

<u>Remark.</u> Observe that (3) actually yields good bounds even for the exponential moments $\mathbb{E} \exp(c|X_t - X_{t_0}|^2_{\ell^\infty(D)})$ assuming $c < S^2$?.

<u>Proof of Thm 6.2</u> We apply the same notations as in the proof of Thm 8.1. As before, since there are $\le N_j^2$ different pairs among $f_{\pi_j(t)} - f_{\pi_j(s)}$, $s, t \in T'$, we obtain that

$$\mathbb{P}\Big(\exists s,t \in T' : |X_{\pi_j(t)} - X_{\pi_j(s)}| \ge \sqrt{2}\, d_X(\pi_j(t), \pi_j(s)) \underbrace{\sqrt{\log(2^j N_j^2)}}_{\ge 1 \text{ for } j \ge 1}\Big)$$

$$\le N_j^2 \exp\Big(-\frac{(\sqrt{2})^2 \log(2^j N_j^2)}{2}\Big)$$
$$= 2^{-j}$$

Hence, by Borel-Cantelli there exist a random index $j_0 \ge 1$ such that

(4) $$j \ge j_0 \Rightarrow |X_{\pi_j(t)} - X_{\pi_j(s)}| \le \sqrt{2}\, d_X(\pi_j(t), \pi_j(s)) \sqrt{\log(2^j N_j^2)}$$

Exactly in a similar manner we obtain a random $j_0' \ge 1$ such that (obs. $M_j \le N_j^2$)

(5) $$j \ge j_0' \Rightarrow |X_{\pi_j(t)} - X_{\pi_{j-1}(t)}| \le \sqrt{2}\, d_X(\pi_j(t), \pi_{j-1}(t)) \sqrt{\log(2^j N_j^2)}$$

By replacing j_0 by $\max(j_0, j_0')$ we may assume $j_0 = j_0'$. Define

$$\eta := 2^{-j_0 - 1}$$

If then $s, t \in D$ are arbitrary such that

$$d_X(s,t) \le \eta,$$

we pick $j_1 > j_0$ so that

$$2^{-j_1-1} < d_X(s,t) \le 2^{-j_1},$$

and observe that

$$d_X(\pi_{j_1}(s), \pi_{j_1}(t)) \le d_X(s, \pi_{j_1}(s)) + d_X(s,t) + d_X(t, \pi_{j_1}(t)) \le 3\, 2^{-j_1}$$

We may now combine (2), (4) and (5) to estimate a.s.

$$|X_t - X_s| \le |X_{\pi_{j_1}(t)} - X_{\pi_{j_1}(s)}| + \left|\sum_{j=j_1+1}^{\infty} X_{\pi_j(t)} - X_{\pi_{j-1}(t)}\right| + \left|\sum_{j=j_1+1}^{\infty} X_{\pi_j(s)} - X_{\pi_{j-1}(s)}\right|$$

$$\le \sqrt{2}\cdot 3\, 2^{-j_1} \sqrt{\log(2^{j_1} N_{j_1}^2)} + 2 \sum_{j=j_1+1}^{\infty} \sqrt{2}\cdot 2^{1-j} \sqrt{\log(2^j N_j^2)}$$

$$\le c \sum_{j=j_1+1}^{\infty} 2^{-j} \sqrt{\log(2^j N_j^2)} .$$

This latter sum is converted to an integral just as in the proof of Thm 8.1. □

Next we look at some corollaries, that give an idea what these results yield in practise.

APPENDIX

A. The Bochner integral

Standing assumption: E a Banach space, $(\Omega, \mathcal{F}, \mu)$ 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
$$x(\omega) = \sum_{k=1}^{n} a_k \chi_{A_k},$$
where $n \geq 1$, $a_k \in E$, $A_k \in \mathcal{F}$ 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. $X^{-1}(B) \in \mathcal{F}$ for every open $B \Leftrightarrow X^{-1}(B) \in \mathcal{F}$ 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 $X^{-1}(V) \in \mathcal{F}$ as X_m:s are clearly measurable.

Assume then that $X(\Omega)$ is separable and $X : (\Omega, \mathcal{F}) \to (E, \mathcal{B}(E))$ 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 $\omega \in \Omega$. $\square$

Theorem A 2. (Pettis measurability thm)

$X : \Omega \to \mathbb{E}$ 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 $\overline{X}^{-1}(\overline{B(0,r)}) = \bigcap_{k=1}^{\infty} \{|e'_k(X(\omega))| \leq r\} \in F$

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. $\square$

Definition A strongly measurable function $X : \Omega \to E$ is Bochner-integrable 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
$$\int_\Omega X(\omega) \, d\mu(\omega) := \lim_{n \to \infty} \int_\Omega X_n(\omega) \, d\mu(\omega),$$
where for the simple function $X_n(\omega) := \sum_{k=1}^{m(n)} a_{n,k} \chi_{A_{n,k}}$ one defines
$$\int_\Omega X_n(\omega) \, d\mu(\omega) := \sum_{k=1}^{m(n)} a_{n,k} \, \mu(A_{n,k}).$$

Theorem A.3. (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
$$e'\left(\int_\Omega X(\omega) \, d\mu(\omega)\right) = \int_\Omega e'(X(\omega)) \, d\mu(\omega) \quad \forall \, e' \in E'$$

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

Proof (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

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

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

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

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

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

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

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

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

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

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

as soon as $n \geq n_0(\varepsilon)$, $m \geq n_0(\varepsilon)$ (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 $\sup_{n\geq 1} \int_\Omega \|X_n\|\,d\mu \leq \int_\Omega \|X_1\|\,d\mu + \sup_{n\geq 1} \int_\Omega \|X_n - X_1\|\,d\mu < \infty$. Hence by Fatou's lemma

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

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 $X_n(\omega) \underset{n\to\infty}{\to} X(\omega)$ 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 $\tilde{X}_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)$ 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 $X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)$

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 $\mathbb{X} \in L^1_{\mathbb{X}}(d\mu)$ if $\mathbb{X}$ is Bochner integrable. More generally,

  $F \in L^p_{\mathbb{X}}(d\mu)$ if F is strongly measurable and

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