1

Euclidean Sections

I Introduction

In this chapter (Section IV), a fundamental theorem of structure of Banach spaces is proved: *Dvoretzky's theorem*. This result is local in nature, stating that every space E of dimension n contains *large* subspaces F (of dimension of the order of $\log n$) that are *almost Hilbertian*: the Banach–Mazur distance d_F between F and $\ell_2^{\dim F}$ is close to 1. The proof uses both compactness (the Dvoretzky–Rogers and Lewis lemmas) and a probabilistic argument.

Two proofs of this theorem are presented. The first, due to Gordon [1985], is valid only in the real case, but can be well adapted to give an isomorphic version to be presented later; it is based on the comparison of Gaussian vectors (the Slepian–Gordon theorem). The second is valid in both cases, and relies on an inequality of concentration of measure due to Maurey and Pisier (see Pisier [1986 b] or Pisier 2). Probability plays an important role in these proofs, so, in Sections II and III, the required probabilistic tools are developed.

The chapter finishes with another theorem of structure, the Lindenstrauss–Tzafriri theorem (Section V), of global nature this time, stating that every Banach space, with all its closed subspaces complemented, is isomorphic to a Hilbert space. The proof essentially relies on Dvoretzky's theorem, plus an argument of compactness (ultraproducts) to pass from local to global nature.

II An Inequality of Concentration of Measure

The phenomenon of *concentration of measure*, dear to V. Milman, turns out to be crucial in the proof of Dvoretzky's theorem, as Milman discovered (Milman [1971 b]). He used an isoperimetric inequality on the Euclidean sphere due to Paul Lévy. Here, an alternative, simpler inequality of this type is presented,

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-107-16262-4 — Introduction to Banach Spaces: Analysis and Probability
Daniel Li , Hervé Queffélec , Translated by Danièle Gibbons , Greg Gibbons
Excerpt
More Information

due to Maurey and Pisier (see Pisier [1986 b] or Pisier 2). Beforehand, a few complements on Gaussian variables are provided, to be used in Section III.

II.1 Asymptotic Behavior of Gaussian Variables

Recall that a standard Gaussian (always denoted g) is a real random variable g with density $(2\pi)^{-1/2}\mathrm{e}^{-x^2/2}$; this variable is not bounded, but everything turns out as if it were "almost" bounded, as the following proposition, the analytic version of the Gaussian "bell curves", shows:

Proposition II.1 *Let g be a standard Gaussian on a space $(\Omega, \mathcal{A}, \mathbb{P})$. Then:*

1) $\mathbb{P}(|g| > t) \leqslant \mathrm{e}^{-t^2/2}$ *for any* $t > 0$;
2) $\mathbb{P}(|g| > t) \sim \sqrt{\frac{2}{\pi}}\, t^{-1}\mathrm{e}^{-t^2/2}$ *when* $t \to +\infty$;
3) *for any* $\delta > 0$, *there exists* $\alpha = \alpha(\delta) > 0$ *such that* $\mathbb{P}(g > t) \geqslant \dfrac{\alpha}{t}\,\mathrm{e}^{-t^2/2}$ *when* $t \geqslant \delta$.

Proof

1) We have:

$$\begin{aligned}\mathbb{P}(g > t) &= \frac{1}{\sqrt{2\pi}}\int_t^{+\infty} \mathrm{e}^{-x^2/2}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_0^{+\infty} \mathrm{e}^{-(x+t)^2/2}\,\mathrm{d}x\\ &= \frac{\mathrm{e}^{-t^2/2}}{\sqrt{2\pi}}\int_0^{+\infty} \mathrm{e}^{-tx}\,\mathrm{e}^{-x^2/2}\,\mathrm{d}x \leqslant \frac{\mathrm{e}^{-t^2/2}}{\sqrt{2\pi}}\int_0^{+\infty}\mathrm{e}^{-x^2/2}\,\mathrm{d}x\\ &= \frac{1}{2}\,\mathrm{e}^{-t^2/2};\end{aligned}$$

then $\mathbb{P}(|g| > t) = 2\,\mathbb{P}(g > t) \leqslant \mathrm{e}^{-t^2/2}$.

2) It follows from the calculation in 1) that:

$$\mathbb{P}(g > t) = \frac{\mathrm{e}^{-t^2/2}}{\sqrt{2\pi}\,t}\int_0^{+\infty}\mathrm{e}^{-y}\,\mathrm{e}^{-y^2/2t^2}\,\mathrm{d}y,$$

a quantity equivalent to:

$$\frac{\mathrm{e}^{-t^2/2}}{\sqrt{2\pi}\,t}\int_0^{+\infty}\mathrm{e}^{-y}\,\mathrm{d}y = \frac{\mathrm{e}^{-t^2/2}}{\sqrt{2\pi}\,t} \qquad \text{when } t \to +\infty.$$

3) The function $t\,\mathrm{e}^{t^2/2}\,\mathbb{P}(g > t)$ is continuous, > 0 on $[\delta, +\infty[$, and tends to $(2\pi)^{-1/2}$ as $t \to +\infty$ by 2), hence the result. □

When we consider n independent copies $g_1, \ldots, g_n$ of a standard Gaussian g, Proposition II.1 leads to the following bounds:

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-16262-4 — Introduction to Banach Spaces: Analysis and Probability
Daniel Li , Hervé Queffélec , Translated by Danièle Gibbons , Greg Gibbons
Excerpt
More Information

Proposition II.2 *Let $g_1, \dots, g_n$ be n independent copies of a standard real Gaussian g, and for $n = 1, 2, \dots$, consider the two maximal functions $M_n = \max(g_1, \dots, g_n)$ and $M_n^* = \max(|g_1|, \dots, |g_n|)$. Then:*

1) $\mathbb{P}\big(M_n > \sqrt{\log n}\big) \geqslant 1 - \varepsilon_n$, *with* $0 < \varepsilon_n < 1$ *and* $\sum_{n=1}^{+\infty} \varepsilon_n < +\infty$;
2) $C_1 \sqrt{\log n} \leqslant \mathbb{E}(M_n) \leqslant \mathbb{E}(M_n^*) \leqslant C_2 \sqrt{\log(n+1)}$, *with C_1 and C_2 positive numerical constants.*

Proof

1) If $n = 1$, the inequality holds with $\varepsilon_1 = 1/2$. If $n \geqslant 2$, we apply 3) of Proposition II.1, with $\delta = \sqrt{\log 2}$ and $t = \sqrt{\log n}$, to obtain:

$$\mathbb{P}(M_n < t) = \big(\mathbb{P}(g_1 < t)\big)^n = \big(1 - \mathbb{P}(g_1 > t)\big)^n \leqslant \mathrm{e}^{-n\mathbb{P}(g_1 > t)}$$
$$\leqslant \exp\left(-n \frac{\alpha}{\sqrt{\log n}} n^{-1/2}\right) = \exp\left(-\alpha \sqrt{\frac{n}{\log n}}\right) = \varepsilon_n,$$

and the sequence $(\varepsilon_n)_{n\geqslant 1}$ thus defined works.

2) The upper bound was proved in Chapter 1 of Volume 1, Corollary IV.4, with $C_2 = \sqrt{8/3}$. For the lower bound, first note that, for n large enough:

$$(*) \qquad \mathbb{E}(M_n^+) \geqslant \delta \sqrt{\log n} \quad \text{and} \quad \mathbb{E}(M_n^-) \leqslant \frac{c}{\sqrt{n}},$$

with δ and c positive constants. In fact,

$$\sum_{n=1}^{+\infty} \mathbb{P}(M_n^+ \leqslant \sqrt{\log n}) \leqslant \sum_{n=1}^{+\infty} \varepsilon_n < +\infty,$$

hence the Borel–Cantelli lemma gives $\liminf_{n\to+\infty} \frac{M_n^+}{\sqrt{\log n}} \geqslant 1$ almost surely, and thus $\liminf_{n\to+\infty} \frac{\mathbb{E}(M_n^+)}{\sqrt{\log n}} \geqslant 1$ by Fatou's lemma. Moreover, by 1) of Proposition II.1:

$$\mathbb{E}(M_n^-) = \int_0^{+\infty} \mathbb{P}(M_n < -t)\,\mathrm{d}t = \int_0^{+\infty} \big(\mathbb{P}(g_1 < -t)\big)^n \mathrm{d}t$$
$$\leqslant \int_0^{+\infty} \big(\mathbb{P}(|g| > t)\big)^n \mathrm{d}t \leqslant \int_0^{+\infty} \mathrm{e}^{-nt^2/2}\,\mathrm{d}t = \frac{c}{\sqrt{n}}.$$

It follows from $(*)$ that:

$$\mathbb{E}(M_n) = \mathbb{E}(M_n^+) - \mathbb{E}(M_n^-) \geqslant \frac{\delta}{2}\sqrt{\log n}$$

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-16262-4 — Introduction to Banach Spaces: Analysis and Probability
Daniel Li , Hervé Queffélec , Translated by Danièle Gibbons , Greg Gibbons
Excerpt
More Information

for $n \geqslant n_0$. To obtain the lower bound of 2), it thus suffices to see that, for $2 \leqslant n < n_0$, $\mathbb{E}(M_n) > 0$. However, for $n \geqslant 2$:

$$\mathbb{E}(M_n) \geqslant \mathbb{E}(M_2) = \frac{1}{2\pi}\, 2 \int_{\mathbb{R}} \left[\int_{x_1 \leqslant x_2} x_2 \mathrm{e}^{-\frac{1}{2}(x_1^2 + x_2^2)}\, \mathrm{d}x_2 \right] \mathrm{d}x_1$$
$$= \frac{1}{\pi} \int_{-\infty}^{+\infty} \mathrm{e}^{-x_1^2/2}\, \mathrm{e}^{-x_1^2/2}\, \mathrm{d}x_1 = \frac{1}{\sqrt{\pi}} > 0. \qquad \square$$

II.2 The Maurey–Pisier Inequality

This inequality is needed for the proof of Dvoretzky's theorem in the complex case. Hence it is stated in this framework, even though $\mathbb{C}$ does not play any particular role.

First, we equip $\mathbb{C}^m$ with its standard Gaussian measure γ, of density:

$$\gamma(z) = \frac{1}{(2\pi)^m} \exp\left(-\frac{1}{2} \sum_{j=1}^{m} |z_j|^2 \right),$$

where $z = (z_1, \ldots, z_m)$. In other words, if we write $z_j = x_j + i\, y_j$, with x_j, $y_j \in \mathbb{R}$, we have:

$$\int_{\mathbb{C}^m} f(z)\, \mathrm{d}\gamma(z) =$$
$$\frac{1}{(2\pi)^m} \int_{\mathbb{R}^{2m}} f(x_1, \ldots, x_m, y_1, \ldots, y_m) \exp\left(-\frac{1}{2} \sum_{j=1}^{m} (x_j^2 + y_j^2) \right) \mathrm{d}x_1 \ldots \mathrm{d}x_m \mathrm{d}y_1 \ldots \mathrm{d}y_m$$

for every function $f : \mathbb{C}^m \to \mathbb{C}$ for which this makes sense. The usual Hermitian norm of $\mathbb{C}^m = \ell_2^m$ is denoted by $\| \,.\, \|_2$.

Theorem II.3 (The Maurey–Pisier Deviation Inequality) *Let $\Phi : \mathbb{C}^m \to \mathbb{R}$ be a σ-Lipschitz function*:

$$|\Phi(z) - \Phi(w)| \leqslant \sigma\, \|z - w\|_2, \qquad \forall\, z, w \in \mathbb{C}^m.$$

If $M = \int_{\mathbb{C}^m} \Phi(z)\, \mathrm{d}\gamma(z)$, then:

$$\gamma\big(|\Phi - M| > t\big) \leqslant 2 \exp\left(-K \frac{t^2}{\sigma^2} \right)$$

for any $t > 0$,with $K > 0$ a numerical constant ($K = 2/\pi^2$ is suitable).

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-107-16262-4 — Introduction to Banach Spaces: Analysis and Probability
Daniel Li , Hervé Queffélec , Translated by Danièle Gibbons , Greg Gibbons
Excerpt
More Information

In particular, this leads to the following corollary:

Theorem II.4 (The Maurey–Pisier Concentration of Measure Inequality) *Let E be a (complex) Banach space, $v_1, \ldots, v_m \in E$, $Z_1, \ldots, Z_m$ independent standard (complex) Gaussians, and $Z = \sum_{j=1}^m Z_j\, v_j$. Set*:

$$\sigma_Z = \sup_{\varphi \in B_{E^*}} \big(\mathbb{E}|\varphi(Z)|^2\big)^{1/2};$$

then:

$$\mathbb{P}\big(\big|\, \|Z\| - \mathbb{E}\|Z\|\, \big| > t\big) \leqslant 2 \exp\left(-K\, \frac{t^2}{\sigma_Z^2}\right)$$

for any $t > 0$.

Proof Take $\Phi(z) = \big\| \sum_{j=1}^m z_j\, v_j \big\|$, for $z = (z_1, \ldots, z_m) \in \mathbb{C}^m$. We have:

$$|\Phi(z) - \Phi(w)| \leqslant \left\| \sum_{j=1}^m (z_j - w_j)\, v_j \right\| \leqslant \sigma_Z\, \|z - w\|_2,$$

since $\sigma_Z = \sup_{a \in B_{\ell_2^m}} \big\| \sum_{j=1}^m a_j\, v_j \big\|$, and the result ensues from Theorem II.3. □

To prove Theorem II.3, we need to establish that Lipschitz functions are differentiable almost everywhere. This is the aim of the following theorem:

Theorem II.5 (The Rademacher Theorem) *Every Lipschitz function $\Phi\colon \mathbb{R}^N \to \mathbb{R}$ is differentiable almost everywhere.*

Proof The case $N = 1$ is assumed well known: every absolutely continuous function $\phi\colon \mathbb{R} \to \mathbb{R}$, in particular every Lipschitz function, is differentiable almost everywhere and $\int_a^b \phi'(t)\, \mathrm{d}t = \phi(b) - \phi(a)$ for any $a, b \in \mathbb{R}$; for this, we refer to Rudin 2, Chapter 7 (Theorem 7.18).

For any $u \in \mathbb{R}^N$ with norm 1, denote by $\partial_u \Phi(x)$ the derivative in the direction u of Φ at $x \in \mathbb{R}^N$, when it exists. Let $\mathcal{N}_u$ be the set of $x \in \mathbb{R}^N$ for which $\partial_u \Phi(x)$ does not exist. We can easily verify that this set is measurable. By applying the single-dimensional case to the function $t \mapsto \Phi(x + tu)$, we obtain, for any $x \in \mathbb{R}^N$, the negligibility of $\mathcal{N}_u \cap (x + \mathbb{R}\, u)$. Then, by Fubini's theorem, the measure of $\mathcal{N}_u$ is null. Hence, for every unitary vector u, $\partial_u \Phi(x)$ exists for almost any $x \in \mathbb{R}^N$.

Now consider the gradient $\nabla \Phi(x) = \big(\partial_1 \Phi(x), \ldots, \partial_N \Phi(x)\big)$, where the $\partial_j \Phi(x)$, $1 \leqslant j \leqslant N$, are the usual partial derivatives: the derivatives in the direction of the vectors e_j of the canonical basis of $\mathbb{R}^N$. We have:

$$\partial_u \Phi(x) = \langle u, \nabla \Phi(x) \rangle$$

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-107-16262-4 — Introduction to Banach Spaces: Analysis and Probability
Daniel Li , Hervé Queffélec , Translated by Danièle Gibbons , Greg Gibbons
Excerpt
More Information

for almost any $x \in \mathbb{R}^N$. In fact, this is well known to be the case for any continuously differentiable function; then, if $\psi\colon \mathbb{R}^N \to \mathbb{R}$ is $\mathcal{C}^1$ smooth and compactly supported, we have, via a change of variables:

$$\begin{aligned}\int_{\mathbb{R}^N} \frac{\Phi(x+hu)-\Phi(x)}{h}\,\psi(x)\,\mathrm{d}x &= -\int_{\mathbb{R}^N} \frac{\psi(x)-\psi(x-hu)}{h}\,\Phi(x)\,\mathrm{d}x \\ &\underset{h\to 0}{\longrightarrow} -\int_{\mathbb{R}^N} \partial_u\psi(x)\,\Phi(x)\,\mathrm{d}x \\ &= -\int_{\mathbb{R}^N} \langle u, \nabla\psi(x)\rangle\Phi(x)\,\mathrm{d}x \\ &= -\sum_{j=1}^N u_j \int_{\mathbb{R}^N} \partial_j\psi(x)\,\Phi(x)\,\mathrm{d}x \\ &= \sum_{j=1}^N u_j \int_{\mathbb{R}^N} \psi(x)\,\partial_j\Phi(x)\,\mathrm{d}x,\end{aligned}$$

by integrating by parts with respect to the j-th variable,

$$= \int_{\mathbb{R}^N} \psi(x)\langle u, \nabla\Phi(x)\rangle \mathrm{d}x,$$

thus the result, since, by the dominated convergence theorem (applicable as Φ is Lipschitz), the first integral tends to $\int_{\mathbb{R}^N} \partial_u\Phi(x)\,\psi(x)\,\mathrm{d}x$.

Now let Δ be a countable dense subset in the unit sphere S of $\mathbb{R}^N$. For each $u \in \Delta$, let A_u be the set of $x \in \mathbb{R}^N$ such that $\nabla\Phi(x)$ and $\partial_u\Phi(x)$ exist and satisfy $\partial_u\Phi(x) = \langle u, \nabla\Phi(x)\rangle$, and let $A = \bigcap_{u\in\Delta} A_u$. By the above, $\mathbb{R}^N \setminus A$ has measure zero. Let us show that Φ is differentiable for every $x \in A$.

Fix $x \in A$, and, for $u \in S$ and $h \neq 0$, set:

$$L_h(u) = \frac{\Phi(x+hu)-\Phi(x)}{h} - \langle u, \nabla\Phi(x)\rangle.$$

It suffices to show that $\lim_{h\to 0} L_h(u) = 0$ uniformly for $u \in S$. Indeed, if C is the Lipschitz constant of Φ, then, for every $u, u' \in S$:

$$|L_h(u) - L_h(u')| \leqslant (N+1)\,C\,\|u-u'\|_2.$$

The set of functions L_h, for $h > 0$, is hence equicontinuous on the compact set S. As it converges to 0 on the dense set Δ, it converges uniformly on S to 0, by Ascoli's theorem, and the proof is thus complete. □

Proof of Theorem II.3 Write $z = x + iy \in \mathbb{C}^m$ with $x, y \in \mathbb{R}^m$, and denote $x = \operatorname{Re} z$ and $y = \operatorname{Im} z$. Also denote:

$$\Phi'_x = \left(\frac{\partial\Phi}{\partial x_1}, \ldots, \frac{\partial\Phi}{\partial x_m}\right) \quad \text{and} \quad \Phi'_y = \left(\frac{\partial\Phi}{\partial y_1}, \ldots, \frac{\partial\Phi}{\partial y_m}\right).$$

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-107-16262-4 — Introduction to Banach Spaces: Analysis and Probability
Daniel Li , Hervé Queffélec , Translated by Danièle Gibbons , Greg Gibbons
Excerpt
More Information

Let $z = x + iy$ and $w = u + iv \in \mathbb{C}^m$; for $0 \leqslant \theta \leqslant 2\pi$, set:

$$z(\theta) = z\,\sin\theta + w\,\cos\theta,$$

so that $z'(\theta) = z\,\cos\theta - w\,\sin\theta$. Since $z(\pi/2) = z$ and $z(0) = w$, we obtain:

$$\begin{aligned}\Phi(z) - \Phi(w) = \Phi\big[z(\pi/2)\big] - \Phi\big[z(0)\big] &= \int_0^{\pi/2} \frac{\mathrm{d}}{\mathrm{d}\theta}\big(\Phi[z(\theta)]\big)\,\mathrm{d}\theta \\ &= \int_0^{\pi/2} \big[\langle \Phi'_x\big(z(\theta)\big), \mathrm{Re}\,z'(\theta)\rangle + \langle \Phi'_y\big(z(\theta)\big), \mathrm{Im}\,z'(\theta)\rangle\big]\,\mathrm{d}\theta,\end{aligned}$$

where $\langle\,\cdot\,,\,\cdot\,\rangle$ is the scalar product in $\mathbb{R}^m$.

For $\lambda \in \mathbb{R}$, we introduce the convex function $\phi_\lambda : t \in \mathbb{R} \mapsto \mathrm{e}^{\lambda t}$. By Jensen's inequality:

$$\begin{aligned}&\phi_\lambda\big(\Phi(z) - \Phi(w)\big) \\ &\quad\leqslant \frac{2}{\pi}\int_0^{\pi/2} \phi_\lambda\left[\frac{\pi}{2}\Big(\langle \Phi'_x\big(z(\theta)\big), \mathrm{Re}\,z'(\theta)\rangle + \langle \Phi'_y\big(z(\theta)\big), \mathrm{Im}\,z'(\theta)\rangle\Big)\right]\mathrm{d}\theta.\end{aligned}$$

The crucial point now is that, for each θ, the Gaussian measure $d\gamma(z)d\gamma(w)$ is invariant under the unitary map $(z, w) \mapsto \big(z(\theta), z'(\theta)\big)$. Integrating the preceding inequality, and using Fubini's theorem, we thus obtain:

$$\begin{aligned}&\iint_{\mathbb{C}^m\times\mathbb{C}^m} \phi_\lambda\big(\Phi(z) - \Phi(w)\big)\,\mathrm{d}\gamma(z)d\gamma(w) \\ &\quad\leqslant \iint_{\mathbb{C}^m\times\mathbb{C}^m} \phi_\lambda\left[\frac{\pi}{2}\big(\langle \Phi'_x(z), \mathrm{Re}\,w\rangle + \langle \Phi'_y(z), \mathrm{Im}\,w\rangle\big)\right]\mathrm{d}\gamma(z)d\gamma(w).\end{aligned}$$

The equality, in which $c \in \mathbb{C}^m$ and $\alpha \in \mathbb{R}$:

$$\int_{\mathbb{C}^m} \exp\,\alpha\,\big(\langle \mathrm{Re}\,c, \mathrm{Re}\,w\rangle + \langle \mathrm{Im}\,c, \mathrm{Im}\,w\rangle\big)\,\mathrm{d}\gamma(w) = \exp\left(\frac{\alpha^2}{2}\,\|c\|_2^2\right),$$

is used here, with $c = \Phi'_x(z) + i\,\Phi'_y(z)$ and $\alpha = \lambda\,\pi/2$, to obtain:

$$\begin{aligned}&\iint_{\mathbb{C}^m\times\mathbb{C}^m} \phi_\lambda\big(\Phi(z) - \Phi(w)\big)\,\mathrm{d}\gamma(z)d\gamma(w) \\ &\qquad\leqslant \int_{\mathbb{C}^m} \exp\left(\frac{\pi^2}{8}\,\lambda^2\,\big(\|\Phi'_x(z)\|_2^2 + \|\Phi'_y(z)\|_2^2\big)\right)\mathrm{d}\gamma(z).\end{aligned}$$

However $\|\Phi'_x(z)\|_2^2 + \|\Phi'_y(z)\|_2^2 \leqslant \sigma^2$ for almost all z, since $\Phi\colon \mathbb{C}^m \to \mathbb{R}$ is σ-Lipschitz; consequently:

$$\iint_{\mathbb{C}^m\times\mathbb{C}^m} \phi_\lambda\big(\Phi(z) - \Phi(w)\big)\,\mathrm{d}\gamma(z)d\gamma(w) \leqslant \exp\left(\frac{\pi^2}{8}\,\lambda^2\sigma^2\right).$$

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-1-107-16262-4 — Introduction to Banach Spaces: Analysis and Probability
Daniel Li , Hervé Queffélec , Translated by Danièle Gibbons , Greg Gibbons
Excerpt
More Information

Again using Jensen's inequality, we obtain, if $M = \int_{\mathbb{C}^m} \Phi(z)\, \mathrm{d}\gamma(z)$:

$$\int_{\mathbb{C}^m} \phi_\lambda\big(\Phi(z) - M\big)\, \mathrm{d}\gamma(z) \leqslant \exp\left(\frac{\pi^2}{8}\lambda^2\sigma^2\right).$$

The rest of the proof is routine: for $\lambda > 0$, Markov's inequality gives:

$$\gamma\big(\Phi - M > t\big) = \gamma\big(\mathrm{e}^{\lambda(\Phi - M)} > \mathrm{e}^{\lambda t}\big) \leqslant \mathrm{e}^{-\lambda t}\int_{\mathbb{C}^m} \mathrm{e}^{\lambda(\Phi(z) - M)}\, \mathrm{d}\gamma(z)$$
$$\leqslant \exp\left(-\lambda t + \frac{\pi^2}{8}\lambda^2\sigma^2\right).$$

We optimize in λ, by taking $\lambda = \dfrac{4}{\pi^2}\dfrac{t}{\sigma^2}$, to obtain:

$$\gamma\big(\Phi - M > t\big) \leqslant \exp\big(-K\, t^2/\sigma^2\big),$$

with $K = 2/\pi^2$. Applying this inequality to $(-\Phi)$, we also have:

$$\gamma\big(\Phi - M < -t\big) \leqslant \exp\big(-K\, t^2/\sigma^2\big),$$

so finally, by addition:

$$\gamma\big(|\Phi - M| > t\big) \leqslant 2\exp\big(-K\, t^2/\sigma^2\big),$$

as claimed. □

Remark By replacing the integration over $[0, \pi/2]$ by Itô's formula for Brownian motion, we obtain the best constant $K = 1/2$ (see Pisier [1986 b]).

III Comparison of Gaussian Vectors

III.1 Statement of the Problem

Let $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ be two centered Gaussian vectors. They can be considered as processes $i \mapsto X_i$ and $i \mapsto Y_i$ indexed by the instants $i = 1, \ldots, n$, and (with no loss of generality) will always be assumed non-degenerate, i.e. possessing a density. The vectors X and Y are determined by their respective covariance matrices $\big(c_{ij}^X\big)_{i,j=1,\ldots,n}$ and $\big(c_{ij}^Y\big)_{i,j=1,\ldots,n}$, where $c_{ij}^X = \mathbb{E}(X_iX_j)$ and $c_{ij}^Y = \mathbb{E}(Y_iY_j)$.

Our goal is the comparison of the expectations $\mathbb{E}[\varphi(X)]$ and $\mathbb{E}[\varphi(Y)]$, where $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a measurable map of moderate growth, meaning that: $|\varphi(x)| \leqslant a\,\mathrm{e}^{b|x|}$, where $|x|$ is the Euclidean norm of $x \in \mathbb{R}^n$. We will use a variational method for this. If:

$$Z_t = \sqrt{1-t}\,X + \sqrt{t}\,Y, \qquad 0 \leqslant t \leqslant 1$$

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE
Cambridge University Press
978-1-107-16262-4 — Introduction to Banach Spaces: Analysis and Probability
Daniel Li , Hervé Queffélec , Translated by Danièle Gibbons , Greg Gibbons
Excerpt
More Information

and $h(t) = \mathbb{E}[\varphi(Z_t)]$, we study the sign of $h'(t)$ (note that $Z_0 = X$ and $Z_1 = Y$). The following notation is useful:

$$\begin{cases} M_{ij} = c_{ij}^Y - c_{ij}^X \\ N_{ij} = \mathbb{E}(Y_i - Y_j)^2 - \mathbb{E}(X_i - X_j)^2 \end{cases}$$

and:

$$\partial_{ij} = \frac{\partial^2}{\partial x_i \partial x_j}, \qquad \partial_t = \frac{\partial}{\partial t}\cdot$$

Note that $N_{ii} = 0$ and that the M_{ij}, N_{ij} are linked by the trivial identity:

$$(*) \qquad M_{ij} = -\frac{1}{2}N_{ij} + \frac{1}{2}(M_{ii} + M_{jj}).$$

III.2 The Comparison Theorem. Applications

Independently of any hypothesis on the variation of φ and the sign of the N_{ij}'s, the following lemma gives a nice expression for $h'(t)$ when X and Y are assumed independent, which can always be done, since this affects neither the hypotheses nor the conclusions of the theorems to follow.

Lemma III.1 *Let X and Y be two independent centered Gaussian vectors in $\mathbb{R}^n$ and $Z_t = \sqrt{1-t}\,X + \sqrt{t}\,Y$ $(0 \leqslant t \leqslant 1)$. For every $\varphi \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ of moderate growth, set $h(t) = \mathbb{E}[\varphi(Z_t)]$. Then, for $0 < t < 1$:*

$$h'(t) = -\frac{1}{4}\mathbb{E}\left(\sum_{i \neq j} N_{ij}\partial_{ij}\varphi(Z_t)\right) + \frac{1}{2}\sum_i \mathbb{E}\left(M_{ii}\sum_j \partial_{ij}\varphi(Z_t)\right).$$

Proof This expression could be obtained as an application of Itô's formula, but we prefer a direct proof. Let $f(t,x)$ be the density of Z_t, and $F(t,u)$ its characteristic function. The function f satisfies the heat equation:

$$(1) \qquad \partial_t f = \frac{1}{2}\sum_{i,j} M_{ij}\partial_{ij} f, \qquad 0 < t < 1.$$

Indeed, the Fourier inversion formula gives:

$$f(t,x) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{i\langle x,u\rangle} F(t,u)\,du,$$

so that:

$$\partial_t f(t,x) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{i\langle x,u\rangle}\,\partial_t F(t,u)\,du.$$

© in this web service Cambridge University Press www.cambridge.org

Cambridge University Press
978-1-107-16262-4 — Introduction to Banach Spaces: Analysis and Probability
Daniel Li , Hervé Queffélec , Translated by Danièle Gibbons , Greg Gibbons
Excerpt
More Information

Calculate $F(t,u)$:

$$F(t,u) = \exp\left(-\frac{1}{2}\sum_{i,j} u_i u_j\, c_{ij}^{Z_t}\right),$$

with (X and Y being independent):

$$c_{ij}^{Z_t} = \mathbb{E}\big[(\sqrt{1-t}\,X_i + \sqrt{t}\,Y_i)(\sqrt{1-t}\,X_j + \sqrt{t}\,Y_j)\big] = (1-t)\,c_{ij}^X + t\,c_{ij}^Y;$$

hence:

$$F(t,u) = \exp\left[-\frac{1}{2}\sum_{i,j} u_i u_j\left((1-t)\,c_{ij}^X + t\,c_{ij}^Y\right)\right],$$

and:

$$\partial_t F = -\frac{1}{2}\sum_{i,j} u_i u_j M_{ij} F.$$

We thus obtain the relation:

$$\partial_t f(t,x) = -\frac{1}{2}\sum_{i,j} M_{ij}\,(2\pi)^{-n}\int_{\mathbb{R}^n} \mathrm{e}^{i\langle x,u\rangle} u_i u_j\, F(t,u)\,\mathrm{d}u.$$

However, by differentiating under the integral, we also have:

$$\partial_{ij} f(t,x) = -(2\pi)^{-n}\int_{\mathbb{R}^n} u_i u_j\, \mathrm{e}^{i\langle x,u\rangle}\, F(t,u)\,\mathrm{d}u,$$

and the comparison of these two formulas leads to the relation (1) as announced.

It is now easy to obtain the expression for $h'(t)$.

In fact, $h(t) = \int_{\mathbb{R}^n} f(t,x)\varphi(x)\,\mathrm{d}x$, hence, via (1):

$$\begin{aligned} h'(t) &= \int_{\mathbb{R}^n} \partial_t f(t,x)\varphi(x)\,\mathrm{d}x = \frac{1}{2}\int_{\mathbb{R}^n}\sum_{i,j} M_{ij}\partial_{ij} f(t,x)\varphi(x)\,\mathrm{d}x \\ &= \frac{1}{2}\sum_{i,j} M_{ij}\int_{\mathbb{R}^n} f(t,x)\,\partial_{ij}\varphi(x)\,\mathrm{d}x = \frac{1}{2}\sum_{i,j} M_{ij}\mathbb{E}[\partial_{ij}\varphi(Z_t)], \end{aligned}$$

after two integrations by parts (we note that $f(t,x)$ and $\partial_t f(t,x)$ are $\mathrm{O}\big(\mathrm{e}^{-\varepsilon|x|^2}\big)$, while $\varphi(x)$ is $\mathrm{O}(\mathrm{e}^{b|x|})$, which validates the preceding formal calculations). Finally, the formula $(*)$ used on $h'(t)$ above leads to the expression in the statement. □

Lemma III.1 will be used with functions φ that are not always $\mathcal{C}^2$; then we must either consider the derivatives as distributions, or regularize and make

© in this web service Cambridge University Press www.cambridge.org