

Gaussian measures, Hermite polynomials, and the Ornstein-Uhlenbeck semigroup

Jordan Bell

June 27, 2015

1 Definitions

For a topological space X , we denote by \mathcal{B}_X the Borel σ -algebra of X .

We write $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. With the order topology, $\overline{\mathbb{R}}$ is a compact metrizable space, and \mathbb{R} has the subspace topology inherited from $\overline{\mathbb{R}}$, namely the inclusion map is an embedding $\mathbb{R} \rightarrow \overline{\mathbb{R}}$. It follows that¹

$$\mathcal{B}_{\mathbb{R}} = \{E \cap \mathbb{R} : E \in \mathcal{B}_{\overline{\mathbb{R}}}\}.$$

If \mathcal{F} is a collection of functions $X \rightarrow \overline{\mathbb{R}}$ on a set X , we define $\bigvee \mathcal{F} : X \rightarrow \overline{\mathbb{R}}$ and $\bigwedge \mathcal{F} : X \rightarrow \overline{\mathbb{R}}$ by

$$\left(\bigvee \mathcal{F}\right)(x) = \sup\{f(x) : f \in \mathcal{F}\}, \quad x \in X$$

and

$$\left(\bigwedge \mathcal{F}\right)(x) = \inf\{f(x) : f \in \mathcal{F}\}, \quad x \in X.$$

If X is a measurable space and \mathcal{F} is a countable collection of measurable functions $X \rightarrow \overline{\mathbb{R}}$, it is a fact that $\bigwedge \mathcal{F}$ and $\bigvee \mathcal{F}$ are measurable $X \rightarrow \overline{\mathbb{R}}$.

2 Kolmogorov's inequality

Kolmogorov's inequality is the following.²

Theorem 1 (Kolmogorov's inequality). *Suppose that (Ω, \mathcal{S}, P) is a probability space, that $X_1, \dots, X_n \in L^2(P)$, that $E(X_1) = 0, \dots, E(X_n) = 0$, and that X_1, \dots, X_n are independent. Let*

$$S_k(\omega) = \sum_{j=1}^k X_j(\omega), \quad \omega \in \Omega,$$

¹Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 138, Lemma 4.20.

²Gerald B. Folland, *Real Analysis: Modern Techniques and Their Applications*, second ed., p. 322, Theorem 10.11.

for $1 \leq k \leq n$. Then for any $\lambda > 0$,

$$P\left(\left\{\omega \in \Omega : \bigvee_{k=1}^n |S_k(\omega)| \geq \lambda\right\}\right) \leq \frac{1}{\lambda^2} \sum_{j=1}^n V(X_j) = \frac{1}{\lambda^2} V(S_n).$$

3 1-dimension

For real a and $\sigma > 0$, one computes that

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{(t-a)^2}{2\sigma^2}\right) dt = 1. \quad (1)$$

Suppose that γ is a Borel probability measure on \mathbb{R} . If

$$\gamma = \delta_a$$

for some $a \in \mathbb{R}$ or has density

$$p(t, a, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-a)^2}{2\sigma^2}\right), \quad t \in \mathbb{R},$$

for some $a \in \mathbb{R}$ and some $\sigma > 0$, with respect to Lebesgue measure on \mathbb{R} , we say that γ is a **Gaussian measure**. We say that δ_a is a Gaussian measure with **mean** a and **variance** 0, and that a Gaussian measure with density $p(\cdot, a, \sigma^2)$ has **mean** a and **variance** σ^2 . A Gaussian measure with mean 0 and variance 1 is said to be **standard**.

One calculates that the **characteristic function** of a Gaussian measure γ with density $p(\cdot, a, \sigma^2)$ is

$$\tilde{\gamma}(y) = \int_{\mathbb{R}} \exp(iyx) d\gamma(x) = \exp\left(iay - \frac{1}{2}\sigma^2 y^2\right), \quad y \in \mathbb{R}. \quad (2)$$

The **cumulative distribution function** of a standard Gaussian measure γ is, for $t \in \mathbb{R}$,

$$\Phi(t) = \gamma(-\infty, t] = \int_{-\infty}^t d\gamma(s) = \int_{-\infty}^t p(s, 0, 1) ds = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) ds.$$

We define $\Phi(-\infty) = 0$ and also define

$$\Phi(\infty) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) ds = 1,$$

using (1).

$\Phi : \mathbb{R} \rightarrow [0, 1]$ is strictly increasing, thus $\Phi^{-1} : [0, 1] \rightarrow \overline{\mathbb{R}}$ makes sense, and is itself strictly increasing. Then $1 - \Phi$ is strictly decreasing. By (1),

$$\begin{aligned} 1 - \Phi(t) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) ds - \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) ds \\ &= \int_t^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) ds. \end{aligned}$$

The following lemma gives an estimate for $1 - \Phi(t)$ that tells us something substantial as $t \rightarrow +\infty$, beyond the immediate fact that $(1 - \Phi)(\infty) = 1 - \Phi(\infty) = 0$.³

Lemma 2. *For $t > 0$,*

$$\frac{1}{\sqrt{2\pi}} \left(\frac{1}{t} - \frac{1}{t^3} \right) e^{-t^2/2} \leq 1 - \Phi(t) \leq \frac{1}{\sqrt{2\pi}} \frac{1}{t} e^{-t^2/2}.$$

Proof. Integrating by parts,

$$\begin{aligned} 1 - \Phi(t) &= \int_t^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) ds \\ &= \int_t^\infty \frac{1}{s\sqrt{2\pi}} \cdot s \exp\left(-\frac{s^2}{2}\right) ds \\ &= -\frac{1}{s\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) \Big|_t^\infty - \int_t^\infty \frac{1}{s^2\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) ds \\ &\leq \frac{1}{t\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right). \end{aligned}$$

On the other hand, using the above work and again integrating by parts,

$$\begin{aligned} 1 - \Phi(t) &= \frac{1}{t\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) - \int_t^\infty \frac{1}{s^3\sqrt{2\pi}} \cdot s \exp\left(-\frac{s^2}{2}\right) ds \\ &= \frac{1}{t\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) + \frac{1}{s^3\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) \Big|_t^\infty \\ &\quad + \int_t^\infty \frac{3}{s^4\sqrt{2\pi}} \exp\left(-\frac{s^2}{2}\right) ds \\ &\geq \frac{1}{t\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) - \frac{1}{t^3\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right). \end{aligned}$$

□

The following theorem shows that if the variances of a sequence of independent centered random variables are summable then the sequence of random variables is summable almost surely.⁴

Theorem 3. *Suppose that $\xi_j \in L^2(\Omega, \mathcal{S}, P)$, $j \geq 1$, are independent random variables each with mean 0. If $\sum_{j=1}^\infty V(\xi_j) < \infty$, then $\sum_{j=1}^\infty \xi_j$ converges almost surely.*

Proof. Define $S_n : \Omega \rightarrow \mathbb{R}$ by

$$S_n(\omega) = \sum_{j=1}^n \xi_j(\omega),$$

³Vladimir I. Bogachev, *Gaussian Measures*, p. 2, Lemma 1.1.3.

⁴Karl R. Stromberg, *Probability for Analysts*, p. 58, Theorem 4.6.

define $Z_n : \Omega \rightarrow [0, \infty]$ by

$$Z_n = \bigvee_{j=1}^{\infty} |S_{n+j} - S_n|,$$

and define $Z : \Omega \rightarrow [0, \infty]$ by

$$Z = \bigwedge_{n=1}^{\infty} Z_n.$$

If $S_n(\omega)$ converges and $\epsilon > 0$, there is some n such that for all $j \geq 1$, $|S_{n+j}(\omega) - S_n(\omega)| < \epsilon$ and so $Z_n(\omega) \leq \epsilon$ and $Z(\omega) \leq \epsilon$. Therefore, if $S_n(\omega)$ converges then $Z(\omega) = 0$. On the other hand, if $Z(\omega) = 0$ and $\epsilon > 0$, there is some n such that $Z_n(\omega) < \epsilon$, hence $|S_{n+j}(\omega) - S_n(\omega)| < \epsilon$ for all $j \geq 1$. That is, $S_n(\omega)$ is a Cauchy sequence in \mathbb{R} , and hence converges. Therefore

$$\{\omega \in \Omega : S_n(\omega) \text{ converges}\} = \{\omega \in \Omega : Z(\omega) = 0\}. \quad (3)$$

Let $\epsilon > 0$. For any n and k , using Kolmogorov's inequality with $X_j = \xi_{n+j}$ for $j = 1, \dots, k$,

$$P\left(\bigvee_{j=1}^k |S_{n+j} - S_n| \geq \epsilon\right) \leq \frac{1}{\epsilon^2} \sum_{j=1}^k V(X_j) \leq \frac{1}{\epsilon^2} \sum_{j=n+1}^{\infty} V(\xi_j).$$

Because this is true for each k , it follows that

$$P(Z_n \geq \epsilon) \leq \frac{1}{\epsilon^2} \sum_{j=n+1}^{\infty} V(\xi_j),$$

hence, for each n ,

$$P(Z \geq \epsilon) \leq P(Z_n \geq \epsilon) \leq \frac{1}{\epsilon^2} \sum_{j=n+1}^{\infty} V(\xi_j).$$

Because $\sum_{j=1}^{\infty} V(\xi_j) < \infty$, $\sum_{j=n+1}^{\infty} V(\xi_j) \rightarrow 0$ as $n \rightarrow \infty$, so

$$P(Z \geq \epsilon) = 0.$$

Because this is true for all $\epsilon > 0$, we get $P(Z > 0) = 0$, i.e. $P(Z = 0) = 1$. By (3), this means that S_n converges almost surely. \square

The following theorem gives conditions under which the converse of the above theorem holds.⁵

⁵Karl R. Stromberg, *Probability for Analysts*, p. 59, Theorem 4.7.

Theorem 4. Suppose that $\xi_j \in L^2(\Omega, \mathcal{S}, P)$, $j \geq 1$, are independent random variables each with mean 0, and let $S_n = \sum_{j=1}^n \xi_j$. If

$$P\left(\bigvee_{n=1}^{\infty} |S_n| < \infty\right) > 0 \quad (4)$$

and there is some $\beta \in [0, \infty)$ such that $\bigvee_{j=1}^{\infty} |\xi_j| \leq \beta$ almost surely, then

$$\sum_{j=1}^{\infty} V(\xi_j) < \infty.$$

Proof. By (4), there is some $\alpha \in [0, \infty)$ such that $P(A) > 0$, for

$$A = \left\{ \omega \in \Omega : \bigvee_{n=1}^{\infty} |S_n(\omega)| \leq \alpha \right\}.$$

For $p \geq 1$, let

$$A_p = \left\{ \omega \in \Omega : \bigvee_{n=1}^p |S_n(\omega)| \leq \alpha \right\},$$

which satisfies $A_p \downarrow A$ as $p \rightarrow \infty$. For each p , the random variables $\chi_{A_p} S_p$ and ξ_{p+1} are independent and the random variables χ_{A_p} and ξ_{p+1}^2 are independent, whence

$$\begin{aligned} E(\chi_{A_p} S_{p+1}^2) &= E(\chi_{A_p} (S_p + \xi_{p+1})(S_p + \xi_{p+1})) \\ &= E(\chi_{A_p} S_p^2 + 2\chi_{A_p} S_p \xi_{p+1} + \chi_{A_p} \xi_{p+1}^2) \\ &= E(\chi_{A_p} S_p^2) + 2E(\chi_{A_p} S_p)E(\xi_{p+1}) + E(\chi_{A_p})E(\xi_{p+1}^2) \\ &= E(\chi_{A_p} S_p^2) + P(A_p)V(\xi_{p+1}) \\ &\geq E(\chi_{A_p} S_p^2) + P(A)V(\xi_{p+1}). \end{aligned}$$

Set $B_p = A_p \setminus A_{p+1}$. For $\omega \in A_p$, $|S_p(\omega)| \leq \alpha$, and for almost all $\omega \in \Omega$, $|\xi_{p+1}(\omega)| \leq \beta$, so for almost all $\omega \in B_p$,

$$|S_{p+1}(\omega)| \leq |S_p(\omega)| + |\xi_{p+1}(\omega)| \leq \alpha + \beta,$$

hence

$$\begin{aligned} P(A)V(\xi_{p+1}) &\leq E((\chi_{B_p} + \chi_{A_{p+1}})S_{p+1}^2) - E(\chi_{A_p} S_p^2) \\ &= E(\chi_{B_p} S_{p+1}^2) + E(\chi_{A_{p+1}} S_{p+1}^2) - E(\chi_{A_p} S_p^2) \\ &\leq P(B_p)(\alpha + \beta)^2 + E(\chi_{A_{p+1}} S_{p+1}^2) - E(\chi_{A_p} S_p^2). \end{aligned}$$

Adding the inequalities for $p = 1, 2, \dots, n-1$, because B_p are pairwise disjoint,

$$\begin{aligned} P(A) \sum_{p=1}^{n-1} V(\xi_{p+1}) &= (\alpha + \beta)^2 \sum_{p=1}^{n-1} P(B_p) + E(\chi_{A_n} S_n^2) - E(\chi_{A_1} S_1^2) \\ &\leq (\alpha + \beta)^2 + E(\chi_{A_n} S_n^2) \\ &\leq (\alpha + \beta)^2 + \alpha^2. \end{aligned}$$

Because this is true for all n and $P(A) > 0$,

$$\sum_{p=1}^{\infty} V(\xi_{p+1}) < \infty,$$

and with $V(\xi_1) < \infty$ this completes the proof. \square

4 n -dimensions

If μ is a finite Borel measure on \mathbb{R}^n , we define the **characteristic function of μ** by

$$\tilde{\mu}(y) = \int_{\mathbb{R}^n} e^{i\langle y, x \rangle} d\mu(x), \quad y \in \mathbb{R}^n.$$

A Borel probability measure γ on \mathbb{R}^n is said to be **Gaussian** if for each $f \in (\mathbb{R}^n)^*$, the pushforward measure $f_*\gamma$ on \mathbb{R} is a Gaussian measure on \mathbb{R} , where

$$(f_*\gamma)(E) = \gamma(f^{-1}(E))$$

for E a Borel set in \mathbb{R} .

We now give a characterization of Gaussian measures on \mathbb{R}^n and their densities.⁶ In the following theorem, the vector $a \in \mathbb{R}^n$ is called the **mean of γ** and the linear transformation $K \in \mathcal{L}(\mathbb{R}^n)$ is called the **covariance operator of γ** . When $a = 0 \in \mathbb{R}^n$ and $K = \text{id}_{\mathbb{R}^n}$, we say that γ is **standard**.

Theorem 5. *A Borel probability measure γ on \mathbb{R}^n is Gaussian if and only if there is some $a \in \mathbb{R}^n$ and some positive semidefinite $K \in \mathcal{L}(\mathbb{R}^n)$ such that*

$$\tilde{\gamma}(y) = \exp \left(i \langle y, a \rangle - \frac{1}{2} \langle Ky, y \rangle \right), \quad y \in \mathbb{R}^n. \quad (5)$$

If γ is a Gaussian measure whose covariance operator K is positive definite, then the density of γ with respect to Lebesgue measure on \mathbb{R}^n is

$$x \mapsto \frac{1}{\sqrt{(2\pi)^n \det K}} \exp \left(-\frac{1}{2} \langle K^{-1}(x - a), x - a \rangle \right), \quad x \in \mathbb{R}^n.$$

Proof. Suppose that (5) is satisfied. Let $f \in (\mathbb{R}^n)^*$, i.e. a linear map $\mathbb{R}^n \rightarrow \mathbb{R}$, and put $\nu = f_*\gamma$. Using the change of variables formula, the characteristic function of ν is

$$\tilde{\nu}(t) = \int_{\mathbb{R}} e^{its} d\nu(s) = \int_{\mathbb{R}^n} e^{itf(x)} d\gamma(x), \quad t \in \mathbb{R}.$$

⁶Vladimir I. Bogachev, *Gaussian Measures*, p. 3, Proposition 1.2.2; Michel Simonnet, *Measures and Probabilities*, p. 303, Theorem 14.5.

Let v be the unique element of \mathbb{R}^n such that $f(x) = \langle v, x \rangle$ for all $x \in \mathbb{R}^n$. Then

$$\tilde{\nu}(t) = \int_{\mathbb{R}^n} e^{i\langle tv, x \rangle} d\gamma(x) = \tilde{\gamma}(tv).$$

so by (5),

$$\tilde{\nu}(t) = \exp \left(i \langle tv, a \rangle - \frac{1}{2} \langle Ktv, tv \rangle \right) = \exp \left(if(a)t - \frac{1}{2} \langle Kv, v \rangle t^2 \right).$$

This implies that ν is a Gaussian measure on \mathbb{R} with mean $f(a)$ and variance $\langle Kv, v \rangle$: if $\langle Kv, v \rangle = 0$ then $\nu = \delta_{f(a)}$, and if $\langle Kv, v \rangle > 0$ then ν has density

$$\frac{1}{\sqrt{\langle Kv, v \rangle} \sqrt{2\pi}} \exp \left(-\frac{(s - f(a))^2}{2 \langle Kv, v \rangle} \right), \quad s \in \mathbb{R},$$

with respect to Lebesgue measure on \mathbb{R} . That is, for any $f \in (\mathbb{R}^n)^*$, the pushforward measure $f_*\gamma$ is a Gaussian measure on \mathbb{R} , which is what it means for γ to be a Gaussian measure on \mathbb{R}^n .

Suppose that γ is Gaussian and let $f \in (\mathbb{R}^n)^*$. Then the pushforward measure $f_*\gamma$ is a Gaussian measure on \mathbb{R} . Let $a(f)$ be the mean of $f_*\gamma$ and let $\sigma^2(f)$ be the variance of $f_*\gamma$, and let v_f be the unique element of \mathbb{R}^n such that $f(x) = \langle x, v_f \rangle$ for all $x \in \mathbb{R}^n$. Using the change of variables formula,

$$a(f) = \int_{\mathbb{R}} td(f_*\gamma)(t) = \int_{\mathbb{R}^n} f(x)d\gamma(x)$$

and

$$\begin{aligned} \sigma^2(f) &= \int_{\mathbb{R}} (t - a(f))^2 d(f_*\gamma)(t) \\ &= \int_{\mathbb{R}^n} (f(x) - a(f))^2 d\gamma(x) \\ &= \int_{\mathbb{R}^n} (f(x)^2 - 2f(x)a(f) + a(f)^2) d\gamma(x). \end{aligned}$$

Because $f \mapsto a(f)$ is linear $(\mathbb{R}^n)^* \rightarrow \mathbb{R}$, there is a unique $a \in \mathbb{R}^n = (\mathbb{R}^n)^{**}$ such that

$$a(f) = \langle v_f, a \rangle, \quad f \in (\mathbb{R}^n)^*.$$

For $f, g \in (\mathbb{R}^n)^*$,

$$\begin{aligned} \sigma^2(f+g) &= \int_{\mathbb{R}^n} (f(x)^2 + 2f(x)g(x) + g(x)^2 \\ &\quad - 2f(x)a(f) - 2f(x)a(g) - 2g(x)a(f) - 2g(x)a(g) \\ &\quad + a(f)^2 + 2a(f)a(g) + a(g)^2) d\gamma(x), \end{aligned}$$

so

$$\begin{aligned}\sigma^2(f+g) - \sigma^2(f) - \sigma^2(g) &= \int_{\mathbb{R}^n} (2f(x)g(x) - 2f(x)a(g) - 2g(x)a(f) \\ &\quad + 2a(f)a(g))d\gamma(x).\end{aligned}$$

$B(f, g) = \frac{1}{2}(\sigma^2(f+g) - \sigma^2(f) - \sigma^2(g))$ is a symmetric bilinear form on \mathbb{R}^n , and

$$B(f, f) = 2 \int_{\mathbb{R}^n} (f(x) - a(f))^2 d\gamma(x) \geq 0,$$

namely, B is positive semidefinite. It follows that there is a unique positive semidefinite $K \in \mathcal{L}(\mathbb{R}^n)$ such that $B(f, g) = \langle Kv_f, v_g \rangle$ for all $f, g \in (\mathbb{R}^n)^*$. For $y \in \mathbb{R}^n$ and for $v_f = y$, using the change of variables formula, using the fact that $f_*\gamma$ is a Gaussian measure on \mathbb{R} with mean

$$a(f) = \langle v_f, a \rangle = \langle y, a \rangle$$

and variance

$$\sigma^2(f) = B(f, f) = \langle Kv_f, v_f \rangle = \langle Ky, y \rangle$$

and using (2),

$$\begin{aligned}\tilde{\gamma}(y) &= \int_{\mathbb{R}^n} e^{if(x)} d\gamma(x) \\ &= \int_{\mathbb{R}} e^{it} d(f_*\gamma)(t) \\ &= \exp\left(i \langle y, a \rangle \cdot 1 - \frac{1}{2} \langle Ky, y \rangle \cdot 1^2\right) \\ &= \exp\left(i \langle y, a \rangle - \frac{1}{2} \langle Ky, y \rangle\right),\end{aligned}$$

which shows that (5) is satisfied.

Suppose that γ is a Gaussian measure and further that the covariance operator K is positive definite. By the spectral theorem, there is an orthonormal basis $\{e_1, \dots, e_n\}$ for \mathbb{R}^n such that $\langle Ke_j, e_j \rangle > 0$ for each $1 \leq j \leq n$. Write $\langle Ke_j, e_j \rangle = \sigma_j^2$, and for $y \in \mathbb{R}^n$ set $y_j = \langle y, e_j \rangle$, with which $y = y_1 e_1 + \dots + y_n e_n$ and then

$$\begin{aligned}\langle Ky, y \rangle &= \langle y_1 K e_1 + \dots + y_n K e_n, y_1 e_1 + \dots + y_n e_n \rangle \\ &= \langle y_1 \sigma_1^2 e_1 + \dots + y_n \sigma_n^2 e_n, y_1 e_1 + \dots + y_n e_n \rangle \\ &= \sigma_1^2 y_1^2 + \dots + \sigma_n^2 y_n^2.\end{aligned}$$

And

$$\langle y, a \rangle = \langle y_1 e_1 + \dots + y_n e_n, a_1 e_1 + \dots + a_n e_n \rangle = a_1 y_1 + \dots + a_n y_n.$$

Let γ_j be the Gaussian measure on \mathbb{R} with mean a_j and variance σ_j^2 . Because $\sigma_j^2 > 0$, the measure γ_j has density $p(\cdot, a_j, \sigma_j^2)$ with respect to Lebesgue measure on \mathbb{R} , and thus

$$\begin{aligned}
\tilde{\gamma}(y) &= \exp \left(i \langle y, a \rangle - \frac{1}{2} \langle Ky, y \rangle \right) \\
&= \exp \left(i \sum_{j=1}^n a_j y_j - \frac{1}{2} \sum_{j=1}^n \sigma_j^2 y_j^2 \right) \\
&= \prod_{j=1}^n \exp \left(i a_j y_j - \frac{1}{2} \sigma_j^2 y_j^2 \right) \\
&= \prod_{j=1}^n \tilde{\gamma}_j(y_j) \\
&= \prod_{j=1}^n \int_{\mathbb{R}} \exp(i y_j t) d\gamma_j(t) \\
&= \prod_{j=1}^n \int_{\mathbb{R}} \exp(i y_j t) p(t, a_j, \sigma_j^2) dt \\
&= \int_{\mathbb{R}^n} \prod_{j=1}^n \exp(i y_j x_j) p(x_j, a_j, \sigma_j^2) dx \\
&= \int_{\mathbb{R}^n} e^{i \langle y, x \rangle} \prod_{j=1}^n p(x_j, a_j, \sigma_j^2) dx.
\end{aligned}$$

This implies that γ has density

$$x \mapsto \prod_{j=1}^n p(x_j, a_j, \sigma_j^2), \quad x \in \mathbb{R}^n,$$

with respect to Lebesgue measure on \mathbb{R}^n . Moreover,

$$\begin{aligned}
\langle K^{-1}(x - a), x - a \rangle &= \left\langle \sum_{j=1}^n \sigma_j^{-2} (x_j - a_j) e_j, \sum_{j=1}^n (x_j - a_j) e_j \right\rangle \\
&= \sum_{j=1}^n \frac{(x_j - a_j)^2}{\sigma_j^2},
\end{aligned}$$

so we have, as $\det K = \prod_{j=1}^n \sigma_j^2$,

$$\begin{aligned}
\prod_{j=1}^n p(x_j, a_j, \sigma_j^2) &= \prod_{j=1}^n \frac{1}{\sigma_j \sqrt{2\pi}} \exp \left(-\frac{(x_j - a_j)^2}{2\sigma_j^2} \right) \\
&= \frac{1}{\sqrt{(2\pi)^n \det K}} \exp \left(-\frac{1}{2} \langle K^{-1}(x - a), x - a \rangle \right).
\end{aligned}$$

□

Because \mathbb{R} is a second-countable topological space, the Borel σ -algebra $\mathcal{B}_{\mathbb{R}^n}$ is equal to the product σ -algebra $\bigotimes_{j=1}^n \mathcal{B}_{\mathbb{R}}$. The density of the standard Gaussian measure γ_n with respect to Lebesgue measure on \mathbb{R}^n is, by Theorem 5,

$$x \mapsto \frac{1}{\sqrt{(2\pi)^n}} \exp\left(-\frac{1}{2} \langle x, x \rangle\right), \quad x \in \mathbb{R}^n.$$

It follows that γ_n is equal to the product measure $\prod_{j=1}^n \gamma_1$, and thus that the probability space $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \gamma_n)$ is equal to the product $\prod_{j=1}^n (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \gamma_1)$.

For $f_1, \dots, f_n \in L^2(\gamma_1)$, we define $f_1 \otimes \dots \otimes f_n \in L^2(\gamma_n)$, called the **tensor product of f_1, \dots, f_n** , by

$$(f_1 \otimes \dots \otimes f_n)(x) = \prod_{j=1}^n f_j(x_j), \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

It is straightforward to check that for $f_1, \dots, f_n, g_1, \dots, g_n \in L^2(\gamma_1)$,

$$\langle f_1 \otimes \dots \otimes f_n, g_1 \otimes \dots \otimes g_n \rangle_{L^2(\gamma_n)} = \prod_{j=1}^n \langle f_j, g_j \rangle_{L^2(\gamma_1)}.$$

One proves that the linear span of the collection of all tensor products is dense in $L^2(\gamma_n)$, and that $\{v_k : k \geq 0\}$ is an orthonormal basis for $L^2(\gamma_1)$, then

$$\{v_{k_1} \otimes \dots \otimes v_{k_n} : (k_1, \dots, k_n) \in \mathbb{Z}_{\geq 0}^n\} \quad (6)$$

is an orthonormal basis for $L^2(\gamma_n)$.

We will later use the following statement about centered Gaussian measures.⁷

Theorem 6. *Let γ be a Gaussian measure on \mathbb{R}^n with mean 0 and let $\theta \in \mathbb{R}$. Then the pushforward of the product measure $\gamma \times \gamma$ on $\mathbb{R}^n \times \mathbb{R}^n$ under the mapping $(u, v) \mapsto u \sin \theta + v \cos \theta$, $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, is equal to γ .*

Proof. Let μ be the pushforward of $\gamma \times \gamma$ under the above mapping. and let $K \in \mathcal{L}(\mathbb{R}^n)$ be the covariance operator of γ . For $y \in \mathbb{R}^n$, using the change of variables formula,

$$\begin{aligned} \int_{\mathbb{R}^n} \exp(i \langle y, x \rangle) d\mu(x) &= \int_{\mathbb{R}^n \times \mathbb{R}^n} \exp(i \langle y, u \sin \theta + v \cos \theta \rangle) d(\gamma \times \gamma)(u, v) \\ &= \left(\int_{\mathbb{R}^n} \exp(i \langle y \sin \theta, u \rangle) d\gamma(u) \right) \\ &\quad \cdot \left(\int_{\mathbb{R}^n} \exp(i \langle y \cos \theta, v \rangle) d\gamma(v) \right) \\ &= \tilde{\gamma}(y \sin \theta) \tilde{\gamma}(y \cos \theta). \end{aligned}$$

⁷Vladimir I. Bogachev, *Gaussian Measures*, p. 5, Lemma 1.2.5.

By Theorem 5,

$$\begin{aligned}\tilde{\gamma}(y \sin \theta) \tilde{\gamma}(y \cos \theta) &= \exp \left(-\frac{1}{2} \langle Ky \sin \theta, y \sin \theta \rangle \right) \exp \left(-\frac{1}{2} \langle Ky \cos \theta, y \cos \theta \rangle \right) \\ &= \exp \left(-\frac{1}{2} \sin^2(\theta) \langle Ky, y \rangle - \frac{1}{2} \cos^2(\theta) \langle Ky, y \rangle \right) \\ &= \exp \left(-\frac{1}{2} \langle Ky, y \rangle \right).\end{aligned}$$

Thus, the characteristic function of μ is

$$\tilde{\mu}(y) = \exp \left(-\frac{1}{2} \langle Ky, y \rangle \right), \quad y \in \mathbb{R}^n,$$

which implies that μ is equal to the Gaussian measure with mean 0 and covariance operator K , i.e., $\mu = \gamma$. \square

5 Hermite polynomials

For $k \geq 0$, we define the **Hermite polynomial** H_k by

$$H_k(t) = \frac{(-1)^k}{\sqrt{k!}} \exp \left(\frac{t^2}{2} \right) \frac{d^k}{dt^k} \exp \left(-\frac{t^2}{2} \right), \quad t \in \mathbb{R}.$$

It is apparent that $H_k(t)$ is a polynomial of degree k .

Theorem 7. For real λ and t ,

$$\exp \left(\lambda t - \frac{1}{2} \lambda^2 \right) = \sum_{k=0}^{\infty} \frac{1}{\sqrt{k!}} H_k(t) \lambda^k.$$

Proof. For $u \in \mathbb{C}$, let $g(u) = \exp \left(-\frac{1}{2} u^2 \right)$. For $t \in \mathbb{R}$,

$$\begin{aligned}g(u) &= \sum_{k=0}^{\infty} \frac{g^{(k)}(t)}{k!} (u-t)^k \\ &= \sum_{k=0}^{\infty} \frac{\sqrt{k!}}{(-1)^k} \exp \left(-\frac{t^2}{2} \right) H_k(t) \frac{1}{k!} (u-t)^k \\ &= \exp \left(-\frac{t^2}{2} \right) \sum_{k=0}^{\infty} \frac{(-1)^k}{\sqrt{k!}} H_k(t) (u-t)^k.\end{aligned}$$

Therefore, for real λ and t ,

$$\begin{aligned}
\exp\left(\lambda t - \frac{1}{2}\lambda^2\right) &= \exp\left(\frac{1}{2}t^2 - \frac{1}{2}(\lambda - t)^2\right) \\
&= \exp\left(\frac{1}{2}t^2\right) g(\lambda - t) \\
&= \exp\left(\frac{1}{2}t^2\right) g(t - \lambda) \\
&= \exp\left(\frac{1}{2}t^2\right) \exp\left(-\frac{t^2}{2}\right) \sum_{k=0}^{\infty} \frac{(-1)^k}{\sqrt{k!}} H_k(t) (-\lambda)^k \\
&= \sum_{k=0}^{\infty} \frac{1}{\sqrt{k!}} H_k(t) \lambda^k.
\end{aligned}$$

□

Theorem 8. Let γ_1 be the standard Gaussian measure on \mathbb{R} , with density $p(t, 0, 1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$. Then

$$\{H_k : k \geq 0\}$$

is an orthonormal basis for $L^2(\gamma_1)$.

Proof. For $\lambda, \mu \in \mathbb{R}$, on the one hand, using (1) with $a = \lambda + \mu$ and $\sigma = 1$,

$$\begin{aligned}
&\int_{\mathbb{R}} \exp\left(\lambda t - \frac{1}{2}\lambda^2\right) \exp\left(\mu t - \frac{1}{2}\mu^2\right) d\gamma_1(t) \\
&= e^{\lambda\mu} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(t - (\lambda + \mu))^2\right) dt \\
&= e^{\lambda\mu}.
\end{aligned}$$

On the other hand, using Theorem 7,

$$\begin{aligned}
&\int_{\mathbb{R}} \exp\left(\lambda t - \frac{1}{2}\lambda^2\right) \exp\left(\mu t - \frac{1}{2}\mu^2\right) d\gamma_1(t) \\
&= \int_{\mathbb{R}} \left(\sum_{k=0}^{\infty} \frac{1}{\sqrt{k!}} H_k(t) \lambda^k\right) \left(\sum_{l=0}^{\infty} \frac{1}{\sqrt{l!}} H_l(t) \mu^l\right) d\gamma_1(t) \\
&= \int_{\mathbb{R}} \sum_{k,l \geq 0} \frac{1}{\sqrt{k!l!}} \lambda^k \mu^l H_k(t) H_l(t) d\gamma_1(t) \\
&= \sum_{k,l \geq 0} \frac{1}{\sqrt{k!l!}} \lambda^k \mu^l \langle H_k, H_l \rangle_{L^2(\gamma_1)}.
\end{aligned}$$

Therefore

$$\sum_{k,l \geq 0} \frac{1}{\sqrt{k!l!}} \lambda^k \mu^l \langle H_k, H_l \rangle_{L^2(\gamma_1)} = \sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k \mu^k.$$

From this, we get that if $k \neq l$ then $\frac{1}{\sqrt{k!l!}} \langle H_k, H_l \rangle_{L^2(\gamma_1)} = 0$, i.e.

$$\langle H_k, H_l \rangle_{L^2(\gamma_1)} = 0.$$

If $k = l$, then $\frac{1}{\sqrt{k!l!}} \langle H_k, H_l \rangle_{L^2(\gamma_1)} = \frac{1}{k!}$, i.e.

$$\langle H_k, H_k \rangle_{L^2(\gamma_1)} = 1.$$

Therefore, $\{H_k : k \geq 0\}$ is an orthonormal set in $L^2(\gamma_1)$.

Suppose that $f \in L^2(\gamma_1)$ satisfies $\langle f, H_k \rangle_{L^2(\gamma_1)} = 0$ for each $k \geq 0$. Because $H_k(t)$ is a polynomial of degree k , for each $k \geq 0$ we have

$$\text{span}\{H_0, H_1, H_2, \dots, H_k\} = \text{span}\{1, t, t^2, \dots, t^k\}.$$

Hence for each $k \geq 0$, $\langle f, t^k \rangle_{L^2(\gamma_1)} = 0$. One then proves that $\text{span}\{1, t, t^2, \dots\}$ is dense in $L^2(\gamma_1)$, from which it follows that the linear span of the Hermite polynomials is dense in $L^2(\gamma_1)$ and thus that they are an orthonormal basis. \square

Lemma 9. For $k \geq 1$,

$$H'_k(t) = \sqrt{k}H_{k-1}(t), \quad H'_k(t) = tH_k(t) - \sqrt{k+1}H_{k+1}(t).$$

Proof. Theorem 7 says

$$\exp\left(\lambda t - \frac{1}{2}\lambda^2\right) = \sum_{k=0}^{\infty} \frac{1}{\sqrt{k!}} H_k(t) \lambda^k.$$

On the one hand,

$$\begin{aligned} \frac{d}{dt} \exp\left(\lambda t - \frac{1}{2}\lambda^2\right) &= \lambda \exp\left(\lambda t - \frac{1}{2}\lambda^2\right) \\ &= \sum_{k=0}^{\infty} \frac{1}{\sqrt{k!}} H_k(t) \lambda^{k+1} \\ &= \sum_{k=1}^{\infty} \frac{1}{\sqrt{(k-1)!}} H_{k-1}(t) \lambda^k. \end{aligned}$$

On the other hand,

$$\frac{d}{dt} \exp\left(\lambda t - \frac{1}{2}\lambda^2\right) = \sum_{k=0}^{\infty} \frac{1}{\sqrt{k!}} H'_k(t) \lambda^k.$$

Therefore, $H'_0(t) = 0$, and for $k \geq 1$,

$$\frac{1}{\sqrt{(k-1)!}} H_{k-1}(t) = \frac{1}{\sqrt{k!}} H'_k(t),$$

i.e.,

$$H'_k(t) = \sqrt{k}H_{k-1}(t).$$

\square

For $\alpha = (k_1, \dots, k_n) \in \mathbb{Z}_{\geq 0}^n$, we define the **Hermite polynomial** H_α by

$$H_\alpha(x) = H_{k_1}(x_1) \cdots H_{k_n}(x_n), \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Because the collection of all Hermite polynomials H_k is an orthonormal basis for the Hilbert space $L^2(\gamma_1)$, following (6) we have that the collection of all Hermite polynomials H_α is an orthonormal basis for the Hilbert space $L^2(\gamma_n)$.

Theorem 10. *For γ_n the standard Gaussian measure on \mathbb{R}^n , with mean $0 \in \mathbb{R}^n$ and covariance operator $\text{id}_{\mathbb{R}^n}$, the collection*

$$\{H_\alpha : \alpha \in \mathbb{Z}_{\geq 0}^n\}$$

is an orthonormal basis for $L^2(\gamma_n)$.

For $\alpha = (k_1, \dots, k_n) \in \mathbb{Z}_{\geq 0}^n$, write $|\alpha| = k_1 + \dots + k_n$. For $k \geq 0$, we define

$$\mathcal{X}_k = \text{span}\{H_\alpha : |\alpha| = k\},$$

which is a subspace of $L^2(\gamma_n)$ of dimension

$$\binom{k+n-1}{k}.$$

As \mathcal{X}_k is a finite dimensional subspace of $L^2(\gamma_n)$, it is closed. $L^2(\gamma_n)$ is equal to the orthogonal direct sum of the \mathcal{X}_k :

$$L^2(\gamma_n) = \bigoplus_{k=0}^{\infty} \mathcal{X}_k.$$

Let

$$I_k : L^2(\gamma_n) \rightarrow \mathcal{X}_k$$

be the orthogonal projection onto \mathcal{X}_k .

6 Ornstein-Uhlenbeck semigroup

Let γ be a Gaussian measure on \mathbb{R}^n with mean 0 and covariance operator K . For $t \geq 0$, we define $M_t : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ by

$$M_t(u, v) = e^{-t}u + \sqrt{1 - e^{-2t}}v, \quad (x, y) \in \mathbb{R}^n \times \mathbb{R}^n.$$

By Theorem 6, $M_{t*}(\gamma \times \gamma) = \gamma$. Therefore, for $p \geq 1$ and $f \in L^p(\gamma)$, using the change of variables formula,

$$\int_{\mathbb{R}^n} |f(x)|^p d\gamma(x) = \int_{\mathbb{R}^n \times \mathbb{R}^n} |f(M_t(u, v))|^p d(\gamma \times \gamma)(u, v).$$

Applying Fubini's theorem, the function

$$u \mapsto \int_{\mathbb{R}^n} |f(M_t(u, v))|^p d\gamma(v) = \int_{\mathbb{R}^n} |f(e^{-t}u + \sqrt{1 - e^{-2t}}v)|^p d\gamma(v)$$

belongs to $L^1(\gamma)$. We define the **Ornstein-Uhlenbeck semigroup** $\{T_t : t \geq 0\}$ on $L^p(\gamma)$, $p \geq 1$, by

$$T_t(f)(u) = \int_{\mathbb{R}^n} f(M_t(u, v)) d\gamma(v) = \int_{\mathbb{R}^n} f\left(e^{-t}u + \sqrt{1 - e^{-2t}}v\right) d\gamma(v),$$

for $u \in \mathbb{R}^n$.

Theorem 11. *Let γ be a Gaussian measure on \mathbb{R}^n with mean 0. If $f \in L^1(\gamma)$, then*

$$\int_{\mathbb{R}^n} (T_t f)(x) d\gamma(x) = \int_{\mathbb{R}^n} f(x) d\gamma(x).$$

Proof. Using Fubini's theorem, then the change of variables formula, then Theorem 6,

$$\begin{aligned} \int_{\mathbb{R}^n} (T_t f)(u) d\gamma(u) &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(M_t(u, v)) d\gamma(v) \right) d\gamma(u) \\ &= \int_{\mathbb{R}^n \times \mathbb{R}^n} f(M_t(u, v)) d(\gamma \times \gamma)(u, v) \\ &= \int_{\mathbb{R}^n} f(x) d(M_{t*}(\gamma \times \gamma))(x) \\ &= \int_{\mathbb{R}^n} f(x) d\gamma(x). \end{aligned}$$

□

Theorem 12. *Let γ be a Gaussian measure on \mathbb{R}^n with mean 0. For $p \geq 1$ and $t \geq 0$, T_t is a bounded linear operator $L^p(\gamma) \rightarrow L^p(\gamma)$ with operator norm 1.*

Proof. For $f \in L^p(\gamma)$, using Jensen's inequality and then Theorem 11,

$$\begin{aligned} \|T_t f\|_{L^p(\gamma)}^p &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(M_t(u, v)) d\gamma(v) \right|^p d\gamma(u) \\ &\leq \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(M_t(u, v))|^p d\gamma(v) \right) d\gamma(u) \\ &= \int_{\mathbb{R}^n} T_t(|f|^p)(u) d\gamma(u) \\ &= \int_{\mathbb{R}^n} |f|^p(u) d\gamma(u) \\ &= \|f\|_{L^p(\gamma)}^p, \end{aligned}$$

i.e. $\|T_t f\|_{L^p(\mu)} \leq \|f\|_{L^p(\mu)}$. This shows that the operator norm of T_t is ≤ 1 . But, as γ is a probability measure,

$$T_t 1 = \int_{\mathbb{R}^n} 1 d\gamma(v) = 1,$$

so T_t has operator norm 1. \square

For a Banach space E , we denote by $\mathcal{B}(E)$ the set of bounded linear operators $E \rightarrow E$. The **strong operator topology on E** is the coarsest topology on E such that for each $x \in E$, the map $A \mapsto Ax$ is continuous $\mathcal{B}(E) \rightarrow E$. To say that a map $Q : [0, \infty) \rightarrow \mathcal{B}(E)$ is **strongly continuous means** that for each $t \in [0, \infty)$, $Q(s) \rightarrow Q_t$ in the strong operator topology as $s \rightarrow t$, i.e., for each $x \in E$, $Q(s)x \rightarrow Q(t)x$ in E .

A **one-parameter semigroup in $\mathcal{B}(E)$** is a map $Q : [0, \infty) \rightarrow \mathcal{B}(E)$ such that (i) $Q(0) = \text{id}_E$ and (ii) for $s, t \geq 0$, $Q(s+t) = Q(s) \circ Q(t)$. For a one-parameter semigroup to be strongly continuous, one proves that it is equivalent that $Q(t) \rightarrow \text{id}_E$ in the strong operator topology as $t \downarrow 0$, i.e. for each $x \in E$, $Q(t)x \rightarrow x$.⁸

We now establish that the $\{T_t : t \geq 0\}$ is indeed a one-parameter semigroup and that it is strongly continuous.⁹

Theorem 13. *Suppose μ is a Gaussian measure on \mathbb{R}^n with mean 0 and let $p \geq 1$. Then $\{T_t : t \geq 0\}$ is a strongly continuous one-parameter semigroup in $\mathcal{B}(L^p(\gamma))$.*

Proof. For $f \in L^p(\gamma)$, because γ is a probability measure,

$$T_0(f)(u) = \int_{\mathbb{R}^n} f(u) d\gamma(v) = f(u),$$

hence $T_0 = \text{id}_{L^p(\mu)}$. For $s, t \geq 0$, define $P : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ by

$$P(u, v) = e^{-s} \frac{\sqrt{1 - e^{-2t}}}{\sqrt{1 - e^{-2t-2s}}} u + \frac{\sqrt{1 - e^{-2s}}}{\sqrt{1 - e^{-2t-2s}}} v.$$

⁸Walter Rudin, *Functional Analysis*, second ed., p. 376, Theorem 13.35.

⁹Vladimir I. Bogachev, *Gaussian Measures*, p. 10, Theorem 1.4.1.

By Theorem 6, $P_*(\gamma \times \gamma) = \gamma$, whence

$$\begin{aligned}
& (T_t(T_s f))(x) \\
&= \int_{\mathbb{R}^n} (T_s f) \left(e^{-t}x + \sqrt{1 - e^{-2t}}y \right) d\gamma(y) \\
&= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f \left(e^{-s} \left(e^{-t}x + \sqrt{1 - e^{-2t}}y \right) + \sqrt{1 - e^{-2s}}w \right) d\gamma(w) \right) d\gamma(y) \\
&= \int_{\mathbb{R}^n \times \mathbb{R}^n} f \left(e^{-s-t}x + \sqrt{1 - e^{-2t-2s}}P(y, w) \right) d(\gamma \times \gamma)(y, w) \\
&= \int_{\mathbb{R}^n \times \mathbb{R}^n} (f \circ M_{s+t})(x, P(y, w)) d(\gamma \times \gamma)(y, w) \\
&= \int_{\mathbb{R}^n} (f \circ M_{s+t})(x, z) d\gamma(z) \\
&= T_{s+t}(f)(x),
\end{aligned}$$

hence $T_t \circ T_s = T_{s+t}$. This establishes that $\{T_t : t \geq 0\}$ is a semigroup.

For $f \in C_b(\mathbb{R}^n)$, $u \in \mathbb{R}^n$, and $v \in \mathbb{R}^n$, as $t \downarrow 0$ we have

$$f \left(e^{-t}u + \sqrt{1 - e^{-2t}}v \right) - f(u) \rightarrow 0,$$

thus by the dominated convergence theorem, since

$$\left| f \left(e^{-t}u + \sqrt{1 - e^{-2t}}v \right) - f(u) \right| \leq 2 \|f\|_\infty$$

and γ is a probability measure, we have

$$\int_{\mathbb{R}^n} \left(f \left(e^{-t}u + \sqrt{1 - e^{-2t}}v \right) - f(u) \right) d\gamma(v) \rightarrow 0,$$

and hence

$$\begin{aligned}
(T_t f - T_0 f)(u) &= \int_{\mathbb{R}^n} f \left(e^{-t}u + \sqrt{1 - e^{-2t}}v \right) d\gamma(v) - \int_{\mathbb{R}^n} f(u) d\gamma(v) \\
&= \int_{\mathbb{R}^n} \left(f \left(e^{-t}u + \sqrt{1 - e^{-2t}}v \right) - f(u) \right) d\gamma(v) \\
&\rightarrow 0.
\end{aligned}$$

Because this is true for each $u \in \mathbb{R}^n$ and

$$|(T_t f - T_0 f)(u)| \leq \int_{\mathbb{R}^n} 2 \|f\|_\infty d\gamma(v) = 2 \|f\|_\infty,$$

by the dominated convergence theorem we then have

$$\|T_t f - T_0 f\|_{L^p(\gamma)} \rightarrow 0. \tag{7}$$

Now let $f \in L^p(\gamma)$. There is a sequence $f_j \in C_b(\mathbb{R}^n)$ satisfying $\|f_j - f\|_{L^p(\gamma)} \rightarrow 0$, with $\|f_j\|_{L^p(\gamma)} \leq 2\|f\|_{L^p(\gamma)}$ for all j . For any $t \geq 0$,

$$\begin{aligned}\|T_t f - T_0 f\|_{L^p(\gamma)} &\leq \|T_t f - T_t f_j\|_{L^p(\gamma)} + \|T_t f_j - T_0 f_j\|_{L^p(\gamma)} + \|T_0 f_j - T_0 f\|_{L^p(\gamma)} \\ &= \|T_t(f - f_j)\|_{L^p(\gamma)} + \|T_t f_j - T_0 f_j\|_{L^p(\gamma)} + \|f_j - f\|_{L^p(\gamma)} \\ &\leq \|f - f_j\|_{L^p(\gamma)} + \|T_t f - T_0 f_j\|_{L^p(\gamma)} + \|f_j - f\|_{L^p(\gamma)}.\end{aligned}$$

Let $\epsilon > 0$ and let j be so large that $\|f - f_j\|_{L^p(\gamma)} < \epsilon$. Because $f_j \in C_b(\mathbb{R}^n)$, by (7) there is some $\delta > 0$ such that when $0 < t < \delta$, $\|T_t f_j - f_j\|_{L^p(\gamma)} < \epsilon$. Then when $0 < t < \delta$,

$$\|T_t f - T_0 f\|_{L^p(\gamma)} \leq \epsilon + \epsilon + \epsilon,$$

which shows that for each $f \in L^p(\gamma)$, $\|T_t f - T_0 f\|_{L^p(\gamma)}$ as $t \downarrow 0$, which suffices to establish that $\{T_t : t \geq 0\}$ is strongly continuous $[0, \infty) \rightarrow \mathcal{B}(L^p(\gamma))$. \square

For $t > 0$, we define $L_t \in \mathcal{B}(L^p(\gamma))$ by

$$L_t f = \frac{1}{t}(T_t f - f), \quad f \in L^p(\gamma).$$

We define $\mathcal{D}(L)$ to be the set of those $f \in L^p(\gamma)$ such that $L_t f$ converges to some element of $L^p(\gamma)$ as $t \downarrow 0$, and we define $L : \mathcal{D}(L) \rightarrow L^p(\gamma)$. This is the **infinitesimal generator** of the semigroup $\{T_t : t \geq 0\}$, and the infinitesimal generator L of the Ornstein-Uhlenbeck semigroup is called the **Ornstein-Uhlenbeck operator**. Because the Ornstein-Uhlenbeck semigroup is strongly continuous, we get the following.¹⁰

Theorem 14. *Suppose μ is a Gaussian measure on \mathbb{R}^n with mean 0, let $p \geq 1$, and let L be the infinitesimal generator of the Ornstein-Uhlenbeck semigroup $\{T_t : t \geq 0\}$. Then:*

1. $\mathcal{D}(L)$ is a dense linear subspace of $L^p(\gamma)$ and $L : \mathcal{D}(L) \rightarrow L^p(\gamma)$ is a closed operator.
2. For each $f \in \mathcal{D}(L)$ and for each $t \geq 0$,

$$\frac{d}{dt}(T_t f) = (L \circ T_t)f = (T_t \circ L)f.$$

3. For $f \in L^p(\gamma)$ and K a compact subset of $[0, \infty)$, $(\exp(tL_\epsilon)f \rightarrow T_t f$ as $\epsilon \downarrow 0$ uniformly for $t \in K$.
4. For $\lambda \in \mathbb{C}$ with $\operatorname{Re} \lambda > 0$, $R(\lambda) : L^p(\gamma) \rightarrow L^p(\gamma)$ defined by

$$R(\lambda)f = \int_0^\infty e^{-\lambda t} T_t f dt, \quad f \in L^p(\gamma),$$

¹⁰Walter Rudin, *Functional Analysis*, second ed., p. 376, Theorem 13.35.

belongs to $\mathcal{B}(L^p(\gamma))$, the range of $R(\lambda)$ is equal to $\mathcal{D}(L)$, and

$$((\lambda I - L) \circ R(\lambda))f = f, \quad f \in L^p(\gamma), \quad (R(\lambda) \circ (\lambda I - L))f = \mathcal{D}(L),$$

where I is the identity operator on $L^p(\gamma)$.

We remind ourselves that if H is a Hilbert space with inner product $\langle \cdot, \cdot \rangle$, an element A of $\mathcal{B}(H)$ is said to be a **positive operator** when $\langle Ax, x \rangle \geq 0$ for all $x \in H$. We prove that each T_t is a positive operator on the Hilbert space $L^2(\gamma)$.¹¹

Theorem 15. *Suppose μ is a Gaussian measure on \mathbb{R}^n with mean 0. For each $t \geq 0$, $T_t \in \mathcal{B}(L^2(\mu))$ is a positive operator.*

Proof. For $t \geq 0$, define $N_t : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^n$ by

$$O_t(x, y) = \left(e^{-t}x + \sqrt{1 - e^{-2t}}y, -\sqrt{1 - e^{-2t}}x + e^{-t}y \right), \quad (x, y) \in \mathbb{R}^n \times \mathbb{R}^n,$$

whose transpose is the linear operator $N_t^* : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^n$ defined by

$$O_t^*(u, v) = \left(e^{-t}u - \sqrt{1 - e^{-2t}}v, \sqrt{1 - e^{-2t}}u + e^{-t}v \right), \quad (u, v) \in \mathbb{R}^n \times \mathbb{R}^n.$$

For $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, we calculate

$$\begin{aligned} & \int_{\mathbb{R}^n \times \mathbb{R}^n} e^{i\langle (x, y), (u, v) \rangle} d(O_{t*}(\gamma \times \gamma))(u, v) \\ &= \int_{\mathbb{R}^n \times \mathbb{R}^n} e^{i\langle (x, y), O_t(u, v) \rangle} d(\gamma \times \gamma)(u, v) \\ &= \int_{\mathbb{R}^n \times \mathbb{R}^n} e^{i\langle O_t^*(x, y), (u, v) \rangle} d(\gamma \times \gamma)(u, v) \\ &= \widetilde{\gamma \times \gamma}(O_t^*(x, y)) \\ &= \widetilde{\gamma \times \gamma}(e^{-t}x - \sqrt{1 - e^{-2t}}y, \sqrt{1 - e^{-2t}}x + e^{-t}y) \\ &= \widetilde{\gamma}(e^{-t}x - \sqrt{1 - e^{-2t}}y) \widetilde{\gamma}(\sqrt{1 - e^{-2t}}x + e^{-t}y) \\ &= \exp \left(-\frac{1}{2} \left\langle K(e^{-t}x - \sqrt{1 - e^{-2t}}y), e^{-t}x - \sqrt{1 - e^{-2t}}y \right\rangle \right) \\ &\quad \cdot \exp \left(-\frac{1}{2} \left\langle K(\sqrt{1 - e^{-2t}}x + e^{-t}y), \sqrt{1 - e^{-2t}}x + e^{-t}y \right\rangle \right) \\ &= \exp \left(-\frac{1}{2} \langle Kx, x \rangle - \frac{1}{2} \langle Ky, y \rangle \right) \\ &= \widetilde{\gamma}(x) \widetilde{\gamma}(y) \\ &= \widetilde{\gamma \times \gamma}(x, y), \end{aligned}$$

¹¹Vladimir I. Bogachev, *Gaussian Measures*, p. 10, Theorem 1.4.1.

which shows that $O_{t*}(\gamma \times \gamma)$ and $\gamma \times \gamma$ have equal characteristic functions and hence are themselves equal.

For $f, g \in L^2(\gamma)$ and $t \geq 0$,

$$\begin{aligned}
\langle T_t f, g \rangle_{L^2(\gamma)} &= \int_{\mathbb{R}^n} (T_t f)(x) g(x) d\mu(x) \\
&= \int_{\mathbb{R}^n \times \mathbb{R}^n} f\left(e^{-t}x + \sqrt{1 - e^{-2t}}y\right) g(x) d(\gamma \times \gamma)(x, y) \\
&= \int_{\mathbb{R}^n \times \mathbb{R}^n} (f \circ \pi_1 \circ O_t)(x, y) (g \circ \pi_1 \circ O_t^{-1})(x, y) d(\gamma \times \gamma)(x, y) \\
&= \int_{\mathbb{R}^n \times \mathbb{R}^n} (f \circ \pi_1)(u, v) (g \circ \pi_1 \circ O_t^{-1})(u, v) d(O_{t*}(\gamma \times \gamma))(u, v) \\
&= \int_{\mathbb{R}^n \times \mathbb{R}^n} (f \circ \pi_1)(u, v) (g \circ \pi_1 \circ O_t^{-1})(u, v) d(\gamma \times \gamma)(u, v) \\
&= \int_{\mathbb{R}^n \times \mathbb{R}^n} f(u) g\left(e^{-t}u - \sqrt{1 - e^{-2t}}v\right) d(\gamma \times \gamma)(u, v) \\
&= \int_{\mathbb{R}^n} f(u) \left(\int_{\mathbb{R}^n} g(M_t(u, -v)) d\gamma(v) \right) d\gamma(u) \\
&= \int_{\mathbb{R}^n} f(u) \left(\int_{\mathbb{R}^n} g(M_t(u, v)) d\gamma(v) \right) d\gamma(u) \\
&= \int_{\mathbb{R}^n} f(u) (T_t g)(u) d\gamma(u) \\
&= \langle f, T_t g \rangle_{L^2(\gamma)},
\end{aligned}$$

which establishes that T_t is a self-adjoint operator on $L^2(\gamma)$.

Furthermore, using that $T_t = T_{t/2} \circ T_{t/2}$ and that $T_{t/2}$ is self-adjoint,

$$\langle T_t f, f \rangle_{L^2(\gamma)} = \langle T_{t/2} T_{t/2} f, f \rangle_{L^2(\gamma)} = \left\langle T_{t/2} f, T_{t/2}^* f \right\rangle_{L^2(\gamma)} = \langle T_{t/2} f, T_{t/2} f \rangle_{L^2(\gamma)},$$

which is ≥ 0 , which establishes that T_t is a positive operator on $L^2(\gamma)$. \square

We now write the Ornstein-Uhlenbeck semigroup using the orthogonal projections $I_k : L^2(\gamma_n) \rightarrow \mathcal{X}_k$, where γ_n is the standard Gaussian measure on \mathbb{R}^n .¹²

Theorem 16. For each $t \geq 0$ and $f \in L^2(\gamma_n)$,

$$T_t f = \sum_{k=0}^{\infty} e^{-kt} I_k(f).$$

Proof. Define $S_t : L^2(\gamma_n) \rightarrow L^2(\gamma_n)$ by $S_t f = \sum_{k=0}^{\infty} e^{-kt} I_k(f)$, which satisfies, using that the subspaces \mathcal{X}_k are pairwise orthogonal,

$$\|S_t f\|_{L^2(\gamma_n)}^2 = \sum_{k=0}^{\infty} e^{-kt} \|I_k(f)\|_{L^2(\gamma_n)}^2 \leq \sum_{k=0}^{\infty} \|I_k(f)\|_{L^2(\gamma_n)}^2 = \|f\|_{L^2(\gamma_n)}^2,$$

¹²Vladimir I. Bogachev, *Gaussian Measures*, p. 11, Theorem 1.4.4.

so $S_t \in \mathcal{B}(L^2(\gamma_n))$. To prove that $T_t = S_t$, it suffices to prove that $T_t H_\alpha = S_t H_\alpha$ for each Hermite polynomial, which are an orthonormal basis for $L^2(\gamma_n)$. For $\alpha = (k_1, \dots, k_n)$ with $k = |\alpha| = k_1 + \dots + k_n$,

$$S_t H_\alpha = e^{-kt} H_\alpha,$$

and

$$\begin{aligned} (T_t H_\alpha)(x) &= \int_{\mathbb{R}^n} H_\alpha \left(e^{-t}x + \sqrt{1 - e^{-2t}}y \right) d\gamma_n(y) \\ &= \int_{\mathbb{R}^n} \prod_{j=1}^n H_{k_j} \left(e^{-t}x_j + \sqrt{1 - e^{-2t}}y_j \right) d\gamma_n(y) \\ &= \prod_{j=1}^n \int_{\mathbb{R}} H_{k_j} \left(e^{-t}x_j + \sqrt{1 - e^{-2t}}y_j \right) d\gamma_1(y_j). \end{aligned}$$

To prove that $T_t H_\alpha = e^{-kt} H_\alpha$, it thus suffices to prove that for any t , for any k_j , and for any x_j ,

$$\int_{\mathbb{R}} H_{k_j} \left(e^{-t}x_j + \sqrt{1 - e^{-2t}}y_j \right) d\gamma_1(y_j) = e^{-k_j t} H_{k_j}(x_j). \quad (8)$$

For $k_j = 0$, as $H_0 = 1$ and γ_1 is a probability measure, (8) is true. Suppose that (8) is true for $\leq k_j$. That is, for each $0 \leq h \leq k_j$, $T_t H_h = e^{-ht} H_h$. For any l , because the Hermite polynomial H_l is a polynomial of degree l , one checks that $T_t H_l(x_j)$ is a polynomial of degree l : using the binomial formula,

$$\int_{\mathbb{R}} (e^{-t}x_j + \sqrt{1 - e^{-2t}}y_j)^l \exp\left(-\frac{y_j^2}{2}\right) d\gamma_1(y_j)$$

is a polynomial in x_j of degree l . Hence $T_t H_l$ a linear combination of H_0, H_1, \dots, H_l . For $0 \leq h \leq k_j$,

$$\langle T_t H_{k_j+1}, H_h \rangle_{L^2(\gamma_1)} = \langle H_{k_j+1}, T_t H_h \rangle_{L^2(\gamma_1)} = \langle H_{k_j+1}, e^{-ht} H_h \rangle_{L^2(\gamma_1)} = 0.$$

Therefore there is some $c \in \mathbb{R}$ such that $T_t H_{k_j+1} = c H_{k_j+1}$. Then check that $c = e^{-(k_j+1)t}$. \square

We now give an explicit expression for the domain $\mathcal{D}(L)$ of the Ornstein-Uhlenbeck operator L and for L applied to an element of its domain.¹³

Theorem 17.

$$\mathcal{D}(L) = \left\{ f \in L^2(\gamma_n) : \sum_{k=0}^{\infty} k^2 \|I_k(f)\|_{L^2(\gamma_n)}^2 < \infty \right\}.$$

For $f \in \mathcal{D}(L)$,

$$Lf = - \sum_{k=0}^{\infty} k I_k(f).$$

¹³Vladimir I. Bogachev, *Gaussian Measures*, p. 12, Proposition 1.4.5.

Proof. Let $f \in \mathcal{D}(L)$, i.e. $\frac{T_t f - f}{t} \rightarrow Lf$ in $L^2(\gamma_n)$ as $t \downarrow 0$. For any $k \geq 0$, using Theorem 16,

$$\begin{aligned}
I_k Lf &= I_k \left(\lim_{t \downarrow 0} \frac{T_t f - f}{t} \right) \\
&= \lim_{t \downarrow 0} \frac{I_k T_t f - I_k f}{t} \\
&= \lim_{t \downarrow 0} \frac{T_t I_k f - I_k f}{t} \\
&= \lim_{t \downarrow 0} \frac{e^{-kt} I_k f - I_k f}{t} \\
&= \left(\lim_{t \downarrow 0} \frac{e^{-kt} - 1}{t} \right) I_k f \\
&= (e^{-kt})' \big|_{t=0} I_k f \\
&= -k I_k f.
\end{aligned}$$

Using this,

$$\begin{aligned}
\sum_{k=0}^{\infty} k^2 \|I_k f\|_{L^2(\gamma_n)}^2 &= \sum_{k=0}^{\infty} \|I_k Lf\|_{L^2(\gamma_n)}^2 \\
&= \left\| \sum_{k=0}^{\infty} I_k Lf \right\|_{L^2(\gamma_n)}^2 \\
&= \|Lf\|_{L^2(\gamma_n)}^2 \\
&< \infty.
\end{aligned}$$

Moreover,

$$Lf = L \left(\sum_{k=0}^{\infty} I_k f \right) = \sum_{k=0}^{\infty} L I_k f = \sum_{k=0}^{\infty} I_k Lf = \sum_{k=0}^{\infty} -k I_k f.$$

Let $f \in L^2(\gamma_n)$ satisfy

$$\sum_{k=0}^{\infty} k^2 \|I_k f\|_{L^2(\gamma_n)}^2 < \infty.$$

For $t > 0$,

$$\begin{aligned}
\left\| \frac{T_t f - f}{t} + \sum_{k=0}^{\infty} k I_k f \right\|_{L^2(\gamma_n)}^2 &= \left\| \sum_{k=0}^{\infty} \left(\frac{e^{-kt} I_k f - I_k f}{t} + k I_k f \right) \right\|_{L^2(\gamma_n)}^2 \\
&= \sum_{k=0}^{\infty} \left| \frac{e^{-kt} - 1}{t} + k \right|^2 \|I_k f\|_{L^2(\gamma_n)}^2.
\end{aligned}$$

For $t > 0$ and $k \geq 0$,

$$|t^{-1}(e^{-kt} - 1)| \leq k,$$

and thus

$$\sum_{k=0}^{\infty} \left| \frac{e^{-kt} - 1}{t} + k \right|^2 \|I_k f\|_{L^2(\gamma_n)}^2 \leq \sum_{k=0}^{\infty} (2k)^2 \|I_k f\|_{L^2(\gamma_n)}^2 < \infty.$$

For each $k \geq 0$, as $t \downarrow 0$,

$$\frac{e^{-kt} - 1}{t} + k \rightarrow 0,$$

thus as $t \downarrow 0$,

$$\sum_{k=0}^{\infty} \left| \frac{e^{-kt} - 1}{t} + k \right|^2 \|I_k f\|_{L^2(\gamma_n)}^2 \rightarrow 0$$

and hence

$$\left\| \frac{T_t f - f}{t} + \sum_{k=0}^{\infty} k I_k f \right\|_{L^2(\gamma_n)}^2 \rightarrow 0.$$

This means that $\frac{T_t f - f}{t}$ converges in $L^2(\gamma_n)$ to $-\sum_{k=0}^{\infty} k I_k f$ as $t \downarrow 0$, and since $\frac{T_t f - f}{t}$ converges, $f \in \mathcal{D}(L)$. \square