SPDE's Class Note

Hong Sheng Tan

January 11, 2026

Contents

Acknowledgements **2**

1 Gaussian Measures Theory **2**
- 1.1 Definitions and Basic Properties 2
- 1.2 Fernique's Theorem 5
- 1.3 Kolmogorov's Continuity Theorem 11
- 1.4 Cameron-Martin Space 14
- 1.5 Images of Gaussian measures under linear maps 24
- 1.6 Cylindrical Wiener processes and stochastic integration 29

2 Semigroup Theory **38**
- 2.1 Strongly continuous semigroups 38
- 2.2 Adjoint semigroups on dual spaces 50
- 2.3 Analytic semigroups 57
- 2.4 Interpolation Spaces 67

3 Linear SPDEs **81**
- 3.1 Space-time regularity of solutions 83
- 3.2 Markov semigroups and invariant measures 93

4 Semilinear SPDEs **106**
- 4.1 Local solutions 106
- 4.2 Sobelev embedding theorem 113

Acknowledgements

This note is a summary of the Stochastic PDEs reading course with Professor Konstantin Matetski at Michigan State University, conducted from Summer 2025 to Fall 2025.

The primary reference is [Hairer(2009)]. Most of the material (about 95%) follows this reference closely. The remaining portions (about 5%) consist of the author's own additions, written in an attempt to bridge gaps in his own understanding, particularly in probability and functional analysis.

1 Gaussian Measures Theory

This section is devoted to the theory of Gaussian measures on Banach spaces. In most cases, we will work with a separable Banach space, denoted by $\mathcal{B}$. This means that $\mathcal{B}$ contains a countable dense subset.

1.1 Definitions and Basic Properties

Recall the following definition of Gaussian measures on $\mathbb{R}$.

Definition 1.1. A probability measure μ on $\mathbb{R}$ is called a **Gaussian measure** if either

(i) $\mu = \delta_m$ for some $m \in \mathbb{R}$, or

(ii) μ is absolutely continuous with respect to the Lebesgue measure, with density
$$f(x) = \frac{1}{\sqrt{2\pi q}} \exp\left(-\frac{(x-m)^2}{2q}\right),$$
for some $m \in \mathbb{R}$ and $q > 0$.

In the second case, we write $\mu = \mathcal{N}(m, q)$, where m is the mean and q is the variance. A Gaussian measure is called **centered** if $m = 0$.

A Gaussian measure on $\mathbb{R}^n$ is characterized by the property that every one-dimensional projection is Gaussian. This viewpoint extends naturally to Banach spaces.

Let $(X_1, \mathcal{A}_1)$ and $(X_2, \mathcal{A}_2)$ be measurable spaces, $T : X_1 \to X_2$ measurable, and μ a measure on $(X_1, \mathcal{A}_1)$. The **push-forward measure** $T^\sharp\mu$ on $(X_2, \mathcal{A}_2)$ is defined by
$$T^\sharp\mu(A) = \mu\big(T^{-1}(A)\big), \qquad A \in \mathcal{A}_2.$$

We have also the following change-of-variables formula: for every measurable function $f : X_2 \to \mathbb{C}$,

$$\int_{X_2} f(y)\, T^\sharp\mu(dy) = \int_{X_1} f\big(T(x)\big)\, \mu(dx).$$

Now, let us define Gaussian measures on Banach spaces.

Definition 1.2. A probability measure μ on a separable Banach space $\mathcal{B}$ is called a **Gaussian measure** if for every $\ell \in \mathcal{B}^*$, the push-forward measure $\ell^\sharp\mu$ is a Gaussian measure on $\mathbb{R}$. We call μ **centered** if $\ell^\sharp\mu$ is centered for every $\ell \in \mathcal{B}^*$.

Throughout these notes, we assume that all Gaussian measures are centered unless stated otherwise. For this definition to be well posed, we need to know that the family $(\ell^\sharp\mu)_{\ell\in\mathcal{B}^*}$ determines the measure μ. This is a consequence of the following theorem, whose proof will be omitted.

Theorem 1.3. *Let μ and ν be two probability measures on a separable Banach space $\mathcal{B}$. If $\ell^\sharp\mu = \ell^\sharp\nu$ for every $\ell \in \mathcal{B}^*$, then $\mu = \nu$.*

The covariance operator C_μ is the infinite-dimensional analogue of a covariance matrix: it records the covariance of every pair of linear observations $\ell_1(X)$ and $\ell_2(X)$.

Definition 1.4. Let μ be a centered Gaussian measure on $\mathcal{B}$. The **covariance operator** $C_\mu : \mathcal{B}^* \times \mathcal{B}^* \to \mathbb{R}$ is defined by

$$C_\mu(\ell_1, \ell_2) = \int_{\mathcal{B}} \ell_1(x)\ell_2(x)\, \mu(dx), \qquad \ell_1, \ell_2 \in \mathcal{B}^*.$$

Remark. Given a Gaussian measure μ on $\mathcal{B}$, let X be a random variable with distribution μ. For every $\ell_1, \ell_2 \in \mathcal{B}^*$, we have

$$C_\mu(\ell_1, \ell_2) = \mathbb{E}[\ell_1(X)\ell_2(X)].$$

Remark. We can identify C_μ with an operator $\hat{C}_\mu : \mathcal{B}^* \to \mathcal{B}^{**}$ by setting

$$\hat{C}_\mu(\ell_1)(\ell_2) = C_\mu(\ell_1, \ell_2), \qquad \ell_1, \ell_2 \in \mathcal{B}^*.$$

In the case where $\mathcal{B} = \mathbb{R}^n$, the Fourier transform of a centered Gaussian measure μ is given by

$$\hat{\mu}(x) = \int_{\mathbb{R}^n} e^{i\langle x,y\rangle}\,\mu(dy), \qquad x \in \mathbb{R}^n.$$

This formula extends naturally to Banach spaces as follows.

Definition 1.5. Let μ be a centered Gaussian measure on a separable Banach space $\mathcal{B}$. The **Fourier transform** of μ is the function $\hat{\mu} : \mathcal{B}^* \to \mathbb{C}$ defined by

$$\hat{\mu}(\ell) = \int_{\mathcal{B}} e^{i\ell(x)}\,\mu(dx), \qquad \ell \in \mathcal{B}^*.$$

The Fourier transform of a centered Gaussian measure on a Banach space has the following explicit form.

Theorem 1.6. *Let μ be a centered Gaussian measure on a separable Banach space $\mathcal{B}$. Then, for every $\ell \in \mathcal{B}^*$,*

$$\hat{\mu}(\ell) = \exp\left(-\frac{1}{2}C_\mu(\ell,\ell)\right). \tag{1.1}$$

Proof. Let us first consider the case where $\mathcal{B} = \mathbb{R}$ and $\mu = \mathcal{N}(0,q)$. Let X be a random variable with distribution μ. Then

$$\begin{aligned}
\mathbb{E}[e^{iX}] &= \int_{\mathbb{R}} e^{ix}\frac{1}{\sqrt{2\pi q}}\exp\left(-\frac{x^2}{2q}\right)dx \\
&= e^{-\frac{q}{2}}\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi q}}\exp\left(-\frac{(x-iq)^2}{2q}\right)dx \\
&= e^{-\frac{q}{2}}.
\end{aligned}$$

Now let $\mathcal{B}$ be a separable Banach space and let μ be a centered Gaussian measure on $\mathcal{B}$. Fix $\ell \in \mathcal{B}^*$, and let X be a random variable with distribution μ. By definition, the random variable $\ell(X)$ has distribution $\ell^\sharp\mu = \mathcal{N}(0,q)$. Note that

$$C_\mu(\ell,\ell) = \int_{\mathcal{B}} \ell(x)^2\,\mu(dx) = \mathbb{E}[\ell(X)^2] = q.$$

Thus,

$$\hat{\mu}(\ell) = \mathbb{E}[e^{i\ell(X)}] = e^{-\frac{q}{2}} = \exp\left(-\frac{1}{2}C_\mu(\ell,\ell)\right).$$

This completes the proof. $\square$

Remark. It is well known that if $X \sim \mathcal{N}(0, q)$, then for every $t \in \mathbb{R}$,

$$\mathbb{E}[e^{tX}] = e^{\frac{qt^2}{2}}.$$

The proof of this fact is similar to that of Theorem 1.6.

The Fourier transform of a probability measure on $\mathcal{B}$ uniquely determines the measure.

Theorem 1.7. *Let μ and ν be two probability measures on a separable Banach space $\mathcal{B}$. If $\hat{\mu}(\ell) = \hat{\nu}(\ell)$ for every $\ell \in \mathcal{B}^*$, then $\mu = \nu$.*

1.2 Fernique's Theorem

This section is devoted to Fernique's theorem. It shows that a Gaussian measure on a separable Banach space satisfies an exponential integrability estimate. As a consequence, Gaussian measures have exponentially decaying tails and finite moments of all orders.

We need the following lemma.

Lemma 1.8. *Let μ be a centered Gaussian measure on $\mathcal{B}$. For every $\varphi \in \mathbb{R}$, define the rotation $R_\varphi : \mathcal{B}^2 \to \mathcal{B}^2$ by*

$$R_\varphi(x, y) = \big(x \sin\varphi + y \cos\varphi,\ x \cos\varphi - y \sin\varphi\big).$$

Then,

$$R_\varphi^\sharp(\mu \otimes \mu) = \mu \otimes \mu.$$

Proof. By Theorem 1.7, it suffices to show that for every $(\ell_1, \ell_2) \in (\mathcal{B}^*)^2$, one has

$$\widehat{\mu \otimes \mu}(\ell_1, \ell_2) = \widehat{R_\varphi^\sharp(\mu \otimes \mu)}(\ell_1, \ell_2).$$

For $(x, y) \in \mathcal{B}^2$, denote $(x', y') = R_\varphi(x, y)$. For $(\ell_1, \ell_2) \in (\mathcal{B}^*)^2$, set

$$\ell_1' = \ell_1 \sin\varphi + \ell_2 \cos\varphi, \qquad \ell_2' = \ell_1 \cos\varphi - \ell_2 \sin\varphi.$$

By Theorem 1.6,

$$\begin{aligned}\widehat{\mu \otimes \mu}(\ell_1, \ell_2) &= \int_{\mathcal{B}^2} e^{i(\ell_1(x) + \ell_2(y))}\, \mu(dx)\mu(dy) \\ &= \hat{\mu}(\ell_1)\hat{\mu}(\ell_2) = \exp\left(-\frac{1}{2}C_\mu(\ell_1, \ell_1) - \frac{1}{2}C_\mu(\ell_2, \ell_2)\right).\end{aligned}$$

Similarly,

$$
\begin{aligned}
R_\varphi^\sharp\widehat{(\mu\otimes\mu)}(\ell_1,\ell_2) &= \int_{\mathcal{B}^2} e^{i(\ell_1(x')+\ell_2(y'))}\,\mu(dx)\mu(dy)\\
&= \int_{\mathcal{B}^2} e^{i(\ell_1'(x)+\ell_2'(y))}\,\mu(dx)\mu(dy)\\
&= \exp\left(-\frac{1}{2}C_\mu(\ell_1',\ell_1')-\frac{1}{2}C_\mu(\ell_2',\ell_2')\right).
\end{aligned}
$$

A direct computation using bilinearity and symmetry of C_μ shows that

$$
C_\mu(\ell_1,\ell_1)+C_\mu(\ell_2,\ell_2)=C_\mu(\ell_1',\ell_1')+C_\mu(\ell_2',\ell_2').
$$

Therefore the Fourier transforms coincide, which concludes the proof. □

The next theorem is Fernique's theorem. It states that a Gaussian measure on a separable Banach space satisfies an exponential integrability estimate.

Theorem 1.9 (Fernique's Theorem)**.** *Let μ be a centered Gaussian measure on a separable Banach space $\mathcal{B}$. Then, there exists $\alpha>0$ such that*

$$
\int_{\mathcal{B}} \exp\left(\alpha\|x\|^2\right)\mu(dx)<\infty.
$$

Proof. Let $t>\tau\geqslant 0$ and set

$$
A=\{(x,y)\in\mathcal{B}^2 : \|x\|>t,\ \|y\|\leqslant\tau\}.
$$

Applying Lemma 1.8 with $\varphi=\pi/4$, we obtain

$$
\begin{aligned}
(\mu\otimes\mu)(A) &= (\mu\otimes\mu)(R_{\pi/4}^{-1}(A))\\
&= (\mu\otimes\mu)\Big\{(x,y) : \Big\|\frac{x+y}{\sqrt{2}}\Big\|>t,\ \Big\|\frac{x-y}{\sqrt{2}}\Big\|\leqslant\tau\Big\}.
\end{aligned}
$$

Equivalently,

$$
(\mu\otimes\mu)(A)=\int_{\|x+y\|>\sqrt{2}t}\int_{\|x-y\|\leqslant\sqrt{2}\tau}\mu(dx)\mu(dy).
$$

By the triangle inequality,

$$
\min\{\|x\|,\|y\|\}\geqslant\frac{1}{2}\left(\|x+y\|-\|x-y\|\right).
$$

Hence, on the set $\{\|x+y\| > \sqrt{2}t,\ \|x-y\| \leqslant \sqrt{2}\tau\}$, we have

$$\|x\| > \frac{t-\tau}{\sqrt{2}} \quad \text{and} \quad \|y\| > \frac{t-\tau}{\sqrt{2}}.$$

Therefore,

$$(\mu \otimes \mu)(A) \leqslant \int_{\|x\| > \frac{t-\tau}{\sqrt{2}}} \int_{\|y\| > \frac{t-\tau}{\sqrt{2}}} \mu(dx)\mu(dy) = \mu\left(\|x\| > \frac{t-\tau}{\sqrt{2}} \right)^2.$$

On the other hand,

$$(\mu \otimes \mu)(A) = \mu(\|x\| > t)\, \mu(\|x\| \leqslant \tau).$$

Combining these two bounds yields

$$\mu(\|x\| > t)\, \mu(\|x\| \leqslant \tau) \leqslant \mu\left(\|x\| > \frac{t-\tau}{\sqrt{2}} \right)^2. \tag{1.2}$$

Choose $\tau > 0$ such that $\mu(\|x\| \leqslant \tau) \geqslant 3/4$. Set $t_0 = \tau$ and define recursively

$$t_{n+1} = \sqrt{2}\, t_n + \tau, \qquad n \geqslant 0.$$

Applying (1.2) with $t = t_{n+1}$ gives

$$\mu(\|x\| > t_{n+1}) \leqslant \frac{1}{\mu(\|x\| \leqslant \tau)}\, \mu\left(\|x\| > \frac{t_{n+1}-\tau}{\sqrt{2}} \right)^2 \leqslant \frac{4}{3}\, \mu(\|x\| > t_n)^2.$$

Let

$$y_n = \frac{4}{3}\, \mu(\|x\| > t_n).$$

Then $y_{n+1} \leqslant y_n^2$ and, since $\mu(\|x\| > \tau) \leqslant 1/4$, we have $y_0 \leqslant 1/3$. Hence,

$$\mu(\|x\| > t_n) = \frac{3}{4} y_n \leqslant \frac{3}{4} y_0^{2^n} \leqslant \frac{3}{4}\, 3^{-2^n}.$$

Moreover, one checks that

$$t_n = \frac{(\sqrt{2})^{n+1} - 1}{\sqrt{2} - 1}\, \tau \leqslant 2^{n/2}(2 + \sqrt{2})\tau \leqslant 5 \cdot 2^{n/2}\tau.$$

In particular, $2^n \geqslant \frac{t_n^2}{25\tau^2}$, and therefore

$$\mu(\|x\| > t_n) \leqslant \frac{3}{4}\,3^{-2^n} \leqslant \frac{3}{4}\,3^{-t_n^2/(25\tau^2)} = \frac{3}{4}\exp\biggl(-\frac{\log 3}{25}\,\frac{t_n^2}{\tau^2}\biggr)\,.$$

Consequently, there exists $\alpha > 0$ such that for every $t \geqslant \tau$,

$$\mu(\|x\| > t) \leqslant \exp\biggl(-2\alpha\,\frac{t^2}{\tau^2}\biggr)\,.$$

Integrating by parts, we obtain

$$\begin{aligned}\int_{\mathcal{B}} \exp\biggl(\frac{\alpha\|x\|^2}{\tau^2}\biggr)\,\mu(dx) &\leqslant e^{\alpha} + \frac{2\alpha}{\tau^2}\int_{\tau}^{\infty} t\,e^{\alpha t^2/\tau^2}\,\mu(\|x\| > t)\,dt \\ &\leqslant e^{\alpha} + 2\alpha\int_{1}^{\infty} t\,e^{-\alpha t^2}\,dt < \infty. \qquad (1.3)\end{aligned}$$

□

There are a few important consequences of Fernique's theorem. Most notably, Gaussian measures on separable Banach spaces have finite moments of all orders, i.e., for every $p > 0$,

$$\int_{\mathcal{B}} \|x\|^p\,\mu(dx) < \infty.$$

In fact, the result can be stronger, we can bound the even moments in terms of the first moment.

Corollary 1.10. *There exist universal constants $\alpha, K > 0$ with the following properties. Let μ be a centred Gaussian measure on a separable Banach space $\mathcal{B}$ and let $f : \mathbf{R}_+ \to \mathbf{R}_+$ be any measurable function such that $f(x) \leqslant C_f \exp\bigl(\alpha x^2\bigr)$ for every $x \geqslant 0$. Define furthermore the first moment of μ by $M = \int_{\mathcal{B}} \|x\|\mu(dx)$. Then, one has the bound $\int_{\mathcal{B}} f(\|x\|/M)\mu(dx) \leqslant KC_f$.*

In particular, the higher moments of μ are bounded by $\int_{\mathcal{B}} \|x\|^{2n}\mu(dx) \leqslant n!K\alpha^{-n}M^{2n}$.

Proof. Consider $\tau = 4M$ in equation (1.3). Then, we have

$$K = \int_{\mathcal{B}} \exp\left(\frac{\alpha}{16}\frac{\|x\|^2}{M^2}\right)\mu(dx) < \infty.$$

Now, let $\alpha_0 = \frac{\alpha}{16}$. For any measurable function $f : \mathbf{R}_+ \to \mathbf{R}_+$ such that

$f(x) \leqslant C_f \exp\left(\alpha_0 x^2\right)$ for every $x \geqslant 0$, we have

$$\begin{aligned}\int_{\mathcal{B}} f(\|x\|/M)\mu(dx) &\leqslant C_f \int_{\mathcal{B}} \exp\left(\alpha_0 \frac{\|x\|^2}{M^2}\right)\mu(dx)\\ &= KC_f,\end{aligned}$$

which proves the corollary. □

The fact that second moments are finite implies that the covariance form C_μ is a bounded bilinear form on $\mathcal{B}^* \times \mathcal{B}^*$.

Corollary 1.11. *There exists a constant $\|C_\mu\| < \infty$ such that $C_\mu(\ell, \ell') \leqslant \|C_\mu\| \, \|\ell\| \, \|\ell'\|$ for any $\ell, \ell' \in \mathcal{B}^*$. Furthermore, the operator $\hat{C}_\mu$ is a continuous operator from $\mathcal{B}^*$ to $\mathcal{B}$.*

Proof. By definition of C_μ and Cauchy-Schwarz inequality, we have

$$\begin{aligned}C_\mu(\ell, \ell') &= \int_{\mathcal{B}} \ell(x)\ell'(x)\mu(dx)\\ &\leqslant \left(\int_{\mathcal{B}} \ell(x)^2\mu(dx)\right)^{1/2}\left(\int_{\mathcal{B}} \ell'(x)^2\mu(dx)\right)^{1/2}\\ &\leqslant \int_{\mathcal{B}} \|x\|^2\mu(dx)\|\ell\| \, \|\ell'\| \, .\end{aligned}$$

The fact that $\hat{C}_\mu$ is continuous follows directly from the boundedness of C_μ. To check the range of $\hat{C}_\mu$ is in $\mathcal{B}$, we just note that we have identity

$$\hat{C}_\mu(\ell) = \int_{\mathcal{B}} x\ell(x)\mu(dx),$$

and it is clear that $\|\hat{C}_\mu(\ell)\| < \infty$. □

When the Banach space $\mathcal{B}$ is a Hilbert space $\mathcal{H}$, the covariance operator $\hat{C}_\mu$ can be identified with a trace-class operator on $\mathcal{H}$. In this case, the covariance operator is defined by the relation $C_\mu(h,k) = \langle \hat{C}_\mu h, k\rangle$ for any $h, k \in \mathcal{H}$.

Corollary 1.12. *Let μ be a centred Gaussian measure on a separable Hilbert space $\mathcal{H}$. Then, the covariance operator $\hat{C}_\mu : \mathcal{H} \to \mathcal{H}$ is trace class and one has identity*

$$\int_{\mathcal{H}} \|x\|^2\mu(dx) = \mathrm{Tr}\left(\hat{C}_\mu\right).$$

Conversely, for every positive trace class symmetric operator K on $\mathcal{H}$, there exists a centred Gaussian measure μ on $\mathcal{H}$ such that $\hat{C}_\mu = K$.

Proof. Fix an orthonormal basis $\{e_n\}_{n\in\mathbf{N}}$ of $\mathcal{H}$. By Fernique's theorem, the second moment of μ is finite and we have

$$\begin{aligned}\operatorname{Tr}\left(\hat{C}_\mu\right) &= \sum_{n\in\mathbf{N}} \left\langle \hat{C}_\mu e_n, e_n \right\rangle = \sum_{n\in\mathbf{N}} C_\mu\left(e_n, e_n\right) \\ &= \sum_{n\in\mathbf{N}} \int_{\mathcal{H}} \left\langle e_n, x\right\rangle^2 \mu(dx) = \int_{\mathcal{H}} \|x\|^2 \mu(dx).\end{aligned}$$

Conversely, let K be a positive trace class symmetric operator on $\mathcal{H}$. By the spectral theorem, there exists an orthonormal basis $\{e_n\}_{n\in\mathbf{N}}$ of $\mathcal{H}$ and a sequence of nonnegative real numbers $\{\lambda_n\}_{n\in\mathbf{N}}$ such that $\sum_{n\in\mathbf{N}} \lambda_n < \infty$ and

$$K e_n = \lambda_n e_n, \qquad n \in \mathbf{N}.$$

Take a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables $\{\xi_n\}_{n\in\mathbf{N}}$ defined on a probability space $(\Omega, \mathcal{F}, P)$. Define the random variable $X_n : \Omega \to \mathcal{H}$ by

$$X_n(\omega) = \sum_{k=1}^{n} \sqrt{\lambda_k}\, \xi_k(\omega) e_k, \qquad \omega \in \Omega.$$

Then, since we have

$$\mathbb{E}\left[\|X_n\|^2\right] = \mathbb{E}\left[\sum_{k=1}^{n} \lambda_k \xi_k^2\right] = \sum_{k=1}^{n} \lambda_k < \infty,$$

the sequence $\{X_n\}_{n\in\mathbf{N}}$ converges in mean square, so that it has a subsequence that converges almost surely to a random variable $X : \Omega \to \mathcal{H}$. Let μ be the distribution of X. Now, for any $h, k \in \mathcal{H}$, we have

$$\begin{aligned}
C_\mu(h,k) &= \mathbb{E}[\langle h, X\rangle \langle k, X\rangle] \\
&= \lim_{n\to\infty} \mathbb{E}\left[\langle h, X_n\rangle \langle k, X_n\rangle\right] \\
&= \lim_{n\to\infty} \sum_{i,j=1}^{n} \sqrt{\lambda_i \lambda_j}\,\mathbb{E}\left[\xi_i \xi_j\right] \langle h, e_i\rangle \langle k, e_j\rangle \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \lambda_i \langle h, e_i\rangle \langle k, e_i\rangle \\
&= \langle Kh, k\rangle.
\end{aligned}$$

Thus, we have $\hat{C}_\mu = K$, which concludes the proof. □

1.3 Kolmogorov's Continuity Theorem

In this subsection, we present Kolmogorov's continuity theorem. Let $\mathcal{B} = \mathcal{C}([0,1]^d, \mathbb{R})$ denote the space of continuous functions from $[0,1]^d$ to $\mathbb{R}$. Equipped with the supremum norm

$$\|f\|_\infty = \sup_{x\in[0,1]^d} |f(x)|,$$

it is a separable Banach space.

By the Riesz-Markov theorem, its dual space can be identified with

$$\mathcal{C}([0,1]^d, \mathbb{R})^* \cong \mathcal{M}([0,1]^d),$$

the space of finite signed Borel measures on $[0,1]^d$. For $f \in \mathcal{C}([0,1]^d, \mathbb{R})$ and $\nu \in \mathcal{M}([0,1]^d)$, the duality pairing is

$$\langle f, \nu\rangle = \int_{[0,1]^d} f(x)\,\nu(dx).$$

Let μ be a centered Gaussian measure on $\mathcal{C}([0,1]^d, \mathbb{R})$. Its covariance operator

$$C_\mu : \mathcal{M}([0,1]^d) \times \mathcal{M}([0,1]^d) \to \mathbb{R}$$

is defined by

$$C_\mu(\nu, \nu') = \int_{\mathcal{C}([0,1]^d,\mathbb{R})} \langle f, \nu\rangle \langle f, \nu'\rangle\, \mu(df).$$

In particular, for $x, y \in [0,1]^d$,

$$\begin{aligned} C_\mu(\delta_x, \delta_y) &= \int_{\mathcal{C}([0,1]^d,\mathbb{R})} \langle f, \delta_x \rangle \langle f, \delta_y \rangle \, \mu(df) \\ &= \int_{\mathcal{C}([0,1]^d,\mathbb{R})} f(x)\, f(y)\, \mu(df). \end{aligned}$$

For brevity, we write $C_\mu(x, y)$ instead of $C_\mu(\delta_x, \delta_y)$.

The Kolmogorov continuity theorem gives sufficient conditions for a function $C : [0,1]^d \times [0,1]^d \to \mathbb{R}$ to be the covariance function of a centered Gaussian measure on $\mathcal{C}([0,1]^d, \mathbb{R})$.

Theorem 1.13 (Kolmogorov's Continuity Theorem)**.** *For $d > 0$, let $C : [0,1]^d \times [0,1]^d \to \mathbf{R}$ be a symmetric function such that, for every finite collection $\{x_i\}_{i=1}^m$ of points in $[0,1]^d$, the matrix $C_{ij} = C(x_i, x_j)$ is positive definite. If furthermore there exists $\alpha > 0$ and a constant $K > 0$ such that $C(x,x)+C(y,y)-2C(x,y) \leqslant K|x-y|^{2\alpha}$ for any two points $x, y \in [0,1]^d$ then there exists a unique centred Gaussian measure μ on $\mathcal{C}\left([0,1]^d, \mathbf{R}\right)$ such that*

$$\int_{\mathcal{C}([0,1]^d,\mathbf{R})} f(x) f(y) \mu(df) = C(x,y)$$

for any two points $x, y \in [0,1]^d$. Furthermore, for every $\beta < \alpha$, one has $\mu\left(\mathcal{C}^\beta\left([0,1]^d, \mathbf{R}\right)\right) = 1$, where $\mathcal{C}^\beta\left([0,1]^d, \mathbf{R}\right)$ is the space of β-Hölder continuous functions.

Example 1.14. Let $\mathcal{B} = \mathcal{C}([0,1], \mathbf{R})$ and define

$$C : [0,1] \times [0,1] \to \mathbf{R}, \qquad C(x,y) := x \wedge y.$$

To apply Kolmogorov's continuity theorem, we need two facts about C. First, it must be positive semidefinite. Second, its increments must satisfy a suitable bound.

For the first point, fix $x_1, \ldots, x_m \in [0,1]$ and $a_1, \ldots, a_m \in \mathbf{R}$. Consider the indicator functions

$$g_i(t) := \mathbf{1}_{[0,x_i]}(t), \qquad G(t) := \sum_{i=1}^m a_i g_i(t).$$

A direct computation shows that

$$\langle g_i, g_j \rangle_{L^2([0,1])} = \int_0^1 \mathbf{1}_{[0,x_i]}(t)\mathbf{1}_{[0,x_j]}(t)\, dt = x_i \wedge x_j,$$

and therefore

$$\sum_{i,j=1}^{m} a_i a_j\, C(x_i, x_j) = \sum_{i,j=1}^{m} a_i a_j (x_i \wedge x_j) = \|G\|^2_{L^2([0,1])} \geqslant 0.$$

So C is positive semidefinite.

Next we check the increment bound. For any $x, y \in [0,1]$,

$$C(x,x) + C(y,y) - 2C(x,y) = x + y - 2(x \wedge y) = |x-y|.$$

In particular, for every $\alpha \in (0, \frac{1}{2}]$ we have $|x-y| \leqslant |x-y|^{2\alpha}$, so the required estimate holds.

With these two checks in hand, Kolmogorov's continuity theorem applies and yields a unique centred Gaussian measure μ on $\mathcal{C}([0,1], \mathbf{R})$ whose covariance is $C(x,y) = x \wedge y$. This Gaussian measure is called the **Wiener measure**.

Remark. In probability theory, **Wiener measure** is often introduced via its values on **cylinder sets**. Fix $0 < t_1 < \cdots < t_n \leqslant 1$ and Borel sets $A_1, \ldots, A_n \subset \mathbf{R}$. Define

$$\mathcal{C}(t_1, \ldots, t_n; A_1, \ldots, A_n) := \{f \in \mathcal{C}([0,1], \mathbf{R}) : f(t_i) \in A_i \text{ for } i = 1, \ldots, n\}.$$

The Wiener measure μ_W is characterized by

$$\mu_W(\mathcal{C}(t_1, \ldots, t_n; A_1, \ldots, A_n)) = \int_{A_1} p(t_1, x_1)\, dx_1 \prod_{i=2}^{n} \int_{A_i} p(t_i - t_{i-1}, x_i - x_{i-1})\, dx_i,$$

where

$$p(t,x) = \frac{1}{\sqrt{2\pi t}} \exp\!\left(-\frac{x^2}{2t}\right)$$

is the density of $\mathcal{N}(0,t)$.

Let $X_t(f) := f(t)$ be the canonical process. Under μ_W, $(X_t)_{t\in[0,1]}$ is a standard Brownian motion: $X_0 = 0$ a.s., and for $0 \leqslant s < t \leqslant 1$ the increment $X_t - X_s \sim \mathcal{N}(0, t-s)$ and is independent of $\sigma(X_r : r \leqslant s)$. In particular, for

$s \leqslant t$,

$$\mathbb{E}[X_s X_t] = \mathbb{E}\big[X_s\big(X_s + (X_t - X_s)\big)\big] = \mathbb{E}[X_s^2] + \mathbb{E}[X_s]\,\mathbb{E}[X_t - X_s] = s,$$

so $\mathbb{E}[X_s X_t] = s \wedge t$. This agrees with the covariance kernel $C(x, y) = x \wedge y$ of the Gaussian measure obtained from Kolmogorov's continuity theorem.

The next theorem is a more general version of Kolmogorov's continuity theorem. It provides sufficient conditions for a $\mathcal{B}$-valued stochastic process to admit a continuous modification.

Let $\{X(x) : \Omega \to \mathcal{B}\}_{x \in [0,1]^d}$ be a collection of $\mathcal{B}$-valued random variables. Under suitable moment assumptions, one can construct another collection $\{Y(x) : \Omega \to \mathcal{B}\}_{x \in [0,1]^d}$ such that

(i) for every $x \in [0, 1]^d$, $X(x)$ and $Y(x)$ are equal in law;

(ii) for $\mathbb{P}$-almost every $\omega \in \Omega$, the map $x \mapsto Y(x, \omega)$ is continuous.

Equivalently, Y can be viewed as a random variable with values in the path space $\mathcal{C}\big([0, 1]^d, \mathcal{B}\big)$, i.e.

$$Y : \Omega \to \mathcal{C}\big([0, 1]^d, \mathcal{B}\big),$$

where $Y(\omega)(x) := Y(x, \omega)$. The distribution of Y is then a Gaussian measure on $\mathcal{C}\big([0, 1]^d, \mathcal{B}\big)$.

Theorem 1.15. *Let $\mathcal{B}$ be a separable Banach space and let $\{X(x)\}_{x \in [0,1]^d}$ be a collection of $\mathcal{B}$-valued Gaussian random variables such that*

$$\mathbb{E}\|X(x) - X(y)\| \leqslant C|x - y|^{\alpha},$$

for some $C > 0$ and some $\alpha \in (0, 1]$. Then there exists a unique Gaussian measure μ on $\mathcal{C}\big([0, 1]^d, \mathcal{B}\big)$ such that, if Y is a random variable with law μ, then $Y(x)$ is equal in law to $X(x)$ for every $x \in [0, 1]^d$. Furthermore,

$$\mu\big(\mathcal{C}^{\beta}\big([0, 1]^d, \mathcal{B}\big)\big) = 1 \quad \textit{for every } \beta < \alpha.$$

1.4 Cameron-Martin Space

Given a centered Gaussian measure μ on a separable Banach space $\mathcal{B}$, the measure μ is generally not invariant under translations. However, there exists a Hilbert subspace $H \subset \mathcal{B}$, called the **Cameron-Martin space**, such that for every $h \in H$ the translated measure $\mu(\cdot - h)$ is equivalent to μ, i.e. they have the same null sets.

Definition 1.16. The **Cameron-Martin space** $\mathcal{H}_\mu$ associated to μ is the completion of the linear subspace $\mathring{\mathcal{H}}_\mu \subset \mathcal{B}$ defined by

$$\mathring{\mathcal{H}}_\mu = \{h \in \mathcal{B} : \ \exists\, h^* \in \mathcal{B}^* \text{ such that } C_\mu(h^*, \ell) = \ell(h) \ \ \forall\, \ell \in \mathcal{B}^*\}\,,$$

under the norm $\|\cdot\|_\mu$ defined as follows. For $h, k \in \mathring{\mathcal{H}}_\mu$ with corresponding representatives $h^*, k^* \in \mathcal{B}^*$, we set

$$\langle h, k\rangle_\mu := C_\mu(h^*, k^*).$$

This is well-defined and yields a pre-Hilbert structure on $\mathring{\mathcal{H}}_\mu$; its completion is the Hilbert space $\mathcal{H}_\mu$.

Remark. The space $\mathring{\mathcal{H}}_\mu$ coincides with the range of the covariance operator $\hat{C}_\mu : \mathcal{B}^* \to \mathcal{B}$ defined by

$$\left(\hat{C}_\mu(\ell_1)\right)(\ell) = C_\mu(\ell_1, \ell) \qquad \forall\, \ell, \ell_1 \in \mathcal{B}^*.$$

Indeed, if $h \in \mathring{\mathcal{H}}_\mu$ and $h^* \in \mathcal{B}^*$ satisfies $C_\mu(h^*, \ell) = \ell(h)$ for every $\ell \in \mathcal{B}^*$, then $\ell(h) = C_\mu(h^*, \ell) = \ell\Big(\hat{C}_\mu(h^*)\Big)$ for every $\ell \in \mathcal{B}^*$, hence $h = \hat{C}_\mu(h^*)$. Conversely, if $h = \hat{C}_\mu(\ell)$ for some $\ell \in \mathcal{B}^*$, then

$$C_\mu(\ell, \ell') = \ell'(h) \qquad \forall\, \ell' \in \mathcal{B}^*,$$

so $h \in \mathring{\mathcal{H}}_\mu$.

Remark. The norm on Cameron-Martin space is well-defined. Indeed, if $h \in \mathring{\mathcal{H}}_\mu$ has two representatives $h_1^*, h_2^* \in \mathcal{B}^*$, consider $k = h_1^* + h_2^*$. Then, we have

$$C_\mu(h_1^*, h_1^*) - C_\mu(h_2^*, h_2^*) = C_\mu(h_1^*, k) - C_\mu(h_2^*, k) = k(h) - k(h) = 0.$$

Thus, the norm is independent of the choice of representative.

The next theorem shows that the Cameron-Martin space uniquely determines a centered Gaussian measure.

Theorem 1.17. *Let μ and ν be centered Gaussian measures on a separable Banach space $\mathcal{B}$. If $\mathcal{H}_\mu = \mathcal{H}_\nu$ and $\|\cdot\|_\mu = \|\cdot\|_\nu$ on this space, then $\mu = \nu$.*

Proof. By Theorem 1.7, it suffices to show that $\hat{\mu}(\ell) = \hat{\nu}(\ell)$ for every $\ell \in \mathcal{B}^*$. Fix $\ell \in \mathcal{B}^*$ and write $\mathcal{H} := \mathcal{H}_\mu = \mathcal{H}_\nu$. Since the norms agree, the inner products on $\mathcal{H}$ also agree.

For any $h \in \mathcal{H}$, the covariance operator satisfies

$$\langle \hat{C}_\mu(\ell), h\rangle_\mu = \ell(h) \qquad \text{and} \qquad \langle \hat{C}_\nu(\ell), h\rangle_\nu = \ell(h).$$

Hence

$$\langle \hat{C}_\mu(\ell) - \hat{C}_\nu(\ell), h\rangle_\mu = 0 \qquad \forall\, h \in \mathcal{H},$$

so $\hat{C}_\mu(\ell) = \hat{C}_\nu(\ell)$ in $\mathcal{H}$.

By (1.1), we have

$$\hat{\mu}(\ell) = \exp\Big(-\tfrac{1}{2}\,\|\hat{C}_\mu(\ell)\|_\mu^2\Big), \qquad \hat{\nu}(\ell) = \exp\Big(-\tfrac{1}{2}\,\|\hat{C}_\nu(\ell)\|_\nu^2\Big).$$

Since $\hat{C}_\mu(\ell) = \hat{C}_\nu(\ell)$ and the norms coincide on $\mathcal{H}$, we get $\hat{\mu}(\ell) = \hat{\nu}(\ell)$. Therefore $\mu = \nu$. □

The next theorem proves that the Cameron-Martin space is indeed a subset of the original Banach space.

Theorem 1.18. *The Cameron-Martin space $\mathcal{H}_\mu$ is a subset of $\mathcal{B}$. Furthermore, there exists a constant $C > 0$ such that for every $h \in \mathcal{H}_\mu$, one has*

$$\|h\|^2 \leqslant C\|h\|_\mu^2. \tag{1.4}$$

This implies that the inclusion map $\iota : \mathcal{H}_\mu \to \mathcal{B}$ is continuous.

Proof. For $h \in \overset{\circ}{\mathcal{H}}_\mu$, we have

$$\begin{aligned}\|h\|^2 &= \sup_{\ell \in \mathcal{B}^*\setminus\{0\}} \frac{\ell(h)^2}{\|\ell\|^2} = \sup_{\ell \in \mathcal{B}^*\setminus\{0\}} \frac{C_\mu\left(h^*, \ell\right)^2}{\|\ell\|^2} \\ &\leqslant \sup_{\ell \in \mathcal{B}^*\setminus\{0\}} \frac{C_\mu\left(h^*, h^*\right) C_\mu(\ell, \ell)}{\|\ell\|^2} \\ &\leqslant \|C_\mu\|\,\langle h, h\rangle_\mu,\end{aligned}$$

this proves (1.4) for $h \in \overset{\circ}{\mathcal{H}}_\mu$. Now, by (1.4), every Cauchy sequence in $\overset{\circ}{\mathcal{H}}_\mu$ with respect to the norm $\|\cdot\|_\mu$ is also a Cauchy sequence in $\mathcal{B}$ with respect to the norm $\|\cdot\|$. Since $\mathcal{B}$ is complete, the limit of this Cauchy sequence in $\mathcal{B}$ exists. This shows that $\mathcal{H}_\mu \subset \mathcal{B}$ and that the inclusion map $\iota : \mathcal{H}_\mu \to \mathcal{B}$ is continuous. □

In general, the representative $h^* \in \mathcal{B}^*$ appearing in the definition of $\mathcal{H}_\mu$ need not be unique. For example, if $\mu = \delta_0$, then $C_\mu \equiv 0$ and $\overset{\circ}{\mathcal{H}}_\mu = \{0\}$; for $h = 0$ the representative can be any $h^* \in \mathcal{B}^*$.

This non-uniqueness disappears after identifying $\mathcal{B}^*$ with a subspace of $L^2(\mathcal{B}, \mu)$ via the embedding

$$\mathcal{B}^* \hookrightarrow L^2(\mathcal{B}, \mu), \qquad \ell \mapsto \ell(\cdot).$$

Let $\mathcal{R}_\mu$ denote the closure of $\mathcal{B}^*$ in $L^2(\mathcal{B}, \mu)$. It sometimes called the **reproducing kernel Hilbert space** associated to μ.

Theorem 1.19. *There exists a canonical isometric isomorphism $\iota : \mathcal{H}_\mu \to \mathcal{R}_\mu$, written $h \mapsto h^*$. In particular, $\mathcal{H}_\mu$ is separable.*

Proof. We first prove that ι is injective. Suppose $h \in \mathcal{H}_\mu$ satisfies $\iota(h) = h^* = 0$ in $L^2(\mathcal{B}, \mu)$. Then for every $\ell \in \mathcal{B}^*$ we have

$$\ell(h) = C_\mu(h^*, \ell) = 0.$$

Since $\mathcal{B}^*$ separates points of $\mathcal{B}$, it follows that $h = 0$. Hence ι is injective.

Next, let $f \in \mathcal{B}^* \subset L^2(\mathcal{B}, \mu)$. Define

$$h := \hat{C}_\mu(f) = \int_{\mathcal{B}} x\, f(x)\, \mu(dx) \in \mathcal{B}.$$

Then for every $\ell \in \mathcal{B}^*$,

$$\ell(h) = \ell\Big(\hat{C}_\mu(f)\Big) = C_\mu(f, \ell),$$

so $h \in \overset{\circ}{\mathcal{H}}_\mu$ with representative $h^* = f$. In particular, $\iota(h) = f$, which shows that ι maps onto $\mathcal{B}^*$.

Finally, ι is an isometry on $\overset{\circ}{\mathcal{H}}_\mu$:

$$\|h\|_\mu^2 = C_\mu(h^*, h^*) = \|h^*\|_{L^2(\mu)}^2,$$

and therefore extends uniquely by continuity to an isometric map $\iota : \mathcal{H}_\mu \to \mathcal{R}_\mu$. Since $\mathcal{B}^*$ is dense in $\mathcal{R}_\mu$ and $\iota(\mathcal{H}_\mu)$ is closed (being complete), we obtain $\iota(\mathcal{H}_\mu) = \mathcal{R}_\mu$, i.e. ι is surjective.

In particular, $\mathcal{H}_\mu$ is separable because $\mathcal{R}_\mu \subset L^2(\mathcal{B}, \mu)$ is separable. $\square$

We need the following lemma for next result.

Lemma 1.20. *Let $\mathcal{H}$ be a Hilbert space and let $D \subset \mathcal{H}$ be dense. Then for every $h \in \mathcal{H}$,*

$$\|h\| = \sup\{\langle h, k\rangle \mid k \in D,\ \|k\| \leqslant 1\}.$$

Proof. Fix $h \in \mathcal{H}$ and $k \in D$ with $\|k\| \leqslant 1$. By Cauchy-Schwarz,

$$\langle h, k\rangle \leqslant \|h\|\,\|k\| \leqslant \|h\|.$$

Taking the supremum over such k gives

$$\sup\{\langle h, k\rangle \mid k \in D,\ \|k\| \leqslant 1\} \leqslant \|h\|.$$

For the reverse inequality, since D is dense, there exists $k_n \in D$ such that

$$\left\|\frac{h}{\|h\|} - k_n\right\| \to 0 \qquad (n \to \infty)$$

(when $h = 0$ the statement is trivial). Then

$$\langle h, k_n\rangle \longrightarrow \langle h, h/\|h\|\rangle = \|h\|,$$

and therefore

$$\|h\| \leqslant \sup\{\langle h, k\rangle \mid k \in D,\ \|k\| \leqslant 1\}.$$

□

The next result provides an alternative characterisation of the Cameron-Martin space.

Theorem 1.21. *Let μ be a centred Gaussian measure on $\mathcal{B}$. For $h \in \mathcal{B}$ set*

$$\|h\|'_\mu := \sup\{\ell(h) \mid \ell \in \mathcal{B}^*,\ C_\mu(\ell, \ell) \leqslant 1\}.$$

Then for every $h \in \mathcal{H}_\mu$ one has $\|h\|_\mu = \|h\|'_\mu$. In particular,

$$\mathcal{H}_\mu = \{\, h \in \mathcal{B} \mid \|h\|'_\mu < \infty \,\}.$$

Proof. Apply Lemma 1.20 to the Hilbert space $\mathcal{H}_\mu$ with dense subspace $D = \mathring{\mathcal{H}}_\mu$. For $h \in \mathcal{H}_\mu$ we obtain

$$\|h\|_\mu = \sup\{\langle h, k\rangle_\mu \mid k \in \mathring{\mathcal{H}}_\mu,\ \|k\|_\mu \leqslant 1\}.$$

Every $k \in \mathring{\mathcal{H}}_\mu$ can be written as $k = \hat{C}_\mu(\ell)$ for some $\ell \in \mathcal{B}^*$, and then

$$\langle h, k\rangle_\mu = \ell(h), \qquad \|k\|_\mu^2 = C_\mu(\ell, \ell).$$

Hence

$$\|h\|_\mu = \sup\{\ell(h) \mid \ell \in \mathcal{B}^*,\ C_\mu(\ell,\ell) \leqslant 1\} = \|h\|'_\mu.$$

This proves

$$\mathcal{H}_\mu \subset \{h \in \mathcal{B} \mid \|h\|'_\mu < \infty\}.$$

Conversely, suppose $h \in \mathcal{B}$ satisfies $\|h\|'_\mu < \infty$. Define a linear map on $\mathring{\mathcal{H}}_\mu$ by

$$T : \mathring{\mathcal{H}}_\mu \to \mathbb{R}, \qquad T(\hat{C}_\mu(\ell)) := \ell(h).$$

This is well defined, and

$$|T(\hat{C}_\mu(\ell))| = |\ell(h)| \leqslant \|h\|'_\mu \sqrt{C_\mu(\ell,\ell)} = \|h\|'_\mu \, \|\hat{C}_\mu(\ell)\|_\mu,$$

so T is bounded. Hence T extends uniquely to a bounded linear functional on $\mathcal{H}_\mu$. By the Riesz representation theorem, there exists a unique $k \in \mathcal{H}_\mu$ such that

$$\ell(h) = \langle k, \hat{C}_\mu(\ell)\rangle_\mu = C_\mu(k^*, \ell) \qquad \forall\, \ell \in \mathcal{B}^*.$$

Thus $\ell(h-k) = 0$ for all $\ell \in \mathcal{B}^*$, and since $\mathcal{B}^*$ separates points, we conclude $h = k \in \mathcal{H}_\mu$. The proof is complete. $\square$

Since $\mathcal{R}_\mu$ is obtained as the $L^2(\mu)$-closure of $\mathcal{B}^*$, it is not surprising that its elements behave like linear functionals on a set of full μ-measure.

Theorem 1.22. *For every $\ell \in \mathcal{R}_\mu$, there exist a measurable linear subspace $V_\ell \subset \mathcal{B}$ and a linear map $\tilde{\ell} : V_\ell \to \mathbf{R}$ such that $\mu(V_\ell) = 1$ and $\ell = \tilde{\ell}$ μ-almost surely.*

Proof. By definition of $\mathcal{R}_\mu$, there exists a sequence $(\ell_n)_{n\geqslant 1} \subset \mathcal{B}^*$ such that

$$\|\ell_n - \ell\|^2_{L^2(\mu)} \leqslant n^{-4}.$$

Set $X_n := \ell_n - \ell$. By Chebyshev's inequality, for every $\varepsilon > 0$,

$$\mu\big(\,|X_n| > \varepsilon\,\big) \leqslant \frac{\|X_n\|^2_{L^2(\mu)}}{\varepsilon^2}.$$

Choosing $\varepsilon = 1/n$ yields

$$\mu\big(\,|X_n| > 1/n\,\big) \leqslant n^2 \|X_n\|^2_{L^2(\mu)} \leqslant n^{-2},$$

and therefore $\sum_{n\geqslant 1} \mu(|X_n| > 1/n) < \infty$. By the Borel-Cantelli lemma, $|X_n(x)| \leqslant 1/n$ for all but finitely many n, for μ-almost every $x \in \mathcal{B}$; in particular, $\ell_n(x) \to \ell(x)$ for μ-almost every x.

Define

$$V_\ell := \Big\{ x \in \mathcal{B} : \lim_{n\to\infty} \ell_n(x) \text{ exists} \Big\}, \qquad \tilde{\ell}(x) := \lim_{n\to\infty} \ell_n(x) \;\text{ for } x \in V_\ell.$$

Then $\mu(V_\ell) = 1$ and $\ell = \tilde{\ell}$ μ-almost surely. Moreover, V_ℓ is a linear subspace and $\tilde{\ell}$ is linear on V_ℓ since it is the pointwise limit of linear maps. □

The next theorem shows that every element of $\mathcal{R}_\mu$ is a centred Gaussian random variable.

Theorem 1.23. *Let $h^* = \iota(h) \in \mathcal{R}_\mu$. Then h^* is a centred Gaussian random variable with variance $\|h\|^2_\mu$. Moreover, for $h^* = \iota(h)$ and $k^* = \iota(k)$ in $\mathcal{R}_\mu$,*

$$\mathbb{E}[h^* k^*] = \langle h, k \rangle_\mu.$$

Proof. By definition of a centred Gaussian measure, $\ell(x)$ is a centred Gaussian random variable for every $\ell \in \mathcal{B}^*$. Since $\mathcal{R}_\mu$ is the $L^2(\mu)$-closure of $\mathcal{B}^*$, for any $h^* \in \mathcal{R}_\mu$ there exists a sequence $(\ell_n)_{n\geqslant 1} \subset \mathcal{B}^*$ such that $\|\ell_n - h^*\|_{L^2(\mu)} \to 0$. Replacing ℓ_n by a suitable rescaling if necessary, we may assume $\|\ell_n\|_{L^2(\mu)} = \|h^*\|_{L^2(\mu)}$ for all n. Hence each ℓ_n has law $\mathcal{N}(0, \|h^*\|^2_{L^2(\mu)})$. Since L^2-convergence implies convergence in distribution, it follows that h^* also has law $\mathcal{N}(0, \|h^*\|^2_{L^2(\mu)})$. Using $\|h^*\|_{L^2(\mu)} = \|h\|_\mu$ gives the variance claim.

For the covariance, note that $h^* + k^* = \iota(h + k)$, so by the first part,

$$\mathbb{E}[(h^* + k^*)^2] = \|h + k\|^2_\mu, \qquad \mathbb{E}[(h^*)^2] = \|h\|^2_\mu, \qquad \mathbb{E}[(k^*)^2] = \|k\|^2_\mu.$$

Therefore, by polarisation,

$$\begin{aligned} \mathbb{E}[h^* k^*] &= \frac{1}{2}\Big(\mathbb{E}[(h^* + k^*)^2] - \mathbb{E}[(h^*)^2] - \mathbb{E}[(k^*)^2] \Big) \\ &= \frac{1}{2}\Big(\|h + k\|^2_\mu - \|h\|^2_\mu - \|k\|^2_\mu \Big) \\ &= \langle h, k \rangle_\mu. \end{aligned}$$

$\square$

Given two measures μ and ν on a measurable space $(\mathcal{B}, \mathcal{F})$, we say that μ and ν are **equivalent** if they have the same null sets (or preserve the full measure set equivalently), i.e. if $\mu(A) = 0$ if and only if $\nu(A) = 0$ for every $A \in \mathcal{F}$. On the other hand, we say that μ and ν are **mutually singular** if there exist $A, B \in \mathcal{F}$ such that $\mu(A) = 1$, $\nu(B) = 1$, and $A \cap B = \varnothing$.

The next theorem illustrates a striking difference between Gaussian measures on infinite-dimensional spaces and those on $\mathbb{R}^n$.

Theorem 1.24. *Let μ be a centred Gaussian measure on a separable Banach space $\mathcal{B}$ with $\dim(\mathcal{H}_\mu) = \infty$. Let $D_c : \mathcal{B} \to \mathcal{B}$ be the dilation map $D_c(x) = cx$. Then, for every $c \neq \pm 1$, the measures μ and $D_c^\sharp \mu$ are mutually singular.*

Proof. Since $\mathcal{R}_\mu$ is a separable infinite-dimensional Hilbert space, it admits an orthonormal basis $\{e_n\}_{n \geqslant 1}$. Define

$$X_N(x) := \frac{1}{N} \sum_{n=1}^{N} |e_n(x)|^2, \qquad x \in \mathcal{B}.$$

Under μ, the random variables e_n are independent centred Gaussians with variance 1, hence $\mathbb{E}_\mu[|e_1|^2] = 1$. By the strong law of large numbers,

$$X_N(x) \longrightarrow 1 \qquad \text{for } \mu\text{-almost every } x \in \mathcal{B}.$$

On the other hand, under $D_c^\sharp \mu$ the random variables e_n remain independent centred Gaussians, but their variance is multiplied by c^2 since $e_n(cx) = c\, e_n(x)$. Thus $\mathbb{E}_{D_c^\sharp \mu}[|e_1|^2] = c^2$, and again by the strong law of large numbers,

$$X_N(x) \longrightarrow c^2 \qquad \text{for } D_c^\sharp \mu\text{-almost every } x \in \mathcal{B}.$$

Let

$$A := \left\{ x \in \mathcal{B} : \lim_{N \to \infty} X_N(x) = 1 \right\}, \qquad B := \left\{ x \in \mathcal{B} : \lim_{N \to \infty} X_N(x) = c^2 \right\}.$$

Then $\mu(A) = 1$, $D_c^\sharp \mu(B) = 1$, and $A \cap B = \varnothing$ whenever $c^2 \neq 1$, i.e. $c \neq \pm 1$. Hence μ and $D_c^\sharp \mu$ are mutually singular. $\square$

The next result is the celebrated Cameron-Martin theorem, which characterises exactly those translations that preserve the null sets of a Gaussian measure.

Theorem 1.25 (Cameron-Martin)**.** *For $h \in \mathcal{B}$, let $T_h : \mathcal{B} \to \mathcal{B}$ be the translation $T_h(x) = x + h$. Then $T_h^\sharp \mu \ll \mu$ if and only if $h \in \mathcal{H}_\mu$. Moreover, if $h \in \mathcal{H}_\mu$, then*

$$\frac{d\,T_h^\sharp \mu}{d\mu}(x) = \exp\left(h^*(x) - \frac{1}{2}\|h\|_\mu^2\right),$$

where $h^* = \iota(h) \in \mathcal{R}_\mu$.

Proof. Fix $h \in \mathcal{H}_\mu$ and write $h^* = \iota(h) \in \mathcal{R}_\mu$. By Theorem 1.23, h^* is a centred Gaussian random variable with variance $\|h\|_\mu^2$. Define

$$D_h(x) := \exp\left(h^*(x) - \frac{1}{2}\|h\|_\mu^2\right), \qquad x \in \mathcal{B}.$$

Then $D_h > 0$ and

$$\mathbb{E}_\mu[D_h] = \exp\left(-\frac{1}{2}\|h\|_\mu^2\right)\mathbb{E}_\mu\left[e^{h^*}\right] = \exp\left(-\frac{1}{2}\|h\|_\mu^2\right)\exp\left(\frac{1}{2}\|h\|_\mu^2\right) = 1,$$

so $D_h \in L^1(\mu)$ defines a probability measure μ_h by

$$\mu_h(A) := \int_A D_h \, d\mu, \qquad A \in \mathcal{F}.$$

In particular, $\mu_h \ll \mu$ with Radon-Nikodym derivative $d\mu_h/d\mu = D_h$.

To identify μ_h, we compare Fourier transforms. For $\ell \in \mathcal{B}^*$,

$$\widehat{\mu_h}(\ell) = \int_{\mathcal{B}} e^{i\ell(x)}\,\mu_h(dx) = \int_{\mathcal{B}} \exp\left(i\ell(x) + h^*(x) - \frac{1}{2}\|h\|_\mu^2\right)\mu(dx).$$

Since $i\ell + h^*$ is a (complex-valued) centred Gaussian random variable under μ, its exponential moment can be computed explicitly, yielding

$$\widehat{\mu_h}(\ell) = \exp\left(-\frac{1}{2}C_\mu(\ell,\ell) + i\,\ell(h)\right).$$

On the other hand,

$$\begin{aligned}\widehat{T_h^\sharp \mu}(\ell) &= \int_{\mathcal{B}} e^{i\ell(x)}\,T_h^\sharp \mu(dx) = \int_{\mathcal{B}} e^{i\ell(x+h)}\,\mu(dx) = e^{i\ell(h)}\int_{\mathcal{B}} e^{i\ell(x)}\,\mu(dx) \\ &= \exp\left(-\frac{1}{2}C_\mu(\ell,\ell) + i\,\ell(h)\right).\end{aligned}$$

Thus $\widehat{\mu_h} = \widehat{T_h^\sharp \mu}$ on $\mathcal{B}^*$, hence $\mu_h = T_h^\sharp \mu$. This proves $T_h^\sharp \mu \ll \mu$ and the claimed Radon-Nikodym derivative.

We use the fact that for every $a \in \mathbb{R}$,

$$\|\mathcal{N}(0,1) - \mathcal{N}(a,1)\|_{\mathrm{TV}} \geqslant 2 - 2\exp\left(-\frac{a^2}{8}\right).$$

Assume now that $h \notin \mathcal{H}_\mu$. By Theorem 1.21, for each $n \geqslant 1$ there exists $\ell_n \in \mathcal{B}^*$ such that $C_\mu(\ell_n, \ell_n) = 1$ and $\ell_n(h) \geqslant n$. Then

$$\ell_n^\sharp \mu = \mathcal{N}(0,1), \qquad \ell_n^\sharp (T_h^\sharp \mu) = \mathcal{N}(\ell_n(h), 1).$$

Since total variation distance decreases under pushforward,

$$\begin{aligned} \|\mu - T_h^\sharp \mu\|_{\mathrm{TV}} &\geqslant \|\ell_n^\sharp \mu - \ell_n^\sharp (T_h^\sharp \mu)\|_{\mathrm{TV}} = \|\mathcal{N}(0,1) - \mathcal{N}(\ell_n(h),1)\|_{\mathrm{TV}} \\ &\geqslant 2 - 2\exp\left(-\frac{\ell_n(h)^2}{8}\right) \geqslant 2 - 2\exp\left(-\frac{n^2}{8}\right). \end{aligned}$$

Letting $n \to \infty$ yields $\|\mu - T_h^\sharp \mu\|_{\mathrm{TV}} = 2$, which implies that μ and $T_h^\sharp \mu$ are mutually singular. $\square$

Using the Cameron-Martin theorem, one can give another criterion for the Cameron-Martin. It can be used to check the dimension of the Cameron-Martin space is infinite.

Theorem 1.26. *The space $\mathcal{H}_\mu \subset \mathcal{B}$ is the intersection of all (measurable) linear subspaces of full measure. However, if $\mathcal{H}_\mu$ is infinite-dimensional, then one has $\mu(\mathcal{H}_\mu) = 0$.*

Proof. Let $V \subset \mathcal{B}$ be a full measure linear subspace and $h \in \mathcal{H}_\mu$. By the Cameron-Martin theorem, we should have $T_h^\sharp \mu(V) = 1$, which implies $\mu(V - h) = 1$. If $h \notin V$, then $V \cap (V - h) = \varnothing$, this will give a contracdiction since $\mu(V \cup (V - h)) = 2$. Thus, we have $h \in V$, which shows that $\mathcal{H}_\mu$ is contained in the intersection of all full measure linear subspaces.

Conversely, let $h \notin \mathcal{H}_\mu$, we will construct a full measure linear subspace V such that $h \notin V$. By Theorem 1.21, there exists $\{\ell_n\} \in \mathcal{B}^*$ such that $C_\mu(\ell_n, \ell_n) = 1$ and $\ell_n(h) \geqslant n$. Define a norm on $\mathcal{B}$ by

$$\|y\|_*^2 := \sum_{n=1}^\infty \frac{\ell_n(y)^2}{n^2}.$$

Then, we have

$$\int_{\mathcal{B}} \|y\|_*^2 \mu(dy) = \sum_{n=1}^{\infty} \frac{C_\mu(\ell_n, \ell_n)}{n^2} = \frac{\pi^2}{6} < \infty,$$

this implies that the linear subspace

$$V := \{y \in \mathcal{B} : \ \|y\|_* < \infty\}$$

is of full measure. However, by construction, we have $\|h\|_* = \infty$, so $h \notin V$.

Now, assume that $\mathcal{H}_\mu$ is infinite-dimensional. We then pick an orthonormal sequence $\{e^*(n)\}_{n\geqslant 1} \subset \mathcal{R}_\mu$ so that $e^*(n) \sim \mathcal{N}(0,1)$ are independent centred Gaussian random variables. Now, let us denote

$$p = \mathbb{P}(|\mathcal{N}(0,1)| \leqslant 1) > 0,$$

and set

$$A_n := \{x \in \mathcal{B} : \ |e^*(n)(x)| \geqslant 1\},$$

then we have $\mu(A_n) = p$ and

$$\sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} p = \infty.$$

By the Borel-Cantelli lemma, we have μ-almost surely that $|e^*(n)(x)| \geqslant 1$ for infinitely many n. This implies that for μ-almost every $x \in \mathcal{B}$, we have

$$\|x\|_\mu^2 = \sum_{n=1}^{\infty} |e^*(n)(x)|^2 = \infty,$$

so $x \notin \mathcal{H}_\mu$. Thus, we have $\mu\left(\mathcal{H}_\mu\right) = 0$. □

1.5 Images of Gaussian measures under linear maps

Let $A : \mathcal{B} \to \mathcal{B}_2$ be a bounded linear operator between two separable Banach spaces. Recall that the **adjoint** operator $A^* : \mathcal{B}_2^* \to \mathcal{B}^*$ is defined by

$$A^*(\ell)(x) := \ell(Ax), \qquad \forall\, \ell \in \mathcal{B}_2^*, \ x \in \mathcal{B}.$$

Let μ be a centered Gaussian measure on $\mathcal{B}$ and denote by $A^\sharp\mu$ its image under A on $\mathcal{B}_2$. Then, for any $\ell_1, \ell_2 \in \mathcal{B}_2^*$, the covariance form of $A^\sharp\mu$ satisfies

$$\begin{aligned} C_{A^\sharp\mu}(\ell_1, \ell_2) &= \int_{\mathcal{B}_2} \ell_1(x)\, \ell_2(x)\, A^\sharp\mu(dx) \\ &= \int_{\mathcal{B}} \ell_1(Ax)\, \ell_2(Ax)\, \mu(dx) \\ &= C_\mu(A^*\ell_1, A^*\ell_2). \end{aligned}$$

Now recall that the Cameron-Martin space $\mathcal{H}_\mu$ can be characterised as the intersection of all linear subspaces of $\mathcal{B}$ having full μ-measure. In particular, if $A, B : \mathcal{B} \to \mathcal{B}_2$ are linear maps that agree on $\mathcal{H}_\mu$, then they agree on a linear subspace of full μ-measure; consequently, their pushforward measures coincide, i.e. $A^\sharp\mu = B^\sharp\mu$. In this sense, the image measure of a linear map is determined by its restriction to $\mathcal{H}_\mu$.

Theorem 1.27. *Let μ be a centred Gaussian probability measure on a separable Banach space $\mathcal{B}$. Let $\mathcal{H}$ be a separable Hilbert space and let $A : \mathcal{H}_\mu \to \mathcal{H}$ be a Hilbert-Schmidt operator (equivalently, $AA^* : \mathcal{H} \to \mathcal{H}$ is trace class). Then there exists a measurable map $\hat{A} : \mathcal{B} \to \mathcal{H}$ such that $\nu = \hat{A}^\sharp\mu$ is Gaussian on $\mathcal{H}$ with covariance*

$$C_\nu(h, k) = \langle A^*h, A^*k\rangle_\mu, \qquad h, k \in \mathcal{H}.$$

Moreover, there exists a measurable linear subspace $V \subset \mathcal{B}$ with $\mu(V) = 1$ such that $\hat{A}|_V$ is linear and $\hat{A}|_{\mathcal{H}_\mu} = A$.

Proof. Let $\{e_n\}_{n\geqslant 1}$ be an orthonormal basis of $\mathcal{H}_\mu$ and let $\{e_n^*\}_{n\geqslant 1} \subset \mathcal{R}_\mu$ denote the corresponding coordinate maps. For $N \geqslant 1$, define $S_N : \mathcal{B} \to \mathcal{H}$ by

$$S_N(x) := \sum_{n=1}^{N} e_n^*(x)\, Ae_n.$$

By Theorem 1.22, for each $n \geqslant 1$ there exists a measurable linear subspace $V_n \subset \mathcal{B}$ with $\mu(V_n) = 1$ such that e_n^* is linear on V_n. Set

$$V_0 := \bigcap_{n=1}^{\infty} V_n.$$

Then $\mu(V_0) = 1$, and for every $N \geqslant 1$ the map S_N is linear on V_0.

Since the random variables $\{e_n^*\}_{n\geqslant 1}$ are i.i.d. $\mathcal{N}(0, 1)$ under μ, the sequence

$\{S_N\}_{N\geqslant 1}$ is an $\mathcal{H}$-valued martingale. Moreover,

$$\mathbb{E}_\mu \|S_N\|_{\mathcal{H}}^2 = \sum_{n=1}^{N} \|Ae_n\|_{\mathcal{H}}^2 \leqslant \sum_{n=1}^{\infty} \|Ae_n\|_{\mathcal{H}}^2 = \|A\|_{HS}^2 < \infty.$$

Hence $\{S_N\}$ is bounded in $L^2(\mu;\mathcal{H})$, and by the martingale convergence theorem $S_N(x)$ converges in $\mathcal{H}$ for μ-almost every x.

Define

$$V := \Big\{x \in V_0 : \lim_{N\to\infty} S_N(x) \text{ exists in } \mathcal{H}\Big\},$$

so that $\mu(V) = 1$, and set

$$\hat{A}(x) := \begin{cases} \lim_{N\to\infty} S_N(x), & x \in V, \\ 0, & x \notin V. \end{cases}$$

Then $\hat{A}$ is measurable, $\hat{A}|_V$ is linear, and one checks that $\hat{A}|_{\mathcal{H}_\mu} = A$.

Finally, fix $h \in \mathcal{H}$. By construction,

$$\langle \hat{A}(x), h\rangle_{\mathcal{H}} = \sum_{n=1}^{\infty} e_n^*(x)\, \langle Ae_n, h\rangle_{\mathcal{H}} \quad \text{in } L^2(\mu),$$

and therefore

$$C_\nu(h,h) = \mathbb{E}_\mu\big[\langle \hat{A}(x), h\rangle_{\mathcal{H}}^2\big] = \sum_{n=1}^{\infty} \langle Ae_n, h\rangle_{\mathcal{H}}^2 = \|A^*h\|_{\mathcal{H}_\mu}^2.$$

By polarisation this yields $C_\nu(h,k) = \langle A^*h, A^*k\rangle_{\mathcal{H}_\mu}$. Moreover, since each finite partial sum is a linear combination of independent Gaussian variables, $\langle \hat{A}(\cdot), h\rangle$ is Gaussian for every $h \in \mathcal{H}$, hence ν is Gaussian. $\square$

In fact, the result above extends to the case where the image space is a general separable Banach space. We omit the proof.

Theorem 1.28. *Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be separable Banach spaces, and let μ be a centred Gaussian probability measure on $\mathcal{B}_1$. Let $A : \mathcal{H}_\mu \to \mathcal{B}_2$ be a bounded linear operator. Assume that there exists a centred Gaussian measure ν on $\mathcal{B}_2$ whose covariance satisfies*

$$C_\nu(\ell_1,\ell_2) = \langle A^*\ell_1, A^*\ell_2\rangle_\mu, \qquad \ell_1,\ell_2 \in \mathcal{B}_2^*,$$

where $A^* : \mathcal{B}_2^* \to \mathcal{H}_\mu$ denotes the adjoint map. Then there exists a measurable map $\hat{A} : \mathcal{B}_1 \to \mathcal{B}_2$ such that $\nu = \hat{A}^\sharp \mu$. Moreover, there exists a measurable linear subspace $V \subset \mathcal{B}_1$ with $\mu(V) = 1$ such that $\hat{A}|_V$ is linear and $\hat{A}|_{\mathcal{H}_\mu} = A$.

Lastly, we show that the extension $\hat{A}$ is unique up to a μ-null set. The main tool is the Borell-Sudakov-Cirel'son inequality. For $\varepsilon > 0$, denote by B_ε the open ball of radius ε centred at 0 in the Cameron-Martin space $\mathcal{H}_\mu$, namely

$$B_\varepsilon := \{h \in \mathcal{H}_\mu : \ \|h\|_\mu < \varepsilon\}.$$

We also write Φ for the cumulative distribution function of the standard normal law,

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

Theorem 1.29 (Borell-Sudakov-Cirel'son inequality)**.** *Let μ be a centred Gaussian measure on a separable Banach space $\mathcal{B}$ with Cameron-Martin space $\mathcal{H}_\mu$. Let $A \subset \mathcal{B}$ be measurable with $\mu(A) = \Phi(\alpha)$ for some $\alpha \in \mathbb{R}$. Then, for every $\varepsilon > 0$,*

$$\mu(A + B_\varepsilon) \geqslant \Phi(\alpha + \varepsilon).$$

A striking consequence is that if A has positive μ-measure, then $A + \mathcal{H}_\mu$ has full μ-measure. Indeed, since $B_\varepsilon \subset \mathcal{H}_\mu$ for every $\varepsilon > 0$, we have

$$\mu(A + \mathcal{H}_\mu) \geqslant \sup_{\varepsilon > 0} \mu(A + B_\varepsilon) \geqslant \sup_{\varepsilon > 0} \Phi(\alpha + \varepsilon) = 1.$$

Another consequence is the following 0-1 law for measurable linear subspaces.

Theorem 1.30. *Let $V \subset \mathcal{B}$ be a measurable linear subspace. Then $\mu(V) = 0$ or $\mu(V) = 1$.*

Proof. First assume that $\mathcal{H}_\mu \subset V$. Then $B_\varepsilon \subset V$ for every $\varepsilon > 0$, hence $V + B_\varepsilon = V$. If $\mu(V) > 0$, write $\mu(V) = \Phi(\alpha)$ for some $\alpha \in \mathbb{R}$. By the Borell-Sudakov-Cirel'son inequality,

$$\mu(V) = \mu(V + B_\varepsilon) \geqslant \Phi(\alpha + \varepsilon) \qquad \forall\, \varepsilon > 0.$$

Letting $\varepsilon \to \infty$ yields $\mu(V) = 1$.

Next assume that $\mathcal{H}_\mu \not\subset V$. Choose $h \in \mathcal{H}_\mu \backslash V$. Since V is a linear subspace, the sets $\{V + th\}_{t \in \mathbb{R}}$ are pairwise disjoint: if $t_1 \neq t_2$ and $(V + t_1 h) \cap (V + t_2 h) \neq \varnothing$, then for some $v_1, v_2 \in V$ we have $v_1 + t_1 h = v_2 + t_2 h$, which implies $(t_1 - t_2)h = v_2 - v_1 \in V$, hence $h \in V$, a contradiction.

If $\mu(V) > 0$, then by the Cameron-Martin theorem each translate $V + th$ has positive μ-measure, and we would obtain uncountably many pairwise disjoint measurable sets of positive measure, which is impossible for a probability measure. Therefore $\mu(V) = 0$. $\square$

The last theorem in this subsection establishes uniqueness of the measurable extension of a linear map defined on the Cameron-Martin space.

Theorem 1.31. *Let μ be a Gaussian measure on a separable Banach space $\mathcal{B}_1$ with Cameron-Martin space $\mathcal{H}_\mu$, and let $A : \mathcal{H}_\mu \to \mathcal{B}_2$ be a linear map satisfying the assumptions of Theorem 1.28. Then the extension $\hat{A}$ of A is unique (up to μ-null sets) within the class of measurable maps for which there exists a measurable linear subspace $V \subset \mathcal{B}_1$ with $\mu(V) = 1$ such that $\hat{A}$ is linear on V and $\hat{A}x = Ax$ for all $x \in \mathcal{H}_\mu \subset V$.*

Proof. We argue by contradiction. Suppose there exist two measurable extensions $\hat{A}_1$ and $\hat{A}_2$ of A satisfying the assumptions of the theorem. Then there exist measurable linear subspaces $V_1, V_2 \subset \mathcal{B}_1$ with $\mu(V_1) = \mu(V_2) = 1$ such that $\hat{A}_1$ is linear on V_1, $\hat{A}_2$ is linear on V_2, and

$$\hat{A}_1 x = \hat{A}_2 x = Ax \qquad \text{for all } x \in \mathcal{H}_\mu.$$

Set $V := V_1 \cap V_2$. Then $\mu(V) = 1$, and both $\hat{A}_1$ and $\hat{A}_2$ are linear on V. Define $\Delta : V \to \mathcal{B}_2$ by

$$\Delta x := \hat{A}_1 x - \hat{A}_2 x.$$

Then Δ is linear on V and $\Delta x = 0$ for every $x \in \mathcal{H}_\mu \subset V$. We claim that $\Delta x = 0$ μ-almost surely.

Fix $\ell \in \mathcal{B}_2^*$ and $c \in \mathbb{R}$, and consider the measurable sets

$$V_\ell^c := \{x \in V : \ \ell(\Delta x) \leqslant c\}.$$

These sets are invariant under translations by elements of $\mathcal{H}_\mu$. Indeed, for $h \in \mathcal{H}_\mu$ and $x \in V$,

$$\ell(\Delta(x + h)) = \ell(\Delta x + \Delta h) = \ell(\Delta x),$$

since $\Delta h = 0$ on $\mathcal{H}_\mu$. Hence $x \in V_\ell^c$ if and only if $x + h \in V_\ell^c$. By Theorem 1.29

we therefore have $\mu(V_\ell^c) \in \{0,1\}$ for every $c \in \mathbb{R}$.

For fixed ℓ, the map $c \mapsto \mu(V_\ell^c)$ is increasing, and by σ-additivity,

$$\lim_{c\to-\infty} \mu(V_\ell^c) = 0, \qquad \lim_{c\to+\infty} \mu(V_\ell^c) = 1.$$

It follows that there exists a unique $c_\ell \in \mathbb{R}$ such that $\mu(V_\ell^c)$ jumps from 0 to 1 at $c = c_\ell$. Equivalently,

$$\ell(\Delta x) = c_\ell \qquad \mu\text{-almost surely.}$$

Since μ is centred Gaussian, it is invariant under the symmetry $x \mapsto -x$. By linearity of Δ we have $\ell(\Delta(-x)) = -\ell(\Delta x)$, and therefore $c_\ell = -c_\ell$, which implies $c_\ell = 0$. Hence for every $\ell \in \mathcal{B}_2^*$,

$$\ell(\Delta x) = 0 \qquad \mu\text{-almost surely.}$$

By Theorem 1.7, this implies that the law of Δx is the Dirac mass at 0, and in particular $\Delta x = 0$ μ-almost surely. Consequently, $\hat{A}_1 = \hat{A}_2$ μ-almost surely, which proves uniqueness. □

1.6 Cylindrical Wiener processes and stochastic integration

In this subsection we introduce cylindrical Wiener processes on a separable Banach space and the associated stochastic integrals.

Definition 1.32. Let $(\Omega, \mathcal{F})$ be a measurable space and let $\mathcal{B}$ be a separable Banach space. A **stochastic process** with values in $\mathcal{B}$ is a family of measurable maps $\{X_t : \Omega \to \mathcal{B}\}_{t\geqslant 0}$.

Consider the path space $\Omega = \mathcal{C}([0,T],\mathcal{B})$ equipped with its Borel σ-algebra $\mathcal{F} = \mathcal{B}(\mathcal{C}([0,T],\mathcal{B}))$. The **canonical process** $\{X_t\}_{t\in[0,T]}$ on Ω is defined by

$$X_t(\omega) = \omega(t), \qquad \omega \in \mathcal{C}([0,T],\mathcal{B}).$$

Recall that a **one-dimensional Wiener process** (Brownian motion) $\{W_t\}_{t\geqslant 0}$ on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ is an $\mathbb{R}$-valued stochastic process such that $W_0 = 0$ $\mathbb{P}$-almost surely, W has independent increments, and for every $0 \leqslant s < t$, $W_t - W_s \sim \mathcal{N}(0, t-s)$.

On a finite time interval $[0,T]$, one can realise W as the canonical process on $\mathcal{C}([0,T],\mathbb{R})$ endowed with the Wiener measure μ_W (cf. Example 1.14). This measure is Gaussian with covariance function

$$C_{\mu_W}(s,t) = s \wedge t = \min\{s,t\}, \qquad s,t \in [0,T].$$

However, this construction does not extend directly to a Wiener process on $\mathbb{R}_+$, since the path space $\mathcal{C}(\mathbb{R}_+,\mathbb{R})$ is not a separable Banach space in the topology that controls behaviour at infinity. To circumvent this issue, we work instead with a weighted path space.

Definition 1.33. Let $\rho(t) = 1 + t^2$ for $t \geqslant 0$. Define

$$\mathcal{C}_\rho(\mathbb{R}_+,\mathbb{R}) := \left\{ x \in \mathcal{C}(\mathbb{R}_+,\mathbb{R}) : \lim_{t\to\infty} \frac{|x(t)|}{\rho(t)} = 0 \right\},$$

equipped with the norm

$$\|x\|_\rho := \sup_{t\geqslant 0} \frac{|x(t)|}{\rho(t)}.$$

The next result shows that one can realise Wiener measure on this Banach space. The proof is omitted.

Theorem 1.34. *There exists a centred Gaussian measure μ_W on $\mathcal{C}_\rho(\mathbb{R}_+,\mathbb{R})$ with covariance function*

$$C_{\mu_W}(s,t) = s \wedge t, \qquad s,t \geqslant 0.$$

Moreover, the canonical process $\{W_t\}_{t\geqslant 0}$ on $\big(\mathcal{C}_\rho(\mathbb{R}_+,\mathbb{R}), \mathcal{B}(\mathcal{C}_\rho(\mathbb{R}_+,\mathbb{R})), \mu_W\big)$ is a one-dimensional Wiener process (Brownian motion) on $\mathbb{R}_+$.

An **$\mathbb{R}^n$-valued Wiener process** is simply given by n independent copies of a one-dimensional Wiener process. More precisely, we write

$$W(t) = (W_1(t), \ldots, W_n(t)),$$

where $\{W_i(t)\}_{1\leqslant i\leqslant n}$ are independent standard one-dimensional Wiener processes. In particular, for all $s,t \geqslant 0$,

$$\mathbb{E}\big[W_i(s)W_j(t)\big] = \delta_{ij}\,(s\wedge t).$$

Consequently, for any $u, v \in \mathbb{R}^n$,

$$\begin{aligned}
\mathbb{E}\big[\langle W(s), u\rangle \, \langle W(t), v\rangle\big] &= \mathbb{E}\left[\left(\sum_{i=1}^{n} u_i W_i(s)\right)\left(\sum_{j=1}^{n} v_j W_j(t)\right)\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} u_i v_j \, \mathbb{E}\big[W_i(s) W_j(t)\big] \\
&= (s \wedge t) \sum_{i=1}^{n} u_i v_i \\
&= (s \wedge t) \, \langle u, v\rangle.
\end{aligned}$$

This characterisation admits a natural extension to arbitrary Hilbert spaces. Let $\mathcal{H}$ be a Hilbert space with orthonormal basis $\{e_n\}_{n\geqslant 1}$. It is tempting to define an $\mathcal{H}$-valued Wiener process by

$$W(t) := \sum_{n=1}^{\infty} W_n(t) \, e_n,$$

where $\{W_n\}_{n\geqslant 1}$ are independent standard one-dimensional Wiener processes. However, this series does not converge in $\mathcal{H}$ for any $t > 0$, since

$$\mathbb{E}\|W(t)\|_{\mathcal{H}}^2 = \sum_{n=1}^{\infty} \mathbb{E}|W_n(t)|^2 = \sum_{n=1}^{\infty} t = +\infty.$$

Therefore, to make sense of such a construction, we must enlarge the state space.

Theorem 1.35. *Let $\mathcal{H}$ be a separable Hilbert space. Then there exists a Hilbert space $\mathcal{H}'$ such that $\mathcal{H}$ is densely embedded in $\mathcal{H}'$ and the inclusion map $\iota : \mathcal{H} \to \mathcal{H}'$ is Hilbert-Schmidt.*

Proof. Let $\{e_n\}_{n\geqslant 1}$ be an orthonormal basis of $\mathcal{H}$ and define a new norm on $\mathcal{H}$ by

$$\|x\|_{\mathcal{H}'}^2 := \sum_{n=1}^{\infty} \frac{1}{n^2} \, \langle x, e_n\rangle_{\mathcal{H}}^2.$$

Let $\mathcal{H}'$ be the completion of $\mathcal{H}$ with respect to this norm. Then $\{ne_n\}_{n\geqslant 1}$ is an orthonormal basis of $\mathcal{H}'$. Moreover,

$$\|\iota\|_{HS}^2 = \sum_{n=1}^{\infty} \|\iota e_n\|_{\mathcal{H}'}^2 = \sum_{n=1}^{\infty} \|e_n\|_{\mathcal{H}'}^2 = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty,$$

so the inclusion $\iota : \mathcal{H} \to \mathcal{H}'$ is Hilbert-Schmidt. $\square$

We can now define cylindrical Wiener processes.

Definition 1.36. Let $\mathcal{H}$ be a separable Hilbert space and let $\mathcal{H}'$ be as in Theorem 1.35, with inclusion map $\iota : \mathcal{H} \to \mathcal{H}'$. A **cylindrical Wiener process on** $\mathcal{H}$ is an $\mathcal{H}'$-valued Gaussian process $\{W(t)\}_{t\geqslant 0}$ such that for all $h, k \in \mathcal{H}'$ and $s, t \geqslant 0$,

$$\mathbb{E}\big[\langle W(s), h\rangle_{\mathcal{H}'} \, \langle W(t), k\rangle_{\mathcal{H}'}\big] = (s \wedge t)\, \langle \iota^* h, \iota^* k\rangle_{\mathcal{H}} = (s \wedge t)\, \langle \iota\iota^* h, k\rangle_{\mathcal{H}'}.$$

By Kolmogorov's continuity theorem, such a process can be realised as the canonical process on a suitable path space $\mathcal{C}_\rho(\mathbb{R}_+, \mathcal{H}')$ endowed with an appropriate Gaussian measure.

Moreover, in the same setting, the smaller Hilbert space $\mathcal{H}$ can be identified with the Cameron-Martin space of the centred Gaussian measure on $\mathcal{H}'$ whose covariance operator is $\iota\iota^*$.

Theorem 1.37. *Let μ be the centred Gaussian measure on $\mathcal{H}'$ with covariance operator $Q = \iota\iota^*$. Then the Cameron-Martin space of μ is $\iota(\mathcal{H})$. Moreover, for every $\hat{h} \in \mathcal{H}$,*

$$\|\iota\hat{h}\|_\mu^2 = \|\hat{h}\|_{\mathcal{H}}^2.$$

Proof. Recall that the pre-Cameron-Martin space $\mathring{\mathcal{H}}_\mu$ is the range of the covariance operator. Hence

$$\mathring{\mathcal{H}}_\mu = \mathrm{Ran}(Q) = \mathrm{Ran}(\iota\iota^*) \subset \mathrm{Ran}(\iota) = \iota(\mathcal{H}).$$

Conversely, let $\hat{h} \in \mathcal{H}$ and set $h := \iota\hat{h} \in \mathcal{H}'$. Since $h \in \mathrm{Ran}(Q)$, we may define

$$h^* := Q^{-1}h = (\iota\iota^*)^{-1}\iota\hat{h} \in \mathcal{H}'.$$

Then for every $k \in \mathcal{H}'$,

$$C_\mu(h^*, k) = \langle Qh^*, k\rangle_{\mathcal{H}'} = \langle h, k\rangle_{\mathcal{H}'},$$

which shows that $h \in \mathring{\mathcal{H}}_\mu$. Therefore

$$\iota(\mathcal{H}) \subset \mathring{\mathcal{H}}_\mu \subset \iota(\mathcal{H}),$$

and hence $\mathring{\mathcal{H}}_\mu = \iota(\mathcal{H})$. Taking the completion with respect to the Cameron-Martin norm yields $\mathcal{H}_\mu = \iota(\mathcal{H})$.

Finally, take $h = \iota\hat{h}$ and $k = \iota\hat{k}$ in $\mathring{\mathcal{H}}_\mu$. Using $h^* = Q^{-1}h$ and $k^* = Q^{-1}k$, we obtain

$$\begin{aligned}
\langle h, k\rangle_\mu &:= C_\mu(h^*, k^*) \\
&= \langle Qh^*, k^*\rangle_{\mathcal{H}'} \\
&= \langle h, Q^{-1}k\rangle_{\mathcal{H}'} \\
&= \langle \iota\hat{h}, (\iota\iota^*)^{-1}\iota\hat{k}\rangle_{\mathcal{H}'} \\
&= \langle \hat{h}, \iota^*(\iota\iota^*)^{-1}\iota\hat{k}\rangle_{\mathcal{H}} = \langle \hat{h}, \hat{k}\rangle_{\mathcal{H}},
\end{aligned}$$

where in the last step we used $\iota^*(\iota\iota^*)^{-1}\iota = I_{\mathcal{H}}$ on $\mathcal{H}$. In particular,

$$\|\iota\hat{h}\|_\mu^2 = \langle \iota\hat{h}, \iota\hat{h}\rangle_\mu = \|\hat{h}\|_{\mathcal{H}}^2,$$

as claimed. □

Remark. The terminology **cylindrical Wiener process on** $\mathcal{H}$ can be motivated as follows. Although a cylindrical Wiener process W is in general only $\mathcal{H}'$-valued, let us pretend that $W(t) \in \mathcal{H}$ for every $t \geqslant 0$ for a while. Then, for any $h, k \in \mathcal{H}$, using the identity $\iota^*(\iota\iota^*)^{-1}\iota = I_{\mathcal{H}}$ (on the appropriate subspace) and the adjoint relation between ι and ι^*, we obtain

$$\begin{aligned}
\mathbb{E}\big[\langle W(s), h\rangle_{\mathcal{H}}\, \langle W(t), k\rangle_{\mathcal{H}}\big] &= \mathbb{E}\Big[\big\langle \iota W(s), (\iota\iota^*)^{-1}\iota h\big\rangle_{\mathcal{H}'}\, \big\langle \iota W(t), (\iota\iota^*)^{-1}\iota k\big\rangle_{\mathcal{H}'}\Big] \\
&= (s \wedge t)\, \Big\langle (\iota\iota^*)(\iota\iota^*)^{-1}\iota h,\ (\iota\iota^*)^{-1}\iota k\Big\rangle_{\mathcal{H}'} \\
&= (s \wedge t)\, \big\langle \iota h, (\iota\iota^*)^{-1}\iota k\big\rangle_{\mathcal{H}'} \\
&= (s \wedge t)\, \langle h, k\rangle_{\mathcal{H}}.
\end{aligned}$$

This is precisely the covariance identity of a standard Wiener process on $\mathcal{H}$, which explains the terminology.

Remark. Let $\mathcal{H}$ and $\mathcal{H}'$ be as in Theorem 1.35, and let $\mathcal{K}$ be another separable Hilbert space. If $A : \mathcal{H} \to \mathcal{K}$ is a bounded linear operator, then by Theorem 1.27 there exists a measurable extension $\widehat{A} : \mathcal{H}' \to \mathcal{K}$ such that, whenever W is a cylindrical Wiener process on $\mathcal{H}$, the composition $\widehat{A}\,W$ is well-defined as a $\mathcal{K}$-valued process. By a slight abuse of notation, we will simply write AW in place of $\widehat{A}\,W$.

Now we give a precise definition of **white noise**.

Definition 1.38. Let $\mathcal{H}$ be a Hilbert space of distributions on $\mathbb{R}$ such that the embedding

$$L^2(\mathbb{R}) \hookrightarrow \mathcal{H}$$

is Hilbert-Schmidt. A **white noise** is an $\mathcal{H}$-valued centred Gaussian random variable ξ such that, for all $g, h \in L^2(\mathbb{R})$,

$$\mathbb{E}\big[\langle g, \xi\rangle \, \langle h, \xi\rangle\big] = \langle g, h\rangle_{L^2(\mathbb{R})},$$

where $\langle g, \xi\rangle$ denotes the dual pairing between $L^2(\mathbb{R})$ and $\mathcal{H}$ (which coincides with the L^2 inner product whenever $\xi \in L^2(\mathbb{R})$).

Remark. Taking $g = \mathbf{1}_{[0,t]}$ and $h = \mathbf{1}_{[0,s]}$ yields

$$\mathbb{E}\big[\langle \mathbf{1}_{[0,t]}, \xi\rangle \, \langle \mathbf{1}_{[0,s]}, \xi\rangle\big] = \langle \mathbf{1}_{[0,t]}, \mathbf{1}_{[0,s]}\rangle_{L^2(\mathbb{R})} = t \wedge s.$$

Consequently, the process

$$W(t) := \langle \mathbf{1}_{[0,t]}, \xi\rangle = \int_0^t \xi(r)\, dr$$

is a standard Brownian motion. In this sense, white noise can be viewed as the (distributional) time-derivative of Brownian motion.

Remark. The covariance identity can be written formally as

$$\mathbb{E}\bigg[\bigg(\int_{\mathbb{R}} g(s)\xi(s)\, ds\bigg)\bigg(\int_{\mathbb{R}} h(t)\xi(t)\, dt\bigg)\bigg] = \int_{\mathbb{R}} g(s)h(s)\, ds.$$

Equivalently, in distributional notation,

$$\mathbb{E}\big[\xi(s)\xi(t)\big] = \delta(t-s),$$

meaning that for all $g, h \in L^2(\mathbb{R})$,

$$\iint_{\mathbb{R}^2} g(s)h(t)\, \mathbb{E}\big[\xi(s)\xi(t)\big]\, ds\, dt = \int_{\mathbb{R}} g(s)h(s)\, ds.$$

Thus, one may think of the law of ξ as a Gaussian measure on a space of distributions whose covariance kernel is the Dirac delta.

Now, let $W(t)$ be a cylindrical Wiener process on a separable Hilbert space

$\mathcal{H}$, it can be view as a cannonical process on $\mathcal{C}_\rho(\mathbb{R}_+, \mathcal{H}')$ endowed with a suitable Gaussian measure, where $\mathcal{H}'$ is as in Theorem 1.35. We also denote by $\mathcal{F}_t$ the filtration generated by $\{W(s) : s \leqslant t\}$. We are now in position to define the stochastic integral with respect to W.

Let us first define the stochastic integral for elementary predictable processes.

Definition 1.39. Fix separable Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$. Let $\{(s_j, t_j]\}_{j=1}^N$ be a finite family of pairwise disjoint intervals in $\mathbb{R}_+$, and let

$$\Phi_j : \Omega \to \mathcal{L}_2(\mathcal{H}, \mathcal{K}), \qquad j = 1, \dots, N,$$

be $\mathcal{F}_{s_j}$-measurable random variables. An **elementary predictable process** is a process

$$\Phi \in L^2\big(\mathbb{R}_+ \times \Omega; \mathcal{L}_2(\mathcal{H}, \mathcal{K})\big)$$

of the form

$$\Phi(t, \omega) = \sum_{j=1}^N \Phi_j(\omega)\, \mathbf{1}_{(s_j, t_j]}(t).$$

Definition 1.40. Let Φ be an elementary predictable process as above and let W be a cylindrical Wiener process on $\mathcal{H}$. We define the **stochastic integral** of Φ with respect to W by

$$\int_0^\infty \Phi(t)\, dW(t) := \sum_{j=1}^N \Phi_j\big(W(t_j) - W(s_j)\big).$$

Equivalently, for $\omega \in \Omega$,

$$\left(\int_0^\infty \Phi(t)\, dW(t)\right)(\omega) = \sum_{j=1}^N \Phi_j(\omega)\big(W(t_j, \omega) - W(s_j, \omega)\big).$$

Remark. Since the cylindrical Wiener process W is $\mathcal{H}'$-valued, the increment $W(t_j) - W(s_j)$ takes values in $\mathcal{H}'$ and a priori cannot be acted on by an operator defined only on $\mathcal{H}$. This is why we require

$$\Phi_j(\omega) \in \mathcal{L}_2(\mathcal{H}, \mathcal{K}),$$

so that, by Theorem 1.27, each $\Phi_j(\omega)$ admits a measurable extension to a map $\widehat{\Phi_j(\omega)} : \mathcal{H}' \to \mathcal{K}$ which is linear on a subspace of full measure. With this convention, the term $\Phi_j(\omega)\big(W(t_j) - W(s_j)\big)$ is understood as $\widehat{\Phi_j(\omega)}\big(W(t_j) -$

$W(s_j)\big)$, and the stochastic integral is therefore well-defined (up to $\mathbb{P}$-null sets).

One of the fundamental properties of the stochastic integral is the Itô isometry. It asserts that the stochastic integral defines an isometry from the space of elementary predictable processes into $L^2(\Omega;\mathcal{K})$.

Theorem 1.41 (Itô isometry for cylindrical Wiener processes)**.** *Let Φ be an elementary predictable process of the form*

$$\Phi(t,\omega) = \sum_{j=1}^{N} \Phi_j(\omega)\,\mathbf{1}_{(s_j,t_j]}(t), \qquad \Phi_j : \Omega \to \mathcal{L}_2(\mathcal{H},\mathcal{K}) \text{ is } \mathcal{F}_{s_j}\text{-measurable},$$

where the intervals $(s_j,t_j]$ are pairwise disjoint. Then

$$\mathbb{E}\left\|\int_0^\infty \Phi(t)\,dW(t)\right\|_{\mathcal{K}}^2 = \mathbb{E}\int_0^\infty \|\Phi(t)\|^2_{\mathcal{L}_2(\mathcal{H},\mathcal{K})}\,dt = \mathbb{E}\int_0^\infty \operatorname{tr}\big(\Phi(t)\Phi(t)^*\big)\,dt.$$

Proof. By definition of the stochastic integral,

$$\int_0^\infty \Phi(t)\,dW(t) = \sum_{j=1}^{N} \Phi_j\big(W(t_j) - W(s_j)\big).$$

Since the increments of W over disjoint intervals are independent and centred, the cross terms vanish, and therefore

$$\begin{aligned}\mathbb{E}\left\|\int_0^\infty \Phi(t)\,dW(t)\right\|_{\mathcal{K}}^2 &= \mathbb{E}\left\|\sum_{j=1}^{N} \Phi_j\big(W(t_j) - W(s_j)\big)\right\|_{\mathcal{K}}^2 \\ &= \sum_{j=1}^{N} \mathbb{E}\left\|\Phi_j\big(W(t_j) - W(s_j)\big)\right\|_{\mathcal{K}}^2.\end{aligned}$$

By Corollary 1.12, for each $j = 1,\dots,N$,

$$\mathbb{E}\left\|\Phi_j\big(W(t_j) - W(s_j)\big)\right\|_{\mathcal{K}}^2 = (t_j - s_j)\,\mathbb{E}\,\operatorname{tr}\big(\Phi_j\Phi_j^*\big).$$

Summing over j yields

$$\mathbb{E}\left\|\int_0^\infty \Phi(t)\,dW(t)\right\|_{\mathcal{K}}^2 = \sum_{j=1}^{N}(t_j - s_j)\,\mathbb{E}\,\operatorname{tr}\big(\Phi_j\Phi_j^*\big).$$

On the other hand, since $\Phi(t) = \Phi_j$ for $t \in (s_j, t_j]$, we have

$$\int_0^\infty \operatorname{tr}\bigl(\Phi(t)\Phi(t)^*\bigr)\,dt = \sum_{j=1}^N \int_{s_j}^{t_j} \operatorname{tr}\bigl(\Phi_j\Phi_j^*\bigr)\,dt = \sum_{j=1}^N (t_j - s_j)\ \operatorname{tr}\bigl(\Phi_j\Phi_j^*\bigr).$$

Taking expectations and combining the preceding identities gives

$$\mathbb{E}\left\|\int_0^\infty \Phi(t)\,dW(t)\right\|_{\mathcal{K}}^2 = \mathbb{E}\int_0^\infty \operatorname{tr}\bigl(\Phi(t)\Phi(t)^*\bigr)\,dt.$$

Finally, $\|\Phi(t)\|^2_{\mathcal{L}_2(\mathcal{H},\mathcal{K})} = \operatorname{tr}\bigl(\Phi(t)\Phi(t)^*\bigr)$, which concludes the proof. □

To extend the stochastic integral to more general integrands, we need the following density result.

Theorem 1.42. *The set of elementary processes is dense in the space* $L^2_{\mathrm{pr}}\left(\mathbf{R}_+ \times \Omega, \mathcal{L}_2(\mathcal{H},\mathcal{K})\right)$ *of all predictable* $\mathcal{L}_2(\mathcal{H},\mathcal{K})$*-valued processes.*

As a consequence, the stochastic integral can be extended by continuity to all predictable processes in $L^2\bigl(\mathbb{R}_+ \times \Omega; \mathcal{L}_2(\mathcal{H},\mathcal{K})\bigr)$.

Corollary 1.43. *The stochastic integral* $\int_0^\infty \Phi(t)dW(t)$ *can be uniquely defined for every process* $\Phi \in L^2_{\mathrm{pr}}\left(\mathbf{R}_+ \times \Omega, \mathcal{L}_2(\mathcal{H},\mathcal{K})\right)$.

2 Semigroup Theory

In this section, we aim to formulate a meaningful and rigorous solution theory for the abstract evolution equation

$$\frac{d}{dt}x(t) = Lx(t), \quad t > 0, \tag{2.1}$$

where $x(t)$ takes values in a Banach space $\mathcal{B}$ and $L : \mathcal{D}(L) \subset \mathcal{B} \to \mathcal{B}$ is a (typically unbounded) linear operator. The central idea is that solutions of (2.1) can often be described via a semigroup of operators generated by L.

Semigroups provide a convenient framework to describe time evolution: they encode how a state changes over time and constitute a natural language for well-posedness of PDEs.

Definition 2.1. Let $\mathcal{B}$ be a Banach space. A **semigroup** on $\mathcal{B}$ is a family of bounded linear operators $\{S(t)\}_{t\geqslant 0} \subset \mathcal{L}(\mathcal{B})$ such that

(i) $S(0) = \mathrm{Id}$;

(ii) $S(t)S(s) = S(t+s)$ for all $s, t \geqslant 0$.

2.1 Strongly continuous semigroups

This subsection is devoted to strongly continuous semigroups, also called C_0-semigroups.

Definition 2.2. A semigroup $\{S(t)\}_{t\geqslant 0}$ on a Banach space $\mathcal{B}$ is called a **strongly continuous semigroup** (or C_0**-semigroup**) if for every $x \in \mathcal{B}$ the map

$$t \mapsto S(t)x$$

is continuous from $[0, \infty)$ into $\mathcal{B}$.

The next theorem gives a useful characterisation of strong continuity.

Theorem 2.3. *A semigroup $\{S(t)\}_{t\geqslant 0}$ on a Banach space $\mathcal{B}$ is strongly continuous if and only if:*

(i) there exists a dense subset $D \subset \mathcal{B}$ such that the map $t \mapsto S(t)x$ is continuous at $t = 0$ for every $x \in D$;

(ii) there exist constants $M \geqslant 1$ and $a \geqslant 0$ such that

$$\|S(t)\| \leqslant Me^{at} \qquad \textit{for all } t \geqslant 0.$$

To prove Theorem 2.3, we use the following standard lemma (we omit its proof).

Lemma 2.4. *If S is a C_0-semigroup, then*

$$K = \sup_{0 \leqslant t \leqslant 1} \|S(t)\| < \infty.$$

Proof of Theorem 2.3. Assume first that S is a C_0-semigroup. Then (i) holds with $D = \mathcal{B}$. For (ii), let $K = \sup_{0 \leqslant t \leqslant 1} \|S(t)\|$ as in Lemma 2.4. Fix $t \geqslant 0$ and set $n = \lfloor t \rfloor + 1$. Then $\frac{t}{n} \in (0, 1]$, hence

$$\|S(t)\| = \left\| S\left(\frac{t}{n}\right)^n \right\| \leqslant \left\| S\left(\frac{t}{n}\right) \right\|^n \leqslant K^n \leqslant Ke^{(\log K)t}.$$

Therefore (ii) holds with $M = K$ and $a = \log K$.

Conversely, assume (i) and (ii). We first show that $t \mapsto S(t)x$ is continuous at $t = 0$ for every $x \in \mathcal{B}$. Fix $\varepsilon > 0$ and choose $y \in D$ such that

$$\|x - y\| \leqslant \min\left(\frac{\varepsilon}{3}, \frac{\varepsilon}{3Me^a}\right).$$

For $0 \leqslant t \leqslant 1$, (ii) yields $\|S(t)\| \leqslant Me^{at} \leqslant Me^a$, and thus

$$\|S(t)x - S(t)y\| \leqslant \|S(t)\| \, \|x - y\| \leqslant Me^a \|x - y\| \leqslant \frac{\varepsilon}{3}.$$

By (i), there exists $\delta > 0$ such that for $0 \leqslant t < \delta$,

$$\|S(t)y - y\| \leqslant \frac{\varepsilon}{3}.$$

Hence for $0 \leqslant t < \min\{1, \delta\}$,

$$\|S(t)x - x\| \leqslant \|S(t)x - S(t)y\| + \|S(t)y - y\| + \|y - x\| < \varepsilon.$$

This proves continuity at $t = 0$ for all $x \in \mathcal{B}$. Let $t_0 \geqslant 0$ and let $h \to 0$ with $t_0 + h \geqslant 0$. By the semigroup property,

$$\|S(t_0 + h)x - S(t_0)x\| = \|S(t_0)\big(S(h)x - x\big)\| \leqslant \|S(t_0)\| \, \|S(h)x - x\|.$$

By (ii), $\|S(t_0)\| \leqslant Me^{at_0} < \infty$, and by the previous step we have $\|S(h)x - x\| \to 0$ as $h \to 0$. Hence $\|S(t_0 + h)x - S(t_0)x\| \to 0$ as $h \to 0$, which shows that $t \mapsto S(t)x$ is continuous at t_0. Since t_0 was arbitrary, S is strongly continuous. $\square$

For a C_0-semigroup $\{S(t)\}_{t\geqslant 0}$, the map $(t, x) \mapsto S(t)x$ is jointly continuous in t and x. Concretely, for every $\varepsilon > 0$, $t_0 \geqslant 0$, and $x_0 \in \mathcal{B}$, there exists $\delta > 0$ such that

$$\|S(t)x - S(t_0)x_0\| < \varepsilon$$

whenever $|t - t_0| < \delta$ and $\|x - x_0\| < \delta$.

Theorem 2.5. *If S is a C_0-semigroup on $\mathcal{B}$, then the map $(t, x) \longmapsto S(t)x \in \mathcal{B}$ is jointly continuous from $\mathbb{R}_+ \times \mathcal{B}$ into $\mathcal{B}$.*

Proof. Fix $(t_0, x_0) \in \mathbb{R}_+ \times \mathcal{B}$ and $\varepsilon > 0$. As in the proof of Theorem 2.3, the local boundedness of C_0-semigroups implies that

$$M := \sup_{|t-t_0|<1} \|S(t)\| < \infty.$$

Choose $\delta_x := \frac{\varepsilon}{2M}$. Then, for every $x \in \mathcal{B}$ with $\|x - x_0\| < \delta_x$ and every t with $|t - t_0| < 1$, we have

$$\|S(t)(x - x_0)\| \leqslant \|S(t)\| \, \|x - x_0\| \leqslant M\|x - x_0\| < \frac{\varepsilon}{2}.$$

By strong continuity at t_0, there exists $\delta_t > 0$ such that whenever $|t - t_0| < \delta_t$,

$$\|S(t)x_0 - S(t_0)x_0\| < \frac{\varepsilon}{2}.$$

Therefore, for $|t - t_0| < \min\{1, \delta_t\}$ and $\|x - x_0\| < \delta_x$, we obtain

$$\|S(t)x - S(t_0)x_0\| \leqslant \|S(t)(x - x_0)\| + \|S(t)x_0 - S(t_0)x_0\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This proves joint continuity. $\square$

The generator of a C_0-semigroup captures its infinitesimal time evolution: it describes the instantaneous rate of change of $S(t)x$ at $t = 0$.

Definition 2.6. Let $\{S(t)\}_{t\geqslant 0}$ be a C_0-semigroup on a Banach space $\mathcal{B}$. The **generator** L is defined by

$$Lx := \lim_{h\downarrow 0} \frac{S(h)x - x}{h},$$

for all x in the domain

$$\mathcal{D}(L) := \left\{ x \in \mathcal{B} : \lim_{h \downarrow 0} \frac{S(h)x - x}{h} \text{ exists in } \mathcal{B} \right\}.$$

A priori, the domain $\mathcal{D}(L)$ could be small. The next theorem shows that this never happens for generators of C_0-semigroups.

Theorem 2.7. *Let L be the generator of a C_0-semigroup $\{S(t)\}_{t \geqslant 0}$ on a Banach space $\mathcal{B}$. Then $\mathcal{D}(L)$ is dense in $\mathcal{B}$.*

Proof. Fix $x \in \mathcal{B}$ and for $t > 0$ define

$$x_t := \int_0^t S(s)x \, ds,$$

where the integral is understood in the Bochner sense. We first note that

$$\frac{1}{t} x_t - x = \frac{1}{t} \int_0^t (S(s)x - x) \, ds.$$

Since $s \mapsto S(s)x$ is continuous and $S(0)x = x$, we have $S(s)x \to x$ as $s \downarrow 0$; hence the right-hand side converges to 0 as $t \downarrow 0$. Therefore,

$$\lim_{t \downarrow 0} \frac{1}{t} x_t = x \quad \text{in } \mathcal{B}.$$

Consequently, it suffices to show that $x_t \in \mathcal{D}(L)$ for every $t > 0$, because then elements of $\mathcal{D}(L)$ approximate an arbitrary $x \in \mathcal{B}$.

To verify $x_t \in \mathcal{D}(L)$, use the semigroup property:

$$S(h)x_t = \int_0^t S(h)S(s)x \, ds = \int_0^t S(h+s)x \, ds = \int_h^{t+h} S(r)x \, dr.$$

Hence

$$S(h)x_t - x_t = \int_h^{t+h} S(r)x \, dr - \int_0^t S(r)x \, dr = \int_t^{t+h} S(r)x \, dr - \int_0^h S(r)x \, dr.$$

Dividing by h and letting $h \downarrow 0$, continuity of $r \mapsto S(r)x$ yields

$$\lim_{h \downarrow 0} \frac{S(h)x_t - x_t}{h} = S(t)x - S(0)x = S(t)x - x.$$

Therefore the defining limit exists, so $x_t \in \mathcal{D}(L)$ for all $t > 0$. Since $\frac{1}{t} x_t \to x$ as

$t \downarrow 0$, this proves that $\mathcal{D}(L)$ is dense in $\mathcal{B}$. $\square$

Let $\{S(t)\}_{t\geqslant 0}$ be a C_0-semigroup on a Banach space $\mathcal{B}$ with generator L. We consider the abstract Cauchy problem

$$\frac{d}{dt}x(t) = Lx(t), \qquad x(0) = x. \tag{2.2}$$

The next theorem shows that for initial data $x \in \mathcal{D}(L)$ (which is dense in $\mathcal{B}$), the semigroup orbit $u(t) = S(t)x$ is a classical solution.

Theorem 2.8. *Let $\{S(t)\}_{t\geqslant 0}$ be a C_0-semigroup with generator L. Then:*

(i) $S(t)\mathcal{D}(L) \subset \mathcal{D}(L)$ for every $t \geqslant 0$;

(ii) for every $x \in \mathcal{D}(L)$ and $t \geqslant 0$,

$$LS(t)x = S(t)Lx;$$

(iii) for every $x \in \mathcal{D}(L)$, the map $t \mapsto S(t)x$ is differentiable on $(0, \infty)$ and

$$\partial_t S(t)x = LS(t)x = S(t)Lx, \qquad t > 0.$$

Proof. Fix $x \in \mathcal{D}(L)$. We first prove (i) and (ii). Let $t \geqslant 0$. By the semigroup property,

$$\frac{S(t+h)x - S(t)x}{h} = S(t)\frac{S(h)x - x}{h}.$$

Since $x \in \mathcal{D}(L)$, the limit of $\frac{S(h)x-x}{h}$ exists in $\mathcal{B}$ as $h \downarrow 0$ and equals Lx. Because $S(t)$ is bounded, we obtain

$$LS(t)x = \lim_{h\downarrow 0} \frac{S(h)S(t)x - S(t)x}{h} = S(t)Lx.$$

In particular, the defining limit exists for $S(t)x$, so $S(t)x \in \mathcal{D}(L)$. This proves (i), and the identity above gives (ii).

We now prove (iii). Fix $t > 0$. For the right derivative, we have

$$\frac{S(t+h)x - S(t)x}{h} = S(t)\frac{S(h)x - x}{h} \xrightarrow[h\downarrow 0]{} S(t)Lx = LS(t)x.$$

For the left derivative, take $0 < h < t$ and write

$$\frac{S(t)x - S(t-h)x}{h} = \frac{S(t-h)S(h)x - S(t-h)x}{h} = S(t-h)\frac{S(h)x - x}{h}.$$

As $h \downarrow 0$, we have $S(t-h) \to S(t)$ strongly and $\frac{S(h)x-x}{h} \to Lx$ in $\mathcal{B}$. Using local boundedness of $\|S(\cdot)\|$ on $[t-1,t]$, the product converges to $S(t)Lx = LS(t)x$. Therefore, the left and right derivatives agree, so $t \mapsto S(t)x$ is differentiable at t and

$$\partial_t S(t)x = LS(t)x = S(t)Lx.$$

$\square$

Corollary 2.9. *Suppose $x : [0,T] \to \mathcal{D}(L)$ satisfies (2.2). Then $x(t) = S(t)x_0$ for all $t \in [0,T]$. In particular, no two distinct C_0-semigroups can have the same generator.*

Proof. Fix $T > 0$ and define $f : [0,T] \to \mathcal{B}$ by

$$f(t) := S(t)\, x(T-t).$$

Let $t \in (0,T)$ and let $h > 0$ be small. Using the semigroup property and adding/subtracting terms, we compute

$$\begin{aligned}\frac{f(t+h)-f(t)}{h} &= \frac{S(t+h)x(T-t-h) - S(t)x(T-t)}{h} \\ &= S(t)\left[\frac{S(h)x(T-t-h) - x(T-t-h)}{h}\right] \\ &\quad + S(t)\left[\frac{x(T-t-h) - x(T-t)}{h}\right].\end{aligned}$$

As $h \downarrow 0$, we have $x(T-t-h) \to x(T-t)$ and $x(T-t) \in \mathcal{D}(L)$, hence

$$\frac{S(h)x(T-t-h) - x(T-t-h)}{h} \to Lx(T-t).$$

Moreover, since x is differentiable and satisfies $\partial_t x = Lx$,

$$\frac{x(T-t-h) - x(T-t)}{h} \to -\partial_t x(T-t) = -Lx(T-t).$$

Therefore,

$$\lim_{h \downarrow 0} \frac{f(t+h)-f(t)}{h} = S(t)Lx(T-t) - S(t)Lx(T-t) = 0.$$

A similar computation for $h \uparrow 0$ yields the same limit. Hence $f'(t) = 0$ for

$t \in (0,T)$, so f is constant on $[0,T]$. Evaluating at $t=0$ and $t=T$ gives

$$x(T) = f(0) = f(T) = S(T)x_0.$$

Since T was arbitrary, it follows that $x(t) = S(t)x_0$ for all $t \in [0,T]$. $\square$

Lastly, we conclude this subsection with the Hille-Yosida theorem, which characterises the generators of C_0-semigroups. Before stating it, we introduce the resolvent of an operator and explain its connection to semigroups via Laplace transforms.

Definition 2.10. Let $L : \mathcal{D}(L) \subset \mathcal{B} \to \mathcal{B}$ be a (not necessarily bounded) linear operator. The **resolvent set** $\rho(L)$ is the set of $\lambda \in \mathbb{C}$ such that

$$\lambda I - L : \mathcal{D}(L) \to \mathcal{B}$$

is bijective and its inverse extends to a bounded operator on $\mathcal{B}$. For $\lambda \in \rho(L)$ we write

$$R_\lambda := (\lambda I - L)^{-1} \in \mathcal{L}(\mathcal{B}),$$

and call R_λ the **resolvent operator**.

The following properties of the resolvent set and resolvent operator are fundamental in semigroup theory. We state them without proof.

Theorem 2.11. *Let $\mathcal{B}$ be a Banach space and let $L : \mathcal{D}(L) \to \mathcal{B}$ be a linear operator.*

(i) ***Openness:*** *$\rho(L)$ is open in $\mathbb{C}$.*

(ii) ***Resolvent identity:*** *for $\lambda, \mu \in \rho(L)$,*

$$R_\lambda - R_\mu = (\mu - \lambda) R_\lambda R_\mu.$$

(iii) ***Uniqueness:*** *knowing R_λ for one $\lambda \in \rho(L)$ determines L uniquely.*

The next theorem makes the link between semigroups and resolvents precise: the resolvent of the generator can be recovered as the Laplace transform of the semigroup. This identity will be the key tool when we later move to the Hille-Yosida theorem, since it translates analytic bounds on the resolvent into growth bounds for the semigroup (and vice versa).

Theorem 2.12. *Let $\{S(t)\}_{t\geqslant 0}$ be a C_0-semigroup on $\mathcal{B}$ with generator L, and assume that $\|S(t)\| \leqslant Me^{at}$ for all $t \geqslant 0$. If $\lambda \in \mathbb{C}$ satisfies $\operatorname{Re}\lambda > a$, then $\lambda \in \rho(L)$ and*

$$R_\lambda x = \int_0^\infty e^{-\lambda t} S(t)x\,dt, \qquad x \in \mathcal{B}, \tag{2.3}$$

where the integral is understood in the Bochner sense in $\mathcal{B}$.

Proof. Fix λ with $\operatorname{Re}\lambda > a$ and define, for $x \in \mathcal{B}$,

$$Z_\lambda x := \int_0^\infty e^{-\lambda t} S(t)x\,dt.$$

We will show that Z_λ is a bounded operator and that it is precisely the inverse of $\lambda I - L$.

We begin by checking that the integral defining Z_λ is well behaved. Using the growth bound on $S(t)$,

$$\|e^{-\lambda t} S(t)\| \leqslant Me^{-(\operatorname{Re}\lambda - a)t},$$

so the integrand decays exponentially and the Bochner integral converges in norm. In particular,

$$\|Z_\lambda\| \leqslant \int_0^\infty Me^{-(\operatorname{Re}\lambda - a)t}\,dt = \frac{M}{\operatorname{Re}\lambda - a}.$$

Next, we verify that Z_λ is a right inverse for $\lambda I - L$. Fix $x \in \mathcal{B}$ and $h > 0$. By the semigroup property and a change of variables,

$$S(h)Z_\lambda x = \int_0^\infty e^{-\lambda t} S(t+h)x\,dt = e^{\lambda h}\int_h^\infty e^{-\lambda t} S(t)x\,dt.$$

Subtracting $Z_\lambda x$ and dividing by h yields

$$\frac{S(h)Z_\lambda x - Z_\lambda x}{h} = \frac{e^{\lambda h} - 1}{h}\int_0^\infty e^{-\lambda t} S(t)x\,dt - \frac{e^{\lambda h}}{h}\int_0^h e^{-\lambda t} S(t)x\,dt.$$

As $h \downarrow 0$, the first term converges to $\lambda Z_\lambda x$. For the second term, continuity of $t \mapsto S(t)x$ at 0 implies

$$\frac{1}{h}\int_0^h e^{-\lambda t} S(t)x\,dt \to x,$$

so the second term converges to x. Therefore the defining limit exists and

$$LZ_\lambda x = \lambda Z_\lambda x - x,$$

which is the same as saying $(\lambda I - L)Z_\lambda x = x$ for every $x \in \mathcal{B}$. Hence $\lambda I - L$ is surjective.

It remains to show injectivity. Suppose $(\lambda I - L)x = 0$ for some $x \in \mathcal{D}(L)$, i.e. $Lx = \lambda x$. Define $u(t) := e^{\lambda t}x$. Then $u'(t) = \lambda u(t) = Lu(t)$ and $u(0) = x$. By uniqueness of classical solutions (Corollary 2.9), we must have $u(t) = S(t)x$. Taking norms gives

$$\|S(t)x\| = \|e^{\lambda t}x\| = e^{\operatorname{Re}\lambda\, t}\|x\|.$$

If $x \neq 0$, then $\|S(t)\| \geqslant e^{\operatorname{Re}\lambda\, t}$ for all $t \geqslant 0$, contradicting $\|S(t)\| \leqslant Me^{at}$ when $\operatorname{Re}\lambda > a$ and t is large. Hence $x = 0$, so $\lambda I - L$ is injective.

Since $\lambda I - L$ is bijective and Z_λ is bounded and satisfies $(\lambda I - L)Z_\lambda = I$, we conclude that $\lambda \in \rho(L)$ and $R_\lambda = Z_\lambda$, i.e.

$$R_\lambda x = \int_0^\infty e^{-\lambda t} S(t)x\, dt.$$

□

Next, we collect a few basic properties of generators of C_0-semigroups. These will be used repeatedly in the proof of the Hille-Yosida theorem, where one works primarily with resolvents and closedness properties rather than with the semigroup directly.

Definition 2.13. An operator $L : \mathcal{D}(L) \to \mathcal{B}$ is **closed** if its graph

$$\mathcal{G}(L) := \{(x, Lx) : x \in \mathcal{D}(L)\} \subset \mathcal{B} \times \mathcal{B}$$

is closed. Equivalently, if $x_n \in \mathcal{D}(L)$, $x_n \to x$, and $Lx_n \to y$ in $\mathcal{B}$, then $x \in \mathcal{D}(L)$ and $Lx = y$.

Theorem 2.14 (Double Cauchy criterion)**.** *An operator L is closed if and only if whenever $\{x_n\} \subset \mathcal{D}(L)$ satisfies that $\{x_n\}$ is Cauchy in $\mathcal{B}$ and $\{Lx_n\}$ is Cauchy in $\mathcal{B}$, then the limits exist and satisfy*

$$x := \lim_{n\to\infty} x_n \in \mathcal{D}(L), \qquad Lx = \lim_{n\to\infty} Lx_n.$$

One of the main properties of generators of C_0-semigroups is that they are always closed operators.

Theorem 2.15. *Let L be the generator of a C_0-semigroup $\{S(t)\}_{t\geqslant 0}$ satisfying $\|S(t)\| \leqslant Me^{at}$. Then L is closed.*

Proof. We may reduce to the case $a = 0$. Indeed, define $\widetilde{S}(t) := e^{-at}S(t)$, which is again a C_0-semigroup. Its generator is $\widetilde{L} = L - aI$, and L is closed if and only if $\widetilde{L}$ is closed.

Assume now $a = 0$. By Theorem 2.12, we have $1 \in \rho(L)$, hence $(I - L)^{-1} \in \mathcal{L}(\mathcal{B})$. Let $\{x_n\} \subset \mathcal{D}(L)$ be such that $x_n \to x$ and $Lx_n \to y$ in $\mathcal{B}$. Set $z_n := (I - L)x_n = x_n - Lx_n$. Then $z_n \to x - y$. Applying $(I - L)^{-1}$ and using continuity, we obtain

$$x_n = (I - L)^{-1}z_n \longrightarrow (I - L)^{-1}(x - y).$$

By uniqueness of limits, $x = (I-L)^{-1}(x-y)$, so $x \in \mathcal{D}(L)$ and $(I-L)x = x-y$, i.e. $Lx = y$. Thus L is closed. $\square$

The final ingredient we need before stating the Hille-Yosida theorem is a quantitative estimate on the resolvent: not only is R_λ bounded when $\operatorname{Re}\lambda > a$, but its powers satisfy uniform bounds. These bounds will become the resolvent growth condition appearing in Hille-Yosida.

Theorem 2.16. *Let L be the generator of a C_0-semigroup $\{S(t)\}_{t\geqslant 0}$ with $\|S(t)\| \leqslant Me^{at}$. If $\operatorname{Re}\lambda > a$, then for every $n \in \mathbb{N}$,*

$$\|R_\lambda^n\| \leqslant \frac{M}{(\operatorname{Re}\lambda - a)^n}.$$

Proof. Using the (2.3) and iterating n times, we obtain

$$R_\lambda^n x = \int_0^\infty \cdots \int_0^\infty e^{-\lambda(t_1+\cdots+t_n)} S(t_1 + \cdots + t_n)x \, dt_1 \cdots dt_n.$$

Taking norms and using $\|S(t)\| \leqslant Me^{at}$ gives

$$\|R_\lambda^n x\| \leqslant M \int_0^\infty \cdots \int_0^\infty e^{-(\operatorname{Re}\lambda - a)(t_1+\cdots+t_n)} \, dt_1 \cdots dt_n \, \|x\|.$$

Therefore,

$$\|R_\lambda^n\| \leqslant M \int_0^\infty \cdots \int_0^\infty e^{-(\operatorname{Re}\lambda - a)(t_1+\cdots+t_n)} \, dt_1 \cdots dt_n = \frac{M}{(\operatorname{Re}\lambda - a)^n}.$$

$\square$

The Hille-Yosida theorem shows that the properties we have derived for generators of C_0-semigroups are not only necessary but also sufficient. In other words, resolvent bounds completely characterise which (densely defined, closed) operators generate a C_0-semigroup with a prescribed growth bound.

Theorem 2.17 (Hille-Yosida)**.** *Let $L : \mathcal{D}(L) \subset \mathcal{B} \to \mathcal{B}$ be densely defined and closed. Then L generates a C_0-semigroup $\{S(t)\}_{t \geqslant 0}$ satisfying $\|S(t)\| \leqslant Me^{at}$ if and only if*

$$\{\lambda \in \mathbb{C} : \operatorname{Re} \lambda > a\} \subset \rho(L) \quad \text{and} \quad \|R_\lambda^n\| \leqslant \frac{M}{(\operatorname{Re} \lambda - a)^n} \quad \text{for all } n \geqslant 1.$$

Sketch of proof. The necessity of the resolvent conditions was already established in Theorems 2.12 and 2.16. We now explain, at a high level, why these resolvent bounds are also sufficient to construct a C_0-semigroup with generator L.

Assume therefore that L is densely defined and closed, that $\{\operatorname{Re} \lambda > a\} \subset \rho(L)$, and that

$$\|R_\lambda^n\| \leqslant \frac{M}{(\operatorname{Re} \lambda - a)^n}, \qquad n \geqslant 1.$$

For real $\lambda > a$ we introduce the bounded operator

$$L_\lambda := \lambda L R_\lambda \in \mathcal{L}(\mathcal{B}).$$

Using $(\lambda I - L)R_\lambda = I$ we have $LR_\lambda = \lambda R_\lambda - I$, hence the concrete formula

$$L_\lambda = \lambda(\lambda R_\lambda - I) = \lambda^2 R_\lambda - \lambda I.$$

The point is that L_λ is bounded, so it generates a uniformly continuous semigroup, and as $\lambda \to \infty$ it behaves more and more like L on $\mathcal{D}(L)$.

To make this precise, we first recall a standard consequence of the resolvent estimates: for every $x \in \mathcal{B}$,

$$\lambda R_\lambda x \to x \quad \text{in } \mathcal{B} \quad \text{as } \lambda \to \infty.$$

Equivalently, $LR_\lambda x = \lambda R_\lambda x - x \to 0$ for every $x \in \mathcal{B}$. In particular, if $x \in \mathcal{D}(L)$, then $Lx \in \mathcal{B}$ and we may apply this convergence to Lx to obtain

$$\|L_\lambda x - Lx\| = \|\lambda L R_\lambda x - Lx\| = \|\lambda R_\lambda L x - Lx\| = \|(\lambda R_\lambda - I)Lx\| \to 0.$$

So L_λ approximates L on $\mathcal{D}(L)$.

Since L_λ is bounded, we can define the associated uniformly continuous semigroup by the exponential series

$$S_\lambda(t) := e^{tL_\lambda} = \sum_{n=0}^{\infty} \frac{t^n}{n!} L_\lambda^n, \qquad t \geqslant 0.$$

The resolvent bounds imply uniform control of these semigroups. Concretely, one shows that there is a growth estimate of the form

$$\|S_\lambda(t)\| \leqslant M e^{c_\lambda t} \quad \text{for all } t \geqslant 0,$$

with constants c_λ that can be chosen so that $c_\lambda \downarrow a$ as $\lambda \to \infty$. In particular, for λ large we may simply keep the rough bound $\|S_\lambda(t)\| \leqslant M e^{2at}$ (when $a > 0$).

With these uniform bounds in hand, one can take limits. Fix $t \geqslant 0$. Using the approximation $L_\lambda x \to Lx$ on $\mathcal{D}(L)$ together with the uniform growth estimate for S_λ, one proves that for each $x \in \mathcal{D}(L)$ the family $S_\lambda(t)x$ is Cauchy as $\lambda \to \infty$. Since $\mathcal{D}(L)$ is dense and $\sup_\lambda \|S_\lambda(t)\| < \infty$ for fixed t, the limit extends uniquely to every $x \in \mathcal{B}$. We therefore define

$$S(t)x := \lim_{\lambda \to \infty} S_\lambda(t)x, \qquad x \in \mathcal{B}.$$

The semigroup property is inherited from the approximations: since $S_\lambda(t + s) = S_\lambda(t)S_\lambda(s)$ for every λ, passing to the limit yields $S(t + s) = S(t)S(s)$. Strong continuity follows from the fact that each S_λ is strongly continuous and that the convergence $S_\lambda(t)x \to S(t)x$ is uniform for t in compact intervals (for each fixed x), which allows one to interchange limits in t and λ.

Finally, we identify the generator. Let $\widetilde{L}$ denote the generator of the semigroup $\{S(t)\}_{t \geqslant 0}$ just constructed. For the bounded semigroups we have the variation-of-constants identity

$$\frac{S_\lambda(t)x - x}{t} = \frac{1}{t} \int_0^t S_\lambda(s) L_\lambda x \, ds, \qquad x \in \mathcal{B}.$$

Fix $x \in \mathcal{D}(L)$. Using $L_\lambda x \to Lx$ together with the uniform bounds on $S_\lambda(s)$, we may pass $\lambda \to \infty$ inside the integral to obtain

$$\frac{S(t)x - x}{t} = \frac{1}{t} \int_0^t S(s)Lx\, ds.$$

Letting $t \downarrow 0$ then shows that $x \in \mathcal{D}(\widetilde{L})$ and $\widetilde{L}x = Lx$ for all $x \in \mathcal{D}(L)$. To conclude equality of operators (not just an extension), one uses the resolvent characterization: both $\lambda I - L$ and $\lambda I - \widetilde{L}$ are bijective for $\operatorname{Re}\lambda > a$, hence their inverses coincide on $\mathcal{B}$, which forces $\mathcal{D}(L) = \mathcal{D}(\widetilde{L})$ and $\widetilde{L} = L$. $\square$

2.2 Adjoint semigroups on dual spaces

Let $\mathcal{B}$ be a Banach space with dual $\mathcal{B}^*$. We write the duality pairing as

$$\langle \ell, x\rangle := \ell(x), \qquad \ell \in \mathcal{B}^*,\ x \in \mathcal{B}.$$

Definition 2.18. Let $L : \mathcal{D}(L) \subset \mathcal{B} \to \mathcal{B}$ be a linear operator. The **adjoint operator** $L^* : \mathcal{D}(L^*) \subset \mathcal{B}^* \to \mathcal{B}^*$ is defined by the domain

$$\mathcal{D}(L^*) := \{\ell \in \mathcal{B}^* :\ \exists\, \ell' \in \mathcal{B}^* \text{ such that } \langle \ell, Lx\rangle = \langle \ell', x\rangle \ \ \forall\, x \in \mathcal{D}(L)\}\,.$$

For $\ell \in \mathcal{D}(L^*)$, we set $L^*\ell := \ell'$, equivalently

$$\langle L^*\ell, x\rangle = \langle \ell, Lx\rangle, \qquad \forall\, x \in \mathcal{D}(L).$$

If $\{S(t)\}_{t\geqslant 0}$ is a C_0-semigroup on $\mathcal{B}$, its **adjoint family** $\{S^*(t)\}_{t\geqslant 0}$ on $\mathcal{B}^*$ is defined by

$$\langle S^*(t)\ell, x\rangle := \langle \ell, S(t)x\rangle, \qquad \ell \in \mathcal{B}^*,\ x \in \mathcal{B}.$$

Theorem 2.19 (Adjoint relations)**.** *Let $\{S(t)\}_{t\geqslant 0}$ be a C_0-semigroup on $\mathcal{B}$ with generator L. Then for every $\ell \in \mathcal{D}(L^*)$:*

*(i) (*Invariance and commutation*) For every $t \geqslant 0$,*

$$S^*(t)\ell \in \mathcal{D}(L^*) \quad \textit{and} \quad L^*S^*(t)\ell = S^*(t)L^*\ell.$$

*(ii) (*Differentiability of the pairing*) For every $x \in \mathcal{B}$, the map $t \mapsto \langle \ell, S(t)x\rangle$ is differentiable and*

$$\frac{d}{dt}\langle \ell, S(t)x\rangle = \langle L^*\ell, S(t)x\rangle, \qquad t \geqslant 0.$$

Proof. Fix $\ell \in \mathcal{D}(L^*)$. We begin by showing that $\mathcal{D}(L^*)$ is invariant under the adjoint semigroup and that L^* commutes with $S^*(t)$ on this domain. Take $t \geqslant 0$ and $x \in \mathcal{D}(L)$. Using the semigroup property and the identity $LS(t)x = S(t)Lx$, we compute

$$\langle S^*(t)\ell, Lx\rangle = \langle \ell, S(t)Lx\rangle = \langle \ell, LS(t)x\rangle = \langle L^*\ell, S(t)x\rangle = \langle S^*(t)L^*\ell, x\rangle.$$

This shows that $S^*(t)\ell \in \mathcal{D}(L^*)$ and that $L^*S^*(t)\ell = S^*(t)L^*\ell$, proving (i).

With (i) in hand, we next verify the differentiability statement when x lies in the generator domain. Let $x \in \mathcal{D}(L)$. Then

$$\begin{aligned}\frac{d}{dt}\langle \ell, S(t)x\rangle &= \lim_{h\to 0}\frac{\langle \ell, S(t+h)x\rangle - \langle \ell, S(t)x\rangle}{h}\\ &= \lim_{h\to 0}\left\langle \ell, \frac{S(h)S(t)x - S(t)x}{h}\right\rangle\\ &= \langle \ell, LS(t)x\rangle = \langle L^*\ell, S(t)x\rangle,\end{aligned}$$

where the last equality uses $\langle \ell, Ly\rangle = \langle L^*\ell, y\rangle$ with $y = S(t)x \in \mathcal{D}(L)$.

Finally, to extend the differentiability to arbitrary $x \in \mathcal{B}$, fix $T > 0$ and choose $x_n \in \mathcal{D}(L)$ with $x_n \to x$ in $\mathcal{B}$. Define

$$f(t) := \langle \ell, S(t)x\rangle, \qquad f_n(t) := \langle \ell, S(t)x_n\rangle,$$

and

$$g(t) := \langle L^*\ell, S(t)x\rangle, \qquad g_n(t) := \langle L^*\ell, S(t)x_n\rangle.$$

By local boundedness of C_0-semigroups, $\sup_{t\in[0,T]}\|S(t)\| < \infty$, hence $f_n \to f$ and $g_n \to g$ uniformly on $[0,T]$. From the previous paragraph, for each n and all $t \in [0,T]$,

$$f_n(t) - f_n(0) = \int_0^t g_n(s)\,ds.$$

Passing to the limit and using uniform convergence yields

$$f(t) - f(0) = \int_0^t g(s)\,ds.$$

Therefore f is differentiable on $[0,T]$ with $f'(t) = g(t)$. Since $T > 0$ was arbitrary, the identity holds for all $t \geqslant 0$. □

One might hope that $\{S^*(t)\}_{t\geqslant 0}$ is again a C_0-semigroup on $\mathcal{B}^*$. This is, however, false in general, as the following example shows.

Example 2.20 (The adjoint semigroup need not be strongly continuous on $\mathcal{B}^*$)**.** Let $\mathcal{B} = \mathcal{C}([0,1],\mathbb{R})$ equipped with the supremum norm $\|\cdot\|_\infty$. By the Riesz representation theorem, $\mathcal{B}^*$ can be identified with the finite signed Borel measures on $[0,1]$, equipped with the total variation norm

$$\|\mu\|_{TV} := \sup_{\|\phi\|_\infty \leqslant 1} \left| \int_0^1 \phi \, d\mu \right|.$$

Consider the Neumann heat semigroup, defined for $f \in \mathcal{C}([0,1])$ by

$$(S(t)f)(x) = \int_0^1 p_t^N(x,y) f(y) \, dy,$$

where

$$p_t^N(x,y) = 1 + 2\sum_{n=1}^{\infty} e^{-n^2\pi^2 t} \cos(n\pi x)\cos(n\pi y).$$

We claim that

$$S^*(t)\delta_0 \not\to \delta_0 \quad \text{in } \|\cdot\|_{TV} \text{ as } t \downarrow 0,$$

and therefore $\{S^*(t)\}_{t\geqslant 0}$ is not strongly continuous on $\mathcal{B}^*$.

Fix $t>0$. We show that $\|S^*(t)\delta_0 - \delta_0\|_{TV} \geqslant 2$. For $r>0$ define the test function

$$\phi_r(x) = \begin{cases} 1, & x \in [0,r], \\ 1 - 2\frac{x-r}{r}, & x \in [r,2r], \\ -1, & x \in [2r,1]. \end{cases} \qquad \text{so that} \qquad \|\phi_r\|_\infty = 1.$$

By the definition of the total variation norm and of the adjoint semigroup,

$$\begin{aligned} \|S^*(t)\delta_0 - \delta_0\|_{TV} &\geqslant \left| \int_0^1 \phi_r \, d\big(S^*(t)\delta_0 - \delta_0\big) \right| \\ &= \big|(S(t)\phi_r)(0) - \phi_r(0)\big| = \big|(S(t)\phi_r)(0) - 1\big|. \end{aligned}$$

Using $\phi_r \leqslant 1$ on $[0, 2r]$ and $\phi_r = -1$ on $[2r, 1]$, we obtain

$$\begin{aligned}(S(t)\phi_r)(0) &= \int_0^1 p_t^N(0,y)\phi_r(y)\,dy \\ &\leqslant \int_0^{2r} p_t^N(0,y)\,dy - \int_{2r}^1 p_t^N(0,y)\,dy \\ &= 2\int_0^{2r} p_t^N(0,y)\,dy - 1,\end{aligned}$$

where we used $\int_0^1 p_t^N(0,y)\,dy = 1$. Choose $r > 0$ so small that $\int_0^{2r} p_t^N(0,y)\,dy < \varepsilon$. Then

$$(S(t)\phi_r)(0) \leqslant 2\varepsilon - 1 \quad \Longrightarrow \quad \big|(S(t)\phi_r)(0) - 1\big| \geqslant 2 - 2\varepsilon.$$

Letting $\varepsilon \downarrow 0$ yields $\|S^*(t)\delta_0 - \delta_0\|_{TV} \geqslant 2$. In particular, $S^*(t)\delta_0 \not\to \delta_0$ in $\|\cdot\|_{TV}$ as $t \downarrow 0$.

To recover strong continuity, it is natural to restrict the adjoint family $S^*(t)$ to the norm-closure of $\mathcal{D}(L^*)$:

$$\mathcal{B}^\dagger := \overline{\mathcal{D}(L^*)}^{\|\cdot\|_{\mathcal{B}^*}} \subset \mathcal{B}^*.$$

Theorem 2.21 (Phillips)**.** *For every $\ell \in \mathcal{B}^*$ there exists a sequence $(\ell_n) \subset \mathcal{B}^\dagger$ such that*

$$\ell_n(x) \to \ell(x) \qquad \text{for all } x \in \mathcal{B}.$$

Equivalently, $\mathcal{B}^\dagger$ is weak-$*$ dense in $\mathcal{B}^*$.

Proof. Let $R_n := (nI - L)^{-1}$ for n sufficiently large (for instance, $n > a$ under the growth bound), and define

$$\ell_n := n\,R_n^*\ell.$$

Since R_n maps $\mathcal{B}$ into $\mathcal{D}(L)$, its adjoint R_n^* maps $\mathcal{B}^*$ into $\mathcal{D}(L^*)$. Hence $\ell_n \in \mathcal{D}(L^*) \subset \mathcal{B}^\dagger$.

For $x \in \mathcal{B}$ we have

$$\ell_n(x) = \langle \ell_n, x\rangle = \langle \ell, nR_n x\rangle.$$

From the resolvent identity $(nI - L)R_n = I$ we obtain $nR_n - I = LR_n$. Therefore,

$$\|nR_n x - x\| = \|LR_n x\|.$$

As shown in the Hille-Yosida argument, $LR_n x \to 0$ for each fixed $x \in \mathcal{B}$, hence $nR_n x \to x$. Consequently, $\ell_n(x) \to \ell(x)$ for every $x \in \mathcal{B}$. $\square$

Theorem 2.22 (Adjoint semigroup on $\mathcal{B}^\dagger$)**.** *Let $\{S(t)\}_{t\geqslant 0}$ be a C_0-semigroup on $\mathcal{B}$ with generator L, and set $\mathcal{B}^\dagger = \overline{\mathcal{D}(L^*)}^{\|\cdot\|_{\mathcal{B}^*}}$. Then $\{S^*(t)\}_{t\geqslant 0}$ restricts to a C_0-semigroup on the Banach space $\mathcal{B}^\dagger$. Its generator $L^\dagger$ is the part of L^* in $\mathcal{B}^\dagger$, namely*

$$\mathcal{D}(L^\dagger) = \{\ell \in \mathcal{D}(L^*) :\ L^*\ell \in \mathcal{B}^\dagger\}, \qquad L^\dagger \ell := L^*\ell.$$

Proof. We check strong continuity of $S^*(t)$ on $\mathcal{B}^\dagger$ using the standard criterion in Theorem 2.3.

The first condition is growth bound. Since $\|S^*(t)\| = \|S(t)\|$, we have

$$\|S^*(t)\|_{\mathcal{B}^\dagger} \leqslant \|S^*(t)\|_{\mathcal{B}^*} = \|S(t)\| \leqslant Me^{at}, \qquad t \geqslant 0.$$

The second condition is the continuity at $t = 0$ on a dense subset. Let $\ell \in \mathcal{D}(L^*)$ and $x \in \mathcal{B}$. By Theorem 2.19 and the commutation $L^*S^*(s)\ell = S^*(s)L^*\ell$, we have

$$\langle S^*(t)\ell - \ell, x\rangle = \int_0^t \langle L^*S^*(s)\ell, x\rangle\, ds = \int_0^t \langle S^*(s)L^*\ell, x\rangle\, ds.$$

Taking the supremum over $\|x\| \leqslant 1$ yields

$$\begin{aligned}
\|S^*(t)\ell - \ell\| &= \sup_{\|x\|\leqslant 1} \big|\langle S^*(t)\ell - \ell, x\rangle\big| \\
&\leqslant \int_0^t \sup_{\|x\|\leqslant 1} \big|\langle S^*(s)L^*\ell, x\rangle\big|\, ds = \int_0^t \|S^*(s)L^*\ell\|\, ds \\
&\leqslant \|L^*\ell\| \int_0^t \|S^*(s)\|\, ds \leqslant \|L^*\ell\| \int_0^t Me^{as}\, ds \xrightarrow[t\downarrow 0]{} 0.
\end{aligned}$$

Thus $S^*(t)\ell \to \ell$ for all $\ell \in \mathcal{D}(L^*)$, which is dense in $\mathcal{B}^\dagger$ by definition.

The growth bound and continuity at 0 on a dense set imply that $S^*(t)$ is a C_0-semigroup on $\mathcal{B}^\dagger$.

Finally, the resolvent of the generator on $\mathcal{B}^\dagger$ is given by the Laplace formula

$$R_\lambda^\dagger \ell = \int_0^\infty e^{-\lambda t} S^*(t)\ell \, dt,$$

which is exactly the restriction of R_λ^* to $\mathcal{B}^\dagger$. Hence the generator is the part of L^* in $\mathcal{B}^\dagger$, i.e. the stated operator $L^\dagger$. $\square$

Next we discuss the special case of self-adjoint operators on Hilbert spaces. In this setting the picture is considerably simpler, since the Riesz representation theorem allows us to identify $\mathcal{H}$ with its dual.

Let $\mathcal{H}$ be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$.

Definition 2.23. A densely defined operator $L : \mathcal{D}(L) \subset \mathcal{H} \to \mathcal{H}$ is **self-adjoint** if $\mathcal{D}(L) = \mathcal{D}(L^*)$ and $L = L^*$, equivalently,

$$\langle Lx, y \rangle = \langle x, Ly \rangle \qquad \text{for all } x, y \in \mathcal{D}(L).$$

From now on we assume, for convenience, that L is *bounded above*: there exists $C \in \mathbb{R}$ such that

$$\langle Lx, x \rangle \leqslant C\|x\|^2, \qquad x \in \mathcal{D}(L).$$

Replacing L by $L - CI$ reduces us to the case

$$\langle Lx, x \rangle \leqslant 0, \qquad x \in \mathcal{D}(L),$$

which we will refer to as *negative definite* in this context.

Theorem 2.24 (Spectral theorem, simplified form)**.** *If L is self-adjoint on $\mathcal{H}$, then there exist a measure space $(\mathcal{M}, \mu)$ and a unitary operator $K : \mathcal{H} \to L^2(\mathcal{M}, \mu)$ such that*

$$L = K^{-1} M_{f_L} K,$$

where M_{f_L} is the multiplication operator

$$(M_{f_L} g)(m) = f_L(m) g(m),$$

for some real-valued measurable function f_L.

Using the spectral theorem, we may define a functional calculus for self-adjoint operators. For any bounded Borel function $f : \mathbb{R} \to \mathbb{R}$, set

$$f(L) := K^{-1} M_{f\circ f_L} K.$$

Then $(fg)(L) = f(L)g(L)$ and

$$\|f(L)\| = \|f \circ f_L\|_{L^\infty(\mathcal{M},\mu)}.$$

Remark (Semigroup generated by a self-adjoint operator). If L is self-adjoint and $\langle Lx, x\rangle \leqslant 0$, then $S(t) := e^{tL}$ defines a C_0-semigroup of contractions on $\mathcal{H}$, and

$$e^{tL} = K^{-1} M_{e^{tf_L}} K.$$

Remark (Negative definiteness implies $f_L \leqslant 0$). If $\langle Lx, x\rangle \leqslant 0$ for all $x \in \mathcal{D}(L)$, then $f_L(m) \leqslant 0$ for μ-a.e. m. Indeed, for $x \in \mathcal{D}(L)$,

$$0 \geqslant \langle Lx, x\rangle = \langle KLx, Kx\rangle = \langle M_{f_L} Kx, Kx\rangle = \int_{\mathcal{M}} f_L(m)\, |Kx(m)|^2 \, d\mu(m),$$

which forces $f_L \leqslant 0$ almost everywhere.

The next result shows that self-adjoint semigroups have a strong regularising effect.

Theorem 2.25. *Let L be self-adjoint and satisfy $\langle Lx, x\rangle \leqslant 0$, and let $S(t) = e^{tL}$ be the generated semigroup. Then for every $\alpha > 0$ and $t > 0$,*

$$S(t)\mathcal{H} \subset \mathcal{D}\big((1-L)^\alpha\big),$$

and there exists $C_\alpha > 0$ such that for every $t > 0$

$$\|(1-L)^\alpha S(t)\| \leqslant C_\alpha\big(1 + t^{-\alpha}\big).$$

Proof. By the spectral theorem and functional calculus, and using $f_L \leqslant 0$ a.e.,

$$\|(1-L)^\alpha S(t)\| = \|(1-f_L)^\alpha e^{tf_L}\|_{L^\infty} = \sup_{\lambda \geqslant 0} (1+\lambda)^\alpha e^{-\lambda t},$$

where we set $\lambda := -f_L \geqslant 0$.

Choose $C_\alpha > 0$ such that $(1+\lambda)^\alpha \leqslant C_\alpha(1+\lambda^\alpha)$ for all $\lambda \geqslant 0$. Then

$$\sup_{\lambda\geqslant 0}(1+\lambda)^\alpha e^{-\lambda t} \leqslant C_\alpha \left(\sup_{\lambda \geqslant 0} e^{-\lambda t} + \sup_{\lambda\geqslant 0} \lambda^\alpha e^{-\lambda t} \right).$$

The first term is at most 1. For the second term,

$$\sup_{\lambda \geqslant 0} \lambda^\alpha e^{-\lambda t} = t^{-\alpha} \sup_{x \geqslant 0} x^\alpha e^{-x} \leqslant t^{-\alpha}\, \alpha^\alpha e^{-\alpha}.$$

Combining these bounds yields $\|(1-L)^\alpha S(t)\| \leqslant C'_\alpha(1+t^{-\alpha})$. In particular, for every $x \in \mathcal{H}$ we have $(1-L)^\alpha S(t)x \in \mathcal{H}$, i.e. $S(t)x \in \mathcal{D}((1-L)^\alpha)$. $\square$

2.3 Analytic semigroups

An important class of C_0-semigroups are those that extend to holomorphic semigroups in some sector of the complex plane. These semigroups arise naturally when studying parabolic equations, and they have stronger regularising properties than general C_0-semigroups.

Definition 2.26. Let $S(t)$ be a C_0-semigroup on a Banach space $\mathcal{B}$. We say that S is **analytic of angle** $\theta \in (0, \frac{\pi}{2})$ if the following hold:

(i) **Analytic extension:** $z \mapsto S(z)$ is analytic on the sector

$$\Delta_\theta = \{z \in \mathbb{C} : |\arg z| < \theta\}.$$

(ii) **Semigroup law:** $S(z_1 + z_2) = S(z_1)S(z_2)$ for all $z_1, z_2 \in \Delta_\theta$.

(iii) **Rays are C_0:** For each φ with $|\varphi| < \theta$,

$$S_\varphi(t) := S(e^{i\varphi}t), \qquad t > 0,$$

defines a C_0-semigroup on $\mathcal{B}$.

If θ is the largest angle for which these properties hold, we call θ the **angle of analyticity** of S.

The next theorem shows that we can find a uniform bound for analytic semigroups on smaller sectors.

Theorem 2.27 (Uniform bounds on rays)**.** *Let S be an analytic semigroup with angle θ. Then, for every $\theta' < \theta$, there exist constants $M(\theta'), a(\theta') \geqslant 0$ such that*

$$\|S_\varphi(t)\| \leqslant M(\theta')\, e^{a(\theta')t}, \quad \forall t \geqslant 0, |\varphi| \leqslant \theta'.$$

Proof. Fix $\theta' < \theta$. For each $t \geqslant 0$ and $|\varphi| \leqslant \theta'$, one can write

$$te^{i\varphi} = t_+ e^{i\theta'} + t_- e^{-i\theta'},$$

where

$$t_\pm = \frac{t}{2}\left(\frac{\cos\varphi}{\cos\theta'} \pm \frac{\sin\varphi}{\sin\theta'}\right).$$

There exists a constant $C > 0$ such that

$$t_\pm \leqslant \frac{t}{2}\left(\frac{1}{\cos\theta'} + 1\right) \leqslant Ct.$$

Now, since along the boundary rays the semigroups $S_{\theta'}$ and $S_{-\theta'}$ are C_0-semigroups, there exist constants $M_\pm, a_\pm \geqslant 0$ such that

$$\|S_{\pm\theta'}(t)\| \leqslant M_\pm e^{a_\pm t}, \quad t \geqslant 0.$$

Thus, by the semigroup property,

$$\begin{aligned}
\|S_\varphi(t)\| &= \|S_{\theta'}(t_+)\, S_{-\theta'}(t_-)\| \\
&\leqslant \|S_{\theta'}(t_+)\|\, \|S_{-\theta'}(t_-)\| \\
&\leqslant M_+ M_- e^{a_+ t_+ + a_- t_-} \\
&\leqslant M_+ M_- e^{C(a_+ + a_-)t} \\
&= M(\theta')\, e^{a(\theta')t}.
\end{aligned}$$

This proves the claim. □

In fact, we can express the generator of analytic semigroups on each ray in terms of the generator on the positive real axis.

The following theorem compute the generator of $S_\varphi(t)$ in terms of the generator of $S_0(t)$.

Theorem 2.28. *Let S be an analytic semigroup with angle θ and generator L. Then, for $|\varphi| < \theta$, the generator L_φ of S_φ satisfies*

$$L_\varphi = e^{i\varphi} L.$$

Proof. In the following, we write $R_\lambda^\varphi = (\lambda - L_\varphi)^{-1}$. For $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda)$ sufficiently large, we have

$$R_\lambda x = \int_0^\infty e^{-\lambda t} S(t) x dt.$$

Fix $|\varphi| < \theta$. Since the map $z \to e^{-\lambda z} S(z) x$ is analytic on Δ_θ, we then integrate it on the closed contour $\gamma = \gamma_1 + \gamma_2 + \gamma_3$, where

$$\begin{aligned}
&\gamma_1 : [0, s] \to \mathbb{C}, \gamma_1(t) = t \\
&\gamma_2 : [0, \varphi] \to \mathbb{C}, \gamma_2(t) = se^{it} \\
&\gamma_3 : [0, s] \to \mathbb{C}, \gamma_3(t) = e^{i\varphi}(s - t)
\end{aligned}$$

By Cauchy theorem, we have $f_\gamma f(z) dz = 0$ and since, provided again that $\operatorname{Re}\lambda$ is large enough, $f(z)$ decays exponentially to 0 on γ_2 as $s \to \infty$, we can deform the contour of integration to obtain

$$R_\lambda x = e^{i\varphi} \int_0^\infty e^{-\lambda e^{i\varphi t}} S\left(e^{i\varphi} t\right) x dt$$

This shows that

$$R_\lambda = e^{i\varphi} R^\varphi_{\lambda e^{i\varphi}}.$$

This is equivalent to

$$(\lambda - L)^{-1} = e^{i\varphi} \left(\lambda e^{i\varphi} - L_\varphi\right)^{-1} = \left(\lambda - e^{-i\varphi} L_\varphi\right),$$

thus showing that $L_\varphi = e^{i\varphi} L$ as stated. □

Similar to the case of general C_0-semigroups, one can characterise the generators of analytic semigroups via resolvent bounds.

Theorem 2.29 (Hille-Yosida for analytic semigroups)**.** *A closed densely defined operator L on $\mathcal{B}$ is the generator of an analytic semigroup if and only if there exists $\theta \in \left(0, \frac{\pi}{2}\right)$ and $a \geqslant 0$ such that the resolvent set of L contains the sector*

$$V_{\theta,a} = \{a + re^{i\varphi} : r > 0,\ \varphi \in \left(-\frac{\pi}{2} - \theta,\ \frac{\pi}{2} + \theta\right)\}$$

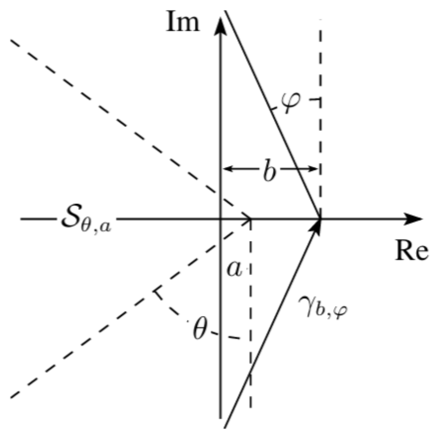
and there exists $M > 0$ such that for every $\lambda \in V_{\theta,a}$, the resolvent R_λ satisfies the bound

$$\|R_\lambda\| \leqslant M d\left(\lambda, V^c_{\theta,a}\right)^{-1}.$$

Proof. The forward direction follows directly from Hille-Yosida of strongly continuous group and Theorem (2.28).

Conversely, We now assume the sectorial resolvent condition and construct an analytic semigroup by a contour integral. Pick $\varphi \in (0, \theta)$ and $b > a$. Let $\gamma_{\varphi,b}$ be the curve oriented counterclockwise along the boundary of the sector $S_{\varphi,b}$ (as in the figure). For every w with $|\arg(w)| < \varphi$, define

$$S(w) := \frac{1}{2\pi i} \int_{\gamma_{\varphi,b}} e^{wz} R_z \, dz.$$



By the assumed resolvent bound, for every $z \in \gamma_{\varphi,b}$,

$$\|R_z\| \leqslant \frac{M}{d(z, V_{\theta,a}^c)} \leqslant \frac{M}{b-a} =: M'.$$

Parametrise $\gamma_{\varphi,b} = \gamma_+ - \gamma_-$ by

$$\gamma_\pm : \; z = b + re^{\pm i(\frac{\pi}{2}+\varphi)}, \qquad r \in [0, \infty).$$

Since $|\arg(w)| < \varphi$, one has $\mathrm{Re}(w\, e^{\pm i(\frac{\pi}{2}+\varphi)}) < 0$, hence $\mathrm{Re}(wz) \to -\infty$ linearly in r along both rays. Therefore the integral defining $S(w)$ converges absolutely and defines a bounded operator. Moreover, the definition is independence of the choice of b and φ.

We now prove it is an analytic semigroup. Pick w_1, w_2 with $|\arg(w_i)| < \varphi$

and choose $b_2 > b_1 > a$. Then

$$\begin{aligned} S(w_1)S(w_2) &= \frac{1}{(2\pi i)^2} \int_{\gamma_{\varphi,b_2}} \int_{\gamma_{\varphi,b_1}} e^{w_1 z + w_2 z'} R_z R_{z'}\, dz\, dz' \\ &= \frac{1}{(2\pi i)^2} \int_{\gamma_{\varphi,b_2}} \int_{\gamma_{\varphi,b_1}} e^{w_1 z + w_2 z'} \frac{R_z - R_{z'}}{z' - z}\, dz\, dz', \end{aligned}$$

where we used the resolvent identity $R_z R_{z'} = \frac{R_z - R_{z'}}{z'-z}$. Splitting the integral, we obtain

$$\begin{aligned} S(w_1)S(w_2) = &\frac{1}{(2\pi i)^2} \int_{\gamma_{\varphi,b_1}} e^{w_1 z} R_z \left(\int_{\gamma_{\varphi,b_2}} \frac{e^{w_2 z'}}{z' - z}\, dz' \right) dz \\ &- \frac{1}{(2\pi i)^2} \int_{\gamma_{\varphi,b_2}} e^{w_2 z'} R_{z'} \left(\int_{\gamma_{\varphi,b_1}} \frac{e^{w_1 z}}{z' - z}\, dz \right) dz'. \end{aligned}$$

For fixed $z \in \gamma_{\varphi,b_1}$, we can close γ_{φ,b_2} to the right; since z lies strictly to the left of γ_{φ,b_2}, the integrand has no pole inside and the inner integral equals 0. For fixed $z' \in \gamma_{\varphi,b_2}$, closing γ_{φ,b_1} to the right encloses the pole at $z = z'$, hence by Cauchy's integral formula,

$$\int_{\gamma_{\varphi,b_1}} \frac{e^{w_1 z}}{z' - z}\, dz = 2\pi i\, e^{w_1 z'}.$$

Plugging these into the previous display yields

$$S(w_1)S(w_2) = \frac{1}{2\pi i} \int_{\gamma_{\varphi,b_2}} e^{(w_1+w_2)z'} R_{z'}\, dz' = S(w_1 + w_2).$$

To prove the strong continuity on each ray, we fix $|\varphi| < \theta$ and pick $b > a$ and $\varphi' > |\varphi|$. For $t \geqslant 0$ we have

$$S_\varphi(t)x = S(e^{i\varphi}t)x = \frac{1}{2\pi i} \int_{\gamma_{\varphi',b}} e^{tz} R_z x\, dz.$$

Since the integrand is dominated by an integrable function independent of t in a neighbourhood of t_0. Hence, by dominated convergence,

$$\lim_{t\to t_0} S_\varphi(t)x = \frac{1}{2\pi i} \int_{\gamma_{\varphi',b}} \lim_{t\to t_0} e^{tz} R_z x\, dz = S_\varphi(t_0)x.$$

Thus S_φ is strongly continuous.

Lastly, we show the generator of S is L. Let $\widehat{L}$ be the generator of the constructed semigroup and denote its resolvent by $\widehat{R}_\lambda$. We show that $\widehat{R}_\lambda = R_\lambda$ for $\operatorname{Re}\lambda > a$. Fix such a λ and choose $b < \operatorname{Re}\lambda$. Then $\operatorname{Re}(z-\lambda) < 0$ for all $z \in \gamma_{\varphi,b}$, and we compute (for $x \in \mathcal{B}$)

$$\begin{aligned}\widehat{R}_\lambda x &= \int_0^\infty e^{-\lambda t} S(t)x\,dt = \frac{1}{2\pi i}\int_0^\infty \int_{\gamma_{\varphi,b}} e^{t(z-\lambda)} R_z x\,dz\,dt \\ &= \frac{1}{2\pi i}\int_{\gamma_{\varphi,b}} \left(\int_0^\infty e^{t(z-\lambda)}\,dt\right) R_z x\,dz = \frac{1}{2\pi i}\int_{\gamma_{\varphi,b}} \frac{R_z x}{\lambda - z}\,dz,\end{aligned}$$

where we used Fubini's theorem (justified by the exponential decay in t and the resolvent bound on $\gamma_{\varphi,b}$). Moreover, the integrand satisfies a decay estimate of the form

$$\left\|\frac{R_z}{\lambda - z}\right\| \lesssim \frac{1}{|z|}\cdot\frac{1}{|\lambda - z|} \sim \frac{1}{|z|^2} \quad \text{as } |z| \to \infty \text{ along the contour,}$$

so we may add a large semicircle on the right and apply the residue theorem to enclose the simple pole at $z = \lambda$. This yields

$$\widehat{R}_\lambda x = R_\lambda x.$$

Therefore $\widehat{L} = L$, and the proof is complete. □

Using the Hille-Yosida theorem for analytic semigroups, we can now test whether specific differential operators generate analytic semigroups. The first example is the translation operator, which does not generate an analytic semigroup, while the second example is the heat operator, which does generate an analytic semigroup.

Example 2.30. Let $L = \frac{d}{dx}$ on $L^2(\mathbf{R})$ with domain $\mathcal{D}(L) = H^1(\mathbf{R})$. Then L cannot be the generator of an analytic semigroup.

We claim that $\sigma(L) = i\mathbf{R}$. To see this, consider the operator

$$L_0 = -i\frac{d}{dx} \quad \text{on } L^2(\mathbf{R}), \qquad \mathcal{D}(L_0) = H^1(\mathbf{R}).$$

We first show that L_0 is self-adjoint. For $f, g \in C_c^\infty(\mathbf{R})$, integration by parts yields

$$\begin{aligned}\langle L_0 f,\, g\rangle &= \int_{\mathbb{R}} -if'(x)\,\overline{g(x)}\,dx \\ &= \int_{\mathbb{R}} f(x)\,\overline{-ig'(x)}\,dx = \langle f,\, L_0 g\rangle\,.\end{aligned}$$

Since $C_c^\infty(\mathbf{R})$ is dense in $H^1(\mathbf{R})$, the identity extends to all $f, g \in H^1(\mathbf{R})$, so L_0 is symmetric. In particular, $\sigma(L_0) \subset \mathbf{R}$.

Next, fix $\lambda \in \mathbf{R}$ and we show that $\lambda \in \sigma(L_0)$. Arguing by contradiction, assume that $(\lambda I - L_0)^{-1}$ exists and is bounded. Pick a nonzero $f \in C_c^\infty(\mathbf{R})$ and, for each $k \in \mathbb{N}$, define

$$g_k(x) = \frac{1}{\sqrt{k}} e^{i\lambda x}\, f\big(k^{-1}x\big)\,.$$

Then $g_k \in H^1(\mathbf{R})$ and $\|g_k\| = \|f\|$ for all k, since

$$\begin{aligned}\|g_k\|^2 &= \int_{\mathbb{R}} |g_k(x)|^2\,dx = \int_{\mathbb{R}} \frac{1}{k}\left|f\big(k^{-1}x\big)\right|^2\,dx \\ &= \int_{\mathbb{R}} |f(y)|^2\,dy = \|f\|^2.\end{aligned}$$

Moreover,

$$(L_0 - \lambda I)g_k(x) = -\frac{i}{k^{3/2}} e^{i\lambda x} f'\big(k^{-1}x\big)\,,$$

so $\|(L_0 - \lambda I)g_k\| = k^{-1}\|f'\|$. Hence,

$$\begin{aligned}\|f\| = \|g_k\| &= \big\|(L_0 - \lambda I)^{-1}(L_0 - \lambda I)g_k\big\| \\ &\leqslant \|(L_0 - \lambda I)^{-1}\|\,\|(L_0 - \lambda I)g_k\| = k^{-1}\|(L_0 - \lambda I)^{-1}\|\,\|f'\|.\end{aligned}$$

Letting $k \to \infty$ forces $\|f\| = 0$, a contradiction. Therefore $\lambda \in \sigma(L_0)$ for every $\lambda \in \mathbf{R}$, and thus $\sigma(L_0) = \mathbf{R}$. Since $L = iL_0$, we obtain $\sigma(L) = i\mathbf{R}$.

Since $\sigma(L) = i\mathbf{R}$, the resolvent set $\rho(L)$ cannot contain any sector

$$V_{a,\theta} := \{a + re^{i\varphi} :\ r > 0,\ |\varphi| < \tfrac{\pi}{2} + \theta\}, \qquad \theta \in \Big(0, \frac{\pi}{2}\Big),$$

based at some $a \geqslant 0$. By the analytic Hille-Yosida theorem, L cannot be the generator of an analytic semigroup.

Example 2.31. Let $L = \frac{d^2}{dx^2}$ on $L^2(\mathbf{R})$ with domain $\mathcal{D}(L) = H^2(\mathbf{R})$. Then L is the generator of an analytic semigroup.

We first note that for $f, g \in C_c^\infty(\mathbf{R})$, integration by parts gives

$$\begin{aligned}\langle Lf,\, g\rangle &= \int_{\mathbb{R}} f''(x)\,\overline{g(x)}\,dx \\ &= -\int_{\mathbb{R}} f'(x)\,\overline{g'(x)}\,dx = \int_{\mathbb{R}} f(x)\,\overline{g''(x)}\,dx = \langle f,\, Lg\rangle\,.\end{aligned}$$

Since $C_c^\infty(\mathbf{R})$ is dense in $H^2(\mathbf{R})$, the identity extends to all $f, g \in H^2(\mathbf{R})$, so L is self-adjoint. Moreover, L is negative definite: for every $f \in H^2(\mathbf{R})$,

$$\langle Lf,\, f\rangle = \int_{\mathbb{R}} f''(x)\,\overline{f(x)}\,dx = -\int_{\mathbb{R}} |f'(x)|^2\,dx \leqslant 0.$$

Hence $\sigma(L) \subset (-\infty, 0]$, so the resolvent set $\rho(L)$ contains $\mathbb{C} \setminus (-\infty, 0]$, in particular the sector $V_{0,\pi/2}$.

Finally, for every $\lambda \in \mathbb{C} \setminus (-\infty, 0]$, we claim that

$$\|R_\lambda\| \leqslant \frac{1}{d(\lambda, (-\infty, 0])}.$$

Fix $u \in \mathcal{D}(L)$ with $u \neq 0$. Then

$$\begin{aligned}\big|\langle(\lambda I - L)u,\, u\rangle\big| &= \big|\langle\lambda u,\, u\rangle - \langle Lu,\, u\rangle\big| = \big|\lambda\|u\|^2 + \|u'\|^2\big| \\ &= \|u\|^2\,|\lambda + a|,\end{aligned}$$

where $a := \|u'\|^2/\|u\|^2 \geqslant 0$. By Cauchy-Schwarz,

$$\begin{aligned}\|(\lambda I - L)u\|\,\|u\| &\geqslant \big|\langle(\lambda I - L)u,\, u\rangle\big| = \|u\|^2\,|\lambda + a| \\ &\geqslant \|u\|^2 \inf_{b \geqslant 0} |\lambda + b| = \|u\|^2\, d(\lambda, (-\infty, 0]).\end{aligned}$$

Dividing by $\|u\|$ yields

$$\|(\lambda I - L)u\| \geqslant d(\lambda, (-\infty, 0])\,\|u\|,$$

and taking the supremum over $u \neq 0$ shows

$$\|R_\lambda\| \leqslant \frac{1}{d(\lambda, (-\infty, 0])}.$$

Therefore, by the analytic Hille-Yosida theorem, L generates an analytic semigroup (the heat semigroup).

The next theorem shows that perturbing the generator of an analytic semigroup by a sufficiently small operator still yields the generator of an analytic semigroup.

Theorem 2.32 (Perturbation of analytic semigroup generators)**.** *Let L_0 be the generator of an analytic semigroup, and let $P : \mathcal{D}(P) \to \mathcal{B}$ be a linear operator (the perturbation) such that:*

(i) $\mathcal{D}(P)$ contains $\mathcal{D}(L_0)$.

(ii) For every $\varepsilon > 0$ there exists $C_\varepsilon > 0$ such that

$$\|Px\| \leqslant \varepsilon \|L_0 x\| + C_\varepsilon \|x\|, \qquad \forall\, x \in \mathcal{D}(L_0).$$

Then the operator $L := L_0 + P$ with domain $\mathcal{D}(L) = \mathcal{D}(L_0)$ is also the generator of an analytic semigroup.

Proof. By the Hille-Yosida theorem for analytic semigroups, there exist $a_0 \geqslant 0$ and $\theta_0 \in \left(0, \frac{\pi}{2}\right)$ such that the resolvent set $\rho(L_0)$ contains the sector V_{a_0,θ_0}. Moreover, there exists $M > 0$ such that for every $\lambda \in V_{a_0,\theta_0}$,

$$\|R_\lambda^0\| \leqslant \frac{M}{d(\lambda, V_{\theta_0,a_0}^c)},$$

where $R_\lambda^0 := (\lambda I - L_0)^{-1}$.

Fix $\lambda \in V_{a_0,\theta_0}$. Since R_λ^0 is bounded and maps $\mathcal{B}$ into $\mathcal{D}(L_0)$, for any $y \in \mathcal{B}$ we look for $x \in \mathcal{D}(L_0)$ solving

$$(\lambda I - L)x = y.$$

Writing $x = R_\lambda^0 z$ for some $z \in \mathcal{B}$, we obtain

$$\begin{aligned}
(\lambda I - L)R_\lambda^0 z &= y,\\
(\lambda I - L_0 - P)R_\lambda^0 z &= y,\\
z - PR_\lambda^0 z &= y,\\
(I - PR_\lambda^0)z &= y.
\end{aligned}$$

We claim that there exist $a \geqslant a_0$, $\theta \in (0, \theta_0)$, and a constant $c \in [0, 1)$ such that for every $\lambda \in V_{a,\theta}$,

$$\|PR_\lambda^0\| \leqslant c < 1.$$

Assume the claim for the moment. Then for $\lambda \in V_{a,\theta}$, the operator $I - PR_\lambda^0$ is invertible and

$$(I - PR_\lambda^0)^{-1} = \sum_{k=0}^{\infty} (PR_\lambda^0)^k, \qquad \|(I - PR_\lambda^0)^{-1}\| \leqslant \frac{1}{1-c}.$$

Thus $z = (I - PR_\lambda^0)^{-1} y$ and $x = R_\lambda^0 z$ solve $(\lambda I - L)x = y$, so $\lambda \in \rho(L)$ and

$$\begin{aligned} \|R_\lambda y\| &= \|x\| = \|R_\lambda^0 (I - PR_\lambda^0)^{-1} y\| \\ &\leqslant \frac{\|R_\lambda^0\|}{1-c} \|y\| \leqslant \frac{M'}{d(\lambda, V_{\theta_0,a_0}^c)} \|y\| \leqslant \frac{M''}{d(\lambda, V_{\theta,a}^c)} \|y\|. \end{aligned}$$

Therefore,

$$\|R_\lambda\| \leqslant \frac{M''}{d(\lambda, V_{\theta,a}^c)},$$

and by the Hille-Yosida theorem for analytic semigroups, L generates an analytic semigroup.

It remains to prove the claim. By the assumption on P, for every $\varepsilon > 0$ and every $x \in \mathcal{D}(L_0)$,

$$\|Px\| \leqslant \varepsilon \|L_0 x\| + C_\varepsilon \|x\|.$$

Apply this with $x = R_\lambda^0 y$ and take the operator norm to obtain

$$\|PR_\lambda^0\| \leqslant \varepsilon \|L_0 R_\lambda^0\| + C_\varepsilon \|R_\lambda^0\|.$$

Using $L_0 R_\lambda^0 = \lambda R_\lambda^0 - I$, we get

$$\|L_0 R_\lambda^0\| \leqslant |\lambda| \, \|R_\lambda^0\| + 1,$$

hence

$$\|PR_\lambda^0\| \leqslant \varepsilon\big(|\lambda| \|R_\lambda^0\| + 1\big) + C_\varepsilon \|R_\lambda^0\| \leqslant \varepsilon + \frac{M(\varepsilon|\lambda| + C_\varepsilon)}{d(\lambda, V_{\theta_0,a_0}^c)}.$$

Now choose $\theta \in (0, \theta_0)$ and then a large enough so that there exists a constant $C' > 0$ with

$$d(\lambda, V_{\theta_0,a_0}^c) \geqslant C'|\lambda| \qquad \text{for all } \lambda \in V_{a,\theta}.$$

Fix $\varepsilon > 0$ small enough so that $\varepsilon < \frac{1}{4}$ and $\frac{\varepsilon}{C'} < \frac{1}{4M}$, and then choose a large enough so that

$$d(\lambda, V_{\theta_0,a_0}^c) > 4MC_\varepsilon \qquad \text{for all } \lambda \in V_{a,\theta}.$$

With these choices, for every $\lambda \in V_{a,\theta}$,

$$\|PR_\lambda^0\| \leqslant \varepsilon + \frac{M\varepsilon|\lambda|}{d(\lambda, V_{\theta_0,a_0}^c)} + \frac{MC_\varepsilon}{d(\lambda, V_{\theta_0,a_0}^c)} \leqslant \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4} < 1.$$

This proves the claim. □

As a consequence of the perturbation theorem, we obtain the following example.

Theorem 2.33. *Let $f \in L^\infty(\mathbf{R})$. Then the operator*

$$(Lg)(x) = g''(x) + f(x)g'(x)$$

on $L^2(\mathbf{R})$, with domain $\mathcal{D}(L) = H^2(\mathbf{R})$, is the generator of an analytic semigroup.

Proof. Let $L_0 = \frac{d^2}{dx^2}$, which is the generator of an analytic semigroup. Define $Pg := fg'$ for $g \in H^2(\mathbf{R})$, so that $L = L_0 + P$.

For $g \in H^2(\mathbf{R})$,

$$\begin{aligned}\|Pg\|^2 &= \int_{\mathbf{R}} |f(x)|^2 \, |g'(x)|^2 \, dx \leqslant \|f\|_{L^\infty}^2 \int_{\mathbf{R}} |g'(x)|^2 \, dx \\ &= -\|f\|_{L^\infty}^2 \, \langle g, g'' \rangle \leqslant \|f\|_{L^\infty}^2 \, \|g\| \, \|g''\| = \|f\|_{L^\infty}^2 \, \|g\| \, \|L_0 g\|.\end{aligned}$$

Using $2xy \leqslant \varepsilon x^2 + y^2/\varepsilon$ for $\varepsilon > 0$, we obtain

$$\|Pg\| \leqslant \varepsilon\|L_0 g\| + \frac{\|f\|_{L^\infty}^2}{4\varepsilon}\|g\|.$$

Thus P satisfies the assumptions of the perturbation theorem, and therefore L is also the generator of an analytic semigroup. □

2.4 Interpolation Spaces

Let L be the generator of an analytic semigroup $\{S(t)\}_{t \geqslant 0}$. In this section, we assume that there exist constants $M \geqslant 1$ and $a > 0$ such that

$$\|S(t)\| \leqslant Me^{-at}, \qquad t \geqslant 0.$$

Under this assumption, the Hille-Yosida theorem implies that $0 \in \rho(L)$. In particular, L^{-1} is well-defined and bounded.

Definition 2.34. For $\alpha > 0$, we define the **negative fractional power of** $-L$ by

$$(-L)^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} S(t)\, dt,$$

where the integral is a Bochner integral in $\mathcal{B}$.

By the assumed decay of $S(t)$, the integral converges and defines a bounded operator for every $\alpha > 0$. Indeed,

$$\begin{aligned}
\|(-L)^{-\alpha}\| &\leqslant \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} \|S(t)\|\, dt \\
&\leqslant \frac{M}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-at}\, dt \\
&= \frac{M}{\Gamma(\alpha) a^\alpha} \int_0^\infty u^{\alpha-1} e^{-u}\, du \\
&= \frac{M}{a^\alpha}.
\end{aligned}$$

The next theorem shows that these operators satisfy the usual algebraic rule for exponents.

Theorem 2.35. *For every* $\alpha, \beta > 0$,

$$(-L)^{-\alpha}(-L)^{-\beta} = (-L)^{-(\alpha+\beta)}.$$

Proof. By definition,

$$\begin{aligned}
(-L)^{-\alpha}(-L)^{-\beta} &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty \int_0^\infty t^{\alpha-1} s^{\beta-1} S(t) S(s)\, dt\, ds \\
&= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty \int_0^\infty t^{\alpha-1} s^{\beta-1} S(t+s)\, dt\, ds,
\end{aligned}$$

where we used the semigroup property. Setting $u = t + s$ and integrating first over $t \in (0,u)$ gives

$$\begin{aligned}
(-L)^{-\alpha}(-L)^{-\beta} &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty \left(\int_0^u t^{\alpha-1}(u-t)^{\beta-1}\,dt \right) S(u)\,du \\
&= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty u^{\alpha+\beta-1} \left(\int_0^1 r^{\alpha-1}(1-r)^{\beta-1}\,dr \right) S(u)\,du \\
&= \frac{B(\alpha,\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty u^{\alpha+\beta-1} S(u)\,du \\
&= \frac{1}{\Gamma(\alpha+\beta)} \int_0^\infty u^{\alpha+\beta-1} S(u)\,du = (-L)^{-(\alpha+\beta)},
\end{aligned}$$

since $B(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$. $\square$

Corollary 2.36. *For every $\alpha > 0$, the operator $(-L)^{-\alpha}$ is injective.*

Proof. Fix $\alpha > 0$ and choose an integer $n > \alpha$. By Theorem 2.35,

$$(-L)^{-n} = (-L)^{-(n-\alpha)}(-L)^{-\alpha}.$$

If $(-L)^{-\alpha}x = 0$, then $(-L)^{-n}x = 0$. On the other hand, since $0 \in \rho(L)$, we have $-L$ invertible and therefore $(-L)^{-1}$ is injective, which implies $(-L)^{-n}$ is injective. Hence $x = 0$, proving that $(-L)^{-\alpha}$ is injective. $\square$

Using injectivity, we define positive fractional powers by inversion:

$$(-L)^{\alpha} := \big((-L)^{-\alpha}\big)^{-1}, \qquad \mathcal{D}((-L)^{\alpha}) = \mathrm{Range}((-L)^{-\alpha}).$$

Remark. This definition agrees with the usual integer powers. For instance, we have $(-L)^1 = -L$. Indeed, since S decays exponentially, the Laplace formula at $\lambda = 0$ yields

$$(-L)^{-1} = \int_0^\infty S(t)\,dt = R_0,$$

so taking inverses gives $(-L)^1 = -L$.

We can now define the interpolation spaces associated with L.

Definition 2.37. For $\alpha > 0$, we define the **interpolation space** $\mathcal{B}_\alpha := \mathcal{D}((-L)^{\alpha})$ equipped with the norm

$$\|x\|_\alpha := \|(-L)^{\alpha}x\|_{\mathcal{B}}.$$

Similarly, define $\mathcal{B}_{-\alpha}$ as the completion of $\mathcal{B}$ with respect to the norm

$$\|x\|_{-\alpha} := \|(-L)^{-\alpha} x\|_{\mathcal{B}}.$$

The next proposition records a basic monotonicity property of these spaces.

Theorem 2.38. *If $\alpha \geqslant \beta$, then*

$$\mathcal{B}_\alpha \subset \mathcal{B}_\beta,$$

regardless of the signs of α and β.

Proof. Let us proof it by 3 cases. The first case is when $\alpha \geqslant \beta \geqslant 0$. In this case,

$$(-L)^{-\alpha} = (-L)^{-\beta}(-L)^{\beta-\alpha},$$

so

$$\mathrm{Range}((-L)^{-\alpha}) \subset \mathrm{Range}((-L)^{-\beta}),$$

which is exactly $\mathcal{B}_\alpha \subset \mathcal{B}_\beta$.

The second case is when $0 \geqslant \alpha \geqslant \beta$. Here,

$$(-L)^{\beta} = (-L)^{\beta-\alpha}(-L)^{\alpha}.$$

Since $(-L)^{\beta-\alpha}$ is bounded, there exists $C > 0$ such that

$$\|x\|_\beta = \|(-L)^\beta x\| \leqslant C\|(-L)^\alpha x\| = C\|x\|_\alpha.$$

Thus the identity map on $\mathcal{B}$ is continuous from $(\mathcal{B}, \|\cdot\|_\alpha)$ into $(\mathcal{B}, \|\cdot\|_\beta)$, and it extends by continuity to the completions. Hence $\mathcal{B}_\alpha \hookrightarrow \mathcal{B}_\beta$.

The last case is when $\alpha \geqslant 0 \geqslant \beta$. This follows from $\mathcal{B}_\alpha \subset \mathcal{B}_0 = \mathcal{B} \subset \mathcal{B}_\beta$. This completes the proof. $\square$

The next proposition gives another useful representation of $(-L)^\alpha$ for $\alpha \in (0,1)$. It has several important corollaries.

Theorem 2.39. *Let $\alpha \in (0,1)$ and $x \in \mathcal{D}(L)$. Then*

$$(-L)^\alpha x = \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty t^{\alpha-1} R_t(-L)x \, dt, \tag{2.4}$$

where $R_t = (tI - L)^{-1}$.

Proof. Denote the right-hand side by

$$A_\alpha x := \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty t^{\alpha-1} R_t(-L)x \, dt.$$

We show that $A_\alpha(-L)^{-\alpha}x = x$ for every $x \in \mathcal{D}(L)$.

Using the Laplace representation $R_t = \int_0^\infty e^{-tu} S(u) \, du$, we obtain

$$\begin{aligned} A_\alpha x &= \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty t^{\alpha-1} \left(\int_0^\infty e^{-tu} S(u) \, du \right) (-L)x \, dt \\ &= \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \left(\int_0^\infty t^{\alpha-1} e^{-tu} \, dt \right) S(u)(-L)x \, du \\ &= \frac{\sin(\pi\alpha)\Gamma(\alpha)}{\pi} \int_0^\infty u^{-\alpha} S(u)(-L)x \, du. \end{aligned}$$

Now plug in the definition

$$(-L)^{-\alpha}x = \frac{1}{\Gamma(\alpha)} \int_0^\infty v^{\alpha-1} S(v)x \, dv$$

to get

$$A_\alpha(-L)^{-\alpha}x = \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \int_0^\infty v^{\alpha-1} u^{-\alpha} S(u+v)(-L)x \, du \, dv.$$

Make the change of variables $r = u + v$ and $w = v/(u+v)$ (so $u = r(1-w)$ and $v = rw$). Then

$$\begin{aligned} A_\alpha(-L)^{-\alpha}x &= \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \int_0^1 (rw)^{\alpha-1} \left[r(1-w)\right]^{-\alpha} (-L)S(r)x \, r \, dw \, dr \\ &= \frac{\sin(\pi\alpha)}{\pi} \left(\int_0^1 w^{\alpha-1}(1-w)^{-\alpha} \, dw \right) \int_0^\infty (-L)S(r)x \, dr \\ &= \frac{\sin(\pi\alpha)}{\pi} B(\alpha, 1-\alpha) \int_0^\infty (-L)S(r)x \, dr \\ &= \frac{\sin(\pi\alpha)\Gamma(\alpha)\Gamma(1-\alpha)}{\pi} \int_0^\infty (-L)S(r)x \, dr. \end{aligned}$$

Using $\Gamma(\alpha)\Gamma(1-\alpha) = \pi / \sin(\pi\alpha)$, this becomes

$$A_\alpha(-L)^{-\alpha}x = \int_0^\infty (-L)S(r)x \, dr.$$

Finally, since $\partial_t S(t)x = LS(t)x$ for $x \in \mathcal{D}(L)$, the fundamental theorem of calculus yields

$$\int_0^T (-L)S(r)x\,dr = x - S(T)x.$$

Letting $T \to \infty$ and we then have $S(T)x \to 0$. This proves the identity. $\square$

The following corollary gives a useful bound on $(-L)^\alpha$.

Corollary 2.40. *For every $\alpha \in (0,1)$ there exists $C_\alpha > 0$ such that*

$$\|(-L)^\alpha x\| \leqslant C_\alpha \|Lx\|^\alpha \|x\|^{1-\alpha}, \qquad x \in \mathcal{D}(L).$$

Proof. Split the integral in (2.4) at $K > 0$:

$$(-L)^\alpha x = \frac{\sin(\pi\alpha)}{\pi} \int_0^K t^{\alpha-1} R_t(-L)x\,dt + \frac{\sin(\pi\alpha)}{\pi} \int_K^\infty t^{\alpha-1} R_t(-L)x\,dt.$$

We use the bound

$$\|R_t\| = \left\| \int_0^\infty e^{-ts} S(s)\,ds \right\| \leqslant \int_0^\infty e^{-ts} M e^{-as}\,ds \leqslant \frac{M}{t}, \qquad t > 0.$$

For the first integral, note that $R_t(-L) = tR_t - I$, so

$$\|R_t(-L)x\| \leqslant (1 + t\|R_t\|)\|x\| \leqslant (1+M)\|x\|.$$

Therefore,

$$\left\| \int_0^K t^{\alpha-1} R_t(-L)x\,dt \right\| \leqslant (1+M) \int_0^K t^{\alpha-1}\,dt\,\|x\| = \frac{1+M}{\alpha} K^\alpha \|x\|.$$

For the second integral, we simply use $\|R_t(-L)x\| \leqslant \|R_t\|\,\|Lx\| \leqslant (M/t)\|Lx\|$ to get

$$\left\| \int_K^\infty t^{\alpha-1} R_t(-L)x\,dt \right\| \leqslant M \int_K^\infty t^{\alpha-2}\,dt\,\|Lx\| = \frac{M}{1-\alpha} K^{\alpha-1} \|Lx\|.$$

Combining these estimates,

$$\|(-L)^\alpha x\| \leqslant C_\alpha \left(K^\alpha \|x\| + K^{\alpha-1} \|Lx\| \right),$$

for a constant C_α depending only on α and M. Choosing

$$K = \frac{1-\alpha}{\alpha}\frac{\|Lx\|}{\|x\|}$$

yields the desired bound. $\square$

The next theorem uses Corollary 2.40 to obtain a more concrete perturbation result.

Theorem 2.41. *Let L_0 be the generator of an analytic semigroup $\{S_0(t)\}_{t\geqslant 0}$. Let P be a linear operator on $\mathcal{B}$ such that there exists $\alpha \in [0,1)$ with*

(i) $\mathcal{B}_\alpha \subset \mathcal{D}(P)$,

(ii) P is bounded from $\mathcal{B}_\alpha$ into $\mathcal{B}$.

Then $L := L_0 + P$ is also the generator of an analytic semigroup $\{S(t)\}_{t\geqslant 0}$. Moreover, for every $x \in \mathcal{B}$,

$$S(t)x = S_0(t)x + \int_0^t S_0(t-s)\, P\, S(s)x\, ds.$$

Proof. By (ii), there exists $C > 0$ such that for every $x \in \mathcal{B}_\alpha$,

$$\|Px\| \leqslant C\|x\|_\alpha.$$

By Corollary 2.40 (applied to L_0), we also have

$$\|Px\| \leqslant C_\alpha \|L_0 x\|^\alpha \|x\|^{1-\alpha}, \qquad x \in \mathcal{D}(L_0).$$

Fix $\varepsilon > 0$. Young's inequality with $p = 1/\alpha$ and $q = 1/(1-\alpha)$ yields

$$\|L_0 x\|^\alpha \|x\|^{1-\alpha} \leqslant \alpha\varepsilon\|L_0 x\| + (1-\alpha)\varepsilon^{-\frac{\alpha}{1-\alpha}}\|x\|.$$

Therefore,

$$\|Px\| \leqslant C_\alpha\left(\alpha\varepsilon\|L_0 x\| + (1-\alpha)\varepsilon^{-\frac{\alpha}{1-\alpha}}\|x\|\right), \qquad x \in \mathcal{D}(L_0),$$

so P satisfies the assumptions of the perturbation theorem. Hence $L = L_0 + P$ generates an analytic semigroup $S(t)$.

Define

$$\Phi(t)x := S_0(t)x + \int_0^t S_0(t-s)\, P\, S(s)x\, ds.$$

We show that Φ has the same Laplace transform as S, which implies $\Phi(t) = S(t)$ by uniqueness of the resolvent.

Let $\operatorname{Re}\lambda$ be sufficiently large. Using Fubini and the semigroup property,

$$
\begin{aligned}
\int_0^\infty \int_0^t e^{-\lambda t} S_0(t-s) P S(s) x \, ds \, dt &= \int_0^\infty \int_s^\infty e^{-\lambda t} S_0(t-s) \, dt \, P S(s) x \, ds \\
&= \int_0^\infty \int_0^\infty e^{-\lambda(u+s)} S_0(u) \, du \, P S(s) x \, ds \\
&= R_\lambda^0 \int_0^\infty e^{-\lambda s} P S(s) x \, ds = R_\lambda^0 P R_\lambda x,
\end{aligned}
$$

where $R_\lambda^0 = (\lambda I - L_0)^{-1}$ and $R_\lambda = (\lambda I - L)^{-1}$. Combining this with $\int_0^\infty e^{-\lambda t} S_0(t) x \, dt = R_\lambda^0 x$, we obtain

$$
\begin{aligned}
\int_0^\infty e^{-\lambda t} \Phi(t) x \, dt &= R_\lambda^0 x + R_\lambda^0 P R_\lambda x \\
&= R_\lambda^0 \big((\lambda I - L_0) R_\lambda + P R_\lambda \big) x \\
&= R_\lambda^0 (\lambda I - L_0 - P) R_\lambda x \\
&= R_\lambda^0 (\lambda I - L) R_\lambda x \\
&= R_\lambda x.
\end{aligned}
$$

Thus the Laplace transform of Φ equals the resolvent of L, so $\Phi(t) = S(t)$. $\square$

In Theorem 2.8, we proved that the semigroup $\{S(t)\}_{t \geqslant 0}$ leaves the domain of L invariant and that

$$
L S(t) = S(t) L, \qquad t \geqslant 0.
$$

In fact, this commutativity extends to the fractional powers of $-L$. We now show that $S(t)$ leaves the interpolation spaces $\mathcal{B}_\alpha = \mathcal{D}((-L)^\alpha)$ invariant for every $\alpha > 0$.

Theorem 2.42. *For every $t > 0$ and every $\alpha \in \mathbb{R}$, one has*

$$
(-L)^\alpha S(t) = S(t)(-L)^\alpha.
$$

In particular, $S(t)\mathcal{B}_\alpha \subset \mathcal{B}_\alpha$ for every $\alpha > 0$.

Proof. For λ with $\operatorname{Re}\lambda$ sufficiently large, we have the Laplace representation

$$R_\lambda = \int_0^\infty e^{-\lambda s} S(s)\, ds,$$

so $S(t)R_\lambda = R_\lambda S(t)$ for every $t \geqslant 0$.

Assume first that $0 < \alpha < 1$. Using the representation proved earlier,

$$(-L)^\alpha x = \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty t^{\alpha-1} R_t(-L)x\, dt,$$

and the commutation $S(t)R_t = R_t S(t)$, we obtain

$$S(t)(-L)^\alpha x = (-L)^\alpha S(t)x.$$

If $\alpha = n \geqslant 0$ is an integer, then $S(t)\mathcal{D}(L) \subset \mathcal{D}(L)$ and $LS(t) = S(t)L$, hence

$$(-L)^n S(t)x = S(t)(-L)^n x.$$

If $\alpha = -n$ with $n \in \mathbb{N}$, then by the definition of negative powers,

$$\begin{aligned}
S(t)(-L)^{-n}x &= \frac{1}{\Gamma(n)}\, S(t) \int_0^\infty s^{n-1} S(s)x\, ds \\
&= \frac{1}{\Gamma(n)} \int_0^\infty s^{n-1} S(t+s)x\, ds \\
&= \frac{1}{\Gamma(n)} \int_0^\infty s^{n-1} S(s)S(t)x\, ds \\
&= (-L)^{-n} S(t)x.
\end{aligned}$$

For general $\alpha \in \mathbb{R}$, write $\alpha = m + \beta$ with $m \in \mathbb{Z}$ and $\beta \in (0,1)$. Then

$$(-L)^\alpha S(t)x = (-L)^m(-L)^\beta S(t)x = (-L)^m S(t)(-L)^\beta x = S(t)(-L)^\alpha x.$$

Thus commutativity holds for all $\alpha \in \mathbb{R}$.

Finally, for $\alpha > 0$ we have $\mathcal{B}_\alpha = \operatorname{Ran}((-L)^{-\alpha})$. The identity

$$S(t)(-L)^{-\alpha}x = (-L)^{-\alpha} S(t)x$$

shows that $S(t)\mathcal{B}_\alpha \subset \mathcal{B}_\alpha$. $\square$

One of the most important properties of analytic semigroups is their smoothing effect.

Theorem 2.43. *For every $\alpha > 0$, the operator $S(t)$ maps $\mathcal{B}$ into $\mathcal{B}_\alpha$. Moreover, there exists a constant C_α such that*

$$\|(-L)^\alpha S(t)\| \leqslant \frac{C_\alpha}{t^\alpha}, \qquad t \in (0,1].$$

Proof. We first consider the case $\alpha = k \in \mathbb{N}$. From the contour representation of analytic semigroups, for every $\varphi \in (0,\theta)$ and $b > 0$,

$$S(t)x = \frac{1}{2\pi i} \int_{\gamma_{\varphi,b}} e^{tz} R_z x \, dz.$$

Applying L and using $LR_z = (zR_z - I)$ gives

$$\begin{aligned} LS(t)x &= \frac{1}{2\pi i} \int_{\gamma_{\varphi,b}} e^{tz} LR_z x \, dz \\ &= \frac{1}{2\pi i} \int_{\gamma_{\varphi,b}} e^{tz} (zR_z - I) x \, dz \\ &= \frac{1}{2\pi i} \int_{\gamma_{\varphi,b}} z e^{tz} R_z x \, dz - \frac{1}{2\pi i} \int_{\gamma_{\varphi,b}} e^{tz} x \, dz. \end{aligned}$$

Closing the contour to the left half-plane shows that the second integral vanishes, hence

$$LS(t)x = \frac{1}{2\pi i} \int_{\gamma_{\varphi,b}} z e^{tz} R_z x \, dz.$$

Iterating k times yields

$$L^k S(t)x = \frac{1}{2\pi i} \int_{\gamma_{\varphi,b}} z^k e^{tz} R_z x \, dz.$$

Parametrize $\gamma_{\varphi,b} = \gamma_+ - \gamma_-$ by

$$\gamma_\pm(r) = b + r e^{\pm i(\pi/2 + \varphi)}, \qquad r \in [0,\infty).$$

Then there exist constants $c_i > 0$ such that

$$\begin{aligned}\|L^k S(t)x\| &\leqslant c_1 \int_0^\infty |\gamma_+(r)|^k e^{t\,\mathrm{Re}(\gamma_+(r))} \|R_{\gamma_+(r)}\| \, \|x\| \, dr \\ &\leqslant c_2 \int_0^\infty (1+r)^k e^{-c_3 tr} \frac{1}{1+r} \, dr \, \|x\| \\ &= c_2 \int_0^\infty (1+r)^{k-1} e^{-c_3 tr} \, dr \, \|x\|.\end{aligned}$$

Integrating by parts $k-1$ times gives

$$\|L^k S(t)x\| \leqslant \frac{c_4}{t^k} \|x\|,$$

as required.

Now let $\alpha > 0$ be non-integer and write

$$(-L)^\alpha = (-L)^{\alpha - m}(-L)^m, \qquad m := \lfloor \alpha \rfloor + 1,$$

so that $\alpha - m \in (-1, 0)$. Set $\alpha' := \alpha - m \in (-1, 0)$. Using the definition of negative powers and the semigroup property,

$$\begin{aligned}\|(-L)^\alpha S(t)x\| &= \|(-L)^{\alpha'}(-L)^m S(t)x\| \\ &= \left\| \frac{1}{\Gamma(-\alpha')} \int_0^\infty s^{-\alpha'-1} S(s)(-L)^m S(t)x \, ds \right\| \\ &= \left\| \frac{1}{\Gamma(-\alpha')} \int_0^\infty s^{-\alpha'-1} (-L)^m S(t+s)x \, ds \right\| \\ &\leqslant C \int_0^\infty s^{-\alpha'-1} \frac{1}{(t+s)^m} \, ds \, \|x\|.\end{aligned}$$

With the change of variables $s = ut$,

$$\|(-L)^\alpha S(t)x\| \leqslant C t^{-\alpha} \left(\int_0^\infty \frac{u^{-\alpha'-1}}{(1+u)^m} \, du \right) \|x\| \leqslant C_\alpha t^{-\alpha} \|x\|.$$

This completes the proof. □

We now record some useful corollaries.

Corollary 2.44. *For every $\alpha, \beta \in \mathbb{R}$ and every $t > 0$, the operator $S(t)$ maps $\mathcal{B}_\alpha$ into $\mathcal{B}_\beta$. Moreover, if $\beta > \alpha$, then there exists $C_{\alpha,\beta} > 0$ such that*

$$\|S(t)x\|_\beta \leqslant C_{\alpha,\beta} \, t^{\alpha-\beta} \|x\|_\alpha, \qquad t \in (0, 1].$$

Proof. Assume $\beta > \alpha$. Then

$$\begin{aligned}\|S(t)x\|_\beta = \|(-L)^\beta S(t)x\| &= \|(-L)^{\beta-\alpha} S(t)(-L)^\alpha x\| \\ &\leqslant \|(-L)^{\beta-\alpha} S(t)\| \, \|x\|_\alpha \leqslant \frac{C_{\alpha,\beta}}{t^{\beta-\alpha}} \|x\|_\alpha.\end{aligned}$$

If $\beta \leqslant \alpha$, then for $x \in \mathcal{B}_\alpha$,

$$\|S(t)x\|_\beta = \|(-L)^{\beta-\alpha} S(t)(-L)^\alpha x\|,$$

and both $(-L)^{\beta-\alpha}$ and $S(t)$ are bounded in this case, so $S(t) : \mathcal{B}_\alpha \to \mathcal{B}_\beta$ is bounded. □

Corollary 2.45. *For every $\alpha \in \mathbb{R}$ and every $\beta \in [\alpha, \alpha+1)$ there exists $C_{\alpha,\beta} > 0$ such that*

$$\|R_t x\|_\beta \leqslant C_{\alpha,\beta}(1+t)^{\beta-\alpha-1} \|x\|_\alpha, \qquad t > 0.$$

Proof. Using the Laplace representation of R_t,

$$\begin{aligned}\|R_t x\|_\beta &\leqslant \int_0^\infty e^{-ts} \|S(s)x\|_\beta \, ds \\ &\leqslant C \int_0^\infty e^{-ts} \|(-L)^{\beta-\alpha} S(s)\| \, ds \, \|x\|_\alpha.\end{aligned}$$

By Theorem 2.43, $\|(-L)^{\beta-\alpha} S(s)\| \leqslant C_{\alpha,\beta} s^{\alpha-\beta}$ for $s \in (0,1]$. For $t \geqslant 1$,

$$\begin{aligned}\int_0^\infty e^{-ts} \|(-L)^{\beta-\alpha} S(s)\| \, ds &\leqslant C_{\alpha,\beta} \int_0^1 e^{-ts} s^{\alpha-\beta} \, ds \\ &\leqslant C_{\alpha,\beta} \Gamma(\beta-\alpha+1)\, t^{\beta-\alpha-1} \leqslant C_{\alpha,\beta}(1+t)^{\beta-\alpha-1}.\end{aligned}$$

If $0 \leqslant t < 1$, split the integral into $\int_0^1 + \int_1^\infty$. The second part is handled as above. For the first part,

$$\int_0^1 e^{-ts} \|(-L)^{\beta-\alpha} S(s)\| \, ds \leqslant C_{\alpha,\beta} \int_0^1 s^{\alpha-\beta} \, ds = \frac{C_{\alpha,\beta}}{\beta-\alpha+1} \leqslant C_{\alpha,\beta}(1+t)^{\beta-\alpha-1}.$$

This proves the claim. □

Corollary 2.46. *Let S be an analytic semigroup with generator L on a Banach space $\mathcal{B}$. Then for every $\alpha \in (0,1)$ there exists $C_\alpha > 0$ such that*

$$\|S(t)x - x\| \leqslant C_\alpha t^\alpha \|x\|_{\mathcal{B}_\alpha}, \qquad x \in \mathcal{B}_\alpha, \quad t \in (0,1].$$

Proof. It suffices to prove the estimate for $x \in \mathcal{D}(L)$, since $\mathcal{D}(L)$ is dense in $\mathcal{B}_\alpha$. For $x \in \mathcal{D}(L)$, we have $\partial_t S(t)x = LS(t)x = S(t)Lx$, hence

$$\begin{aligned}\|S(t)x - x\| &= \left\| \int_0^t S(s)Lx\, ds \right\| = \left\| \int_0^t (-L)^{1-\alpha} S(s)(-L)^\alpha x\, ds \right\| \\ &\leqslant C \int_0^t \|(-L)^{1-\alpha} S(s)\|\, ds\, \|x\|_\alpha \leqslant C \int_0^t s^{\alpha-1}\, ds\, \|x\|_\alpha = Ct^\alpha \|x\|_\alpha.\end{aligned}$$

□

The final theorem shows that, under suitable assumptions, perturbations preserve the interpolation spaces for $\gamma \in [0,1]$. We only sketch the proof.

Theorem 2.47. *Let L_0 be the generator of an analytic semigroup on $\mathcal{B}$ and denote by $\mathcal{B}^0_\gamma$ the corresponding interpolation spaces. Let B be a bounded operator from $\mathcal{B}^0_\alpha$ to $\mathcal{B}$ for some $\alpha \in [0,1)$. Let $\mathcal{B}_\gamma$ be the interpolation spaces associated with $L = L_0 + B$. Then*

$$\mathcal{B}_\gamma = \mathcal{B}^0_\gamma \quad \textit{for every } \gamma \in [0,1].$$

Proof. For $\gamma = 0$ we have $\mathcal{B}_0 = \mathcal{B}^0_0 = \mathcal{B}$. For $\gamma = 1$ we have $\mathcal{B}_1 = \mathcal{D}(L)$ and $\mathcal{B}^0_1 = \mathcal{D}(L_0)$, which coincide.

For $\gamma \in (0,1)$, one shows that there exist constants $C_1, C_2 > 0$ such that

$$C_1 \|(-L_0)^\gamma x\| \leqslant \|(-L)^\gamma x\| \leqslant C_2 \|(-L_0)^\gamma x\|.$$

We use the following estimate (stated without proof): for every $\gamma \in (0,1)$ and $t > 0$ there exists $C > 0$ such that

$$\|BR_t x\| \leqslant C(1+t)^{\alpha-\gamma-1} \|x\|_{\mathcal{B}^0_\gamma}.$$

We also use the resolvent identity

$$R_t = R^0_t + R_t B R^0_t.$$

Applying L and rearranging gives an expression for LR_t in terms of $L_0 R^0_t$ and BR_t. Combining this with the representation formula for fractional powers yields

$$\|x\|_{\mathcal{B}_\gamma} \leqslant C\|x\|_{\mathcal{B}^0_\gamma} + C \int_0^\infty t^{\gamma-1} \|BR_t x\|\, dt.$$

The integral converges by the stated estimate, giving $\|x\|_{\mathcal{B}_\gamma} \leqslant C\|x\|_{\mathcal{B}_\gamma^0}$.

The reverse inequality is obtained similarly, after writing $R_t = R_{t+K} + KR_{t+K}R_t$ for $K > 0$, estimating $\int_0^\infty t^{\gamma-1}\|BR_t x\|\,dt$, and choosing K large enough to absorb a term into the left-hand side. $\square$

3 Linear SPDEs

In this section, we would like to solve the following linear SPDE:

$$dX = LXdt + QdW(t), \quad X(0) = x_0, \tag{3.1}$$

where $X(t) : \Omega \to \mathcal{B}$ is a stochastic process, $L : \mathcal{D}(L) \to \mathcal{B}$ is the generator of a $\mathcal{C}_0$-semigroup on $\mathcal{B}$, W is a cylindrical Wiener process on a Hilbert space $\mathcal{K}$; and $Q : \mathcal{K} \to \mathcal{B}$ is a bounded linear operator.

There are two types of solutions we can consider for the above equation. The first one comes from the notion of weak solutions for PDEs.

Definition 3.1. A $\mathcal{B}$-valued process $x(t)$ is said to be a **weak solution** to (3.1) if, for every $t > 0$, $\int_0^t \|x(s)\| ds < \infty$ almost surely and the identity

$$\langle \ell, x(t) \rangle = \langle \ell, x_0 \rangle + \int_0^t \langle L^* \ell, x(s) \rangle \, ds + \int_0^t \langle Q^* \ell, dW(s) \rangle$$

holds almost surely for every $\ell \in \mathcal{D}(L^*)$.

The second notion is the mild solution.

Definition 3.2. Suppose that there exists a $\mathcal{B}$-valued process $x(t)$ such that, for every $t > 0$, the identity

$$x(t) = S(t)x_0 + \int_0^t S(t-s)QdW(s)$$

holds almost surely (in the sense that it holds when testing against any $\ell \in \mathcal{B}^*$). Then x is said to be the **mild solution** to (3.1).

We will need the following approximation lemma for the proof of next result.

Lemma 3.3 (Approximation Lemma)**.** *For $\varphi \in \mathcal{C}^1([0,t], \mathbb{R})$ and $\ell \in \mathcal{D}\left(L^\dagger\right)$, define*

$$\varphi_\ell(s) = \varphi(s)\ell.$$

Then every element in

$$\mathcal{E} \stackrel{def}{=} \mathcal{C}\left([0,t], \mathcal{D}\left(L^\dagger\right)\right) \cap \mathcal{C}^1\left([0,t], \mathcal{B}^\dagger\right)$$

can be approximated uniformly by linear combinations of functions of the form φ_ℓ.

The next theorem shows the equivalence between the two notions of solutions.

Theorem 3.4. *If the mild solution is almost surely integrable, then it is also a weak solution. Conversely, every weak solution is a mild solution.*

Proof. First of all, we notice that if $y(t)$ with $y(0) = 0$ is a mild solution(or weak solution) to the (3.1), then $x(t) = S(t)x_0 + y(t)$ is also a mild solution(or weak solution) with initial condition x_0. Therefore, without loss of generality, we can assume that $x_0 = 0$.

Suppose now that $x(t)$ with $x(0) = 0$ is a mild solution to (3.1) and is almost surely integrable. Pick any $\ell \in \mathcal{D}(L^\dagger)$, and apply $L^*\ell$ to both sides, we have

$$\langle L^*\ell, x(s)\rangle = \int_0^s \langle L^*\ell, S(s-r)QdW(r)\rangle .$$

Integrate both sides from 0 to t, we have

$$\int_0^t \langle L^*\ell, x(s)\rangle\, ds = \int_0^t \int_0^s \langle L^*\ell, S(s-r)QdW(r)\rangle\, ds$$

By Fubini's theorem, instead of integrating over the region $\{(r, s) : 0 \leqslant r \leqslant s \leqslant t\}$, we can integrate over the region $\{(r, s) : 0 \leqslant r \leqslant t, r \leqslant s \leqslant t\}$, we have

$$\int_0^t \langle L^*\ell, x(s)\rangle\, ds = \int_0^t \left\langle \int_r^t S^*(s-r)L^*\ell ds, QdW(r)\right\rangle$$

Now, since we have $\partial_s S^*(s-r)\ell = L^*S^*(s-r)\ell$, the right hand side becomes

$$\begin{aligned}
&\int_0^t \langle S^*(t-r)\ell, QdW(r)\rangle - \int_0^t \langle \ell, QdW(r)\rangle \\
&= \langle \ell, \int_0^t S(t-r)QdW(r)\rangle - \int_0^t \langle \ell, QdW(r)\rangle \\
&= \langle \ell, x(t)\rangle - \int_0^t \langle \ell, QdW(r)\rangle .
\end{aligned}$$

Combine all the things together, we have for every $\ell \in \mathcal{D}(L^\dagger)$,

$$\langle \ell, x(t)\rangle = \int_0^t \langle L^*\ell, x(s)\rangle\, ds + \int_0^t \langle \ell, QdW(r)\rangle ,$$

this shows that the mild solution is also a weak solution.

Conversely, suppose that $x(t)$ with $x(0) = 0$ is a weak solution to (3.1). Pick any $\ell \in \mathcal{D}(L^\dagger)$ and some time $t > 0$. Define the function $f : [0,t] \to \mathcal{B}^*$ by

$$f(s) = S^*(t-s)\ell.$$

Since we have $S^*\mathcal{D}(L^\dagger) \subset \mathcal{D}(L^\dagger)$ and $\partial_s S^*(t-s)\ell = -L^* S^*(t-s)\ell$, we have $f(s)$ in the set

$$\mathcal{E}{=}\mathcal{C}\left([0,t], \mathcal{D}\left(L^\dagger\right)\right) \cap \mathcal{C}^1\left([0,t], \mathcal{B}^\dagger\right).$$

We claim that for every f in $\mathcal{E}$, we have identity

$$\langle f(t), x(t)\rangle = \int_0^t \left\langle \dot{f}(s) + L^* f(s), x(s)\right\rangle ds + \int_0^t \langle f(s), QdW(s)\rangle.$$

In particular, taking $f(s) = S^*(t-s)\ell$, we have

$$\langle \ell, x(t)\rangle = \int_0^t \langle \ell, S(t-s)QdW(s)\rangle .$$

Since $\mathcal{D}(L^\dagger)$ is dense in $\mathcal{B}^*$, this shows that $x(t)$ is a mild solution to (3.1).

By Lemma 3.3, it suffices to prove the claim for $f = \varphi_\ell$. Now, since we have $x(t)$ is a weak solution, we have

$$d\langle \ell, x(s)\rangle = \langle L^*\ell, x(s)\rangle\, ds + \langle Q^*\ell, dW(s)\rangle .$$

Now, by Itô's product rule, we have

$$d(\varphi(s)\langle \ell, x(s)\rangle) = \dot{\varphi}(s)\langle \ell, x(s)\rangle ds + \varphi(s)\langle L^*\ell, x(s)\rangle ds + \varphi(s)\left\langle Q^*\ell, dW(s)\right\rangle .$$

Integrate both sides from 0 to t, we have

$$\begin{aligned}\langle \varphi(t)\ell, x(t)\rangle &= \int_0^t \dot{\varphi}(s)\langle \ell, x(s)\rangle ds + \int_0^t \varphi(s)\langle L^*\ell, x(s)\rangle ds + \int_0^t \varphi(s)\left\langle Q^*\ell, dW(s)\right\rangle \\ &= \int_0^t \langle \dot{\varphi}(s)\ell + L^*\varphi(s)\ell, x(s)\rangle\, ds + \int_0^t \langle \varphi(s)\ell, QdW(s)\rangle .\end{aligned}$$

This is exactly what we want. This completes the proof. □

3.1 Space-time regularity of solutions

In this section, we study the regularity of the solutions to the linear SPDEs. We will prove that the solutions to the stochastic heat equation are "almost"

$\frac{1}{4}$-Hölder continuous in time and "almost" $\frac{1}{2}$-Hölder continuous in space.

We will need the following lemma for the proof of next theorem.

Lemma 3.5. *For $t \geqslant s \geqslant 0$, we have*

$$\int_s^t (t-r)^{\alpha-1}(r-s)^{-\alpha} dr = \frac{\pi}{\sin(\pi\alpha)}.$$

Proof. Consider the change of variable $u = \dfrac{r-s}{t-s}$, then $(t-r) = (t-s)(1-u)$ and $(r-s) = (t-s)u$. Thus,

$$I = \int_0^1 (1-u)^{\alpha-1} u^{-\alpha} du = B(\alpha, 1-\alpha) = \frac{\Gamma(\alpha)\Gamma(1-\alpha)}{\Gamma(1)} = \frac{\pi}{\sin(\pi\alpha)}.$$

This completes the proof. □

The next theorem shows that the mild solution to (3.1) has continuous sample paths under suitable assumptions.

Theorem 3.6. *Let $\mathcal{H}$ and $\mathcal{K}$ be separable Hilbert spaces, let L be the generator of a $\mathcal{C}_0$-semigroup on $\mathcal{H}$, let $Q : \mathcal{K} \to \mathcal{H}$ be a bounded operator, and let W be a cylindrical Wiener process on $\mathcal{K}$. Assume furthermore that $\|S(t)Q\|_{\mathrm{HS}} < \infty$ for every $t > 0$ and that there exists $\alpha \in \left(0, \frac{1}{2}\right)$ such that*

$$\int_0^1 t^{-2\alpha} \|S(t)Q\|_{\mathrm{HS}}^2 \, dt < \infty.$$

Then the solution x to (3.1) has almost surely continuous sample paths in $\mathcal{H}$.

Proof. Fix $T > 0$ and define the random variable $y : [0,T] \times \Omega \to \mathcal{H}$ by

$$y(t) = \int_0^t (t-s)^{-\alpha} S(t-s) Q \, dW(s).$$

We first establish several properties of $y(t)$. First, we have $\mathbb{E}\|y(t)\|^2 < \infty$ for every $t \in [0,T]$. Consider the estimate

$$\begin{aligned}
\int_0^T r^{-2\alpha}\|S(r)Q\|_{HS}^2\,dr &= \int_0^1 r^{-2\alpha}\|S(r)Q\|_{HS}^2\,dr + \int_1^T r^{-2\alpha}\|S(r)Q\|_{HS}^2\,dr \\
&\leqslant C_1 + \int_1^T \|S(r)Q\|_{HS}^2\,dr \\
&\leqslant C_1 + \|S(1)Q\|_{HS}^2 \int_0^{T-1} \|S(s)\|^2\,ds \leqslant C < \infty.
\end{aligned}$$

Now, by the Itô isometry,

$$\begin{aligned}
E\|y(t)\|^2 &= E\left\|\int_0^t (t-s)^{-\alpha} S(t-s)Q\,dW(s)\right\|^2 \\
&= \int_0^t (t-s)^{-2\alpha}\|S(t-s)Q\|_{HS}^2\,ds \\
&= \int_0^t u^{-2\alpha}\|S(u)Q\|_{HS}^2\,du \leqslant C.
\end{aligned}$$

As a consequence, we claim that for every $p > 0$ there exists $C_p > 0$ such that

$$E\int_0^T \|y(t)\|^p\,dt \leqslant C_p.$$

Indeed, by Corollary 1.10 of Fernique's theorem,

$$E\|y(t)\|^p \leqslant C_p\big(E[\|y(t)\|]\big)^p \leqslant C_p\left(E\|y(t)\|^2\right)^{p/2} \leqslant C_p'.$$

This is enough to proves the claim.

Now we are ready to prove that $x(t)$ has almost surely continuous sample paths. We will use the factorization method. Note that by Lemma 3.5

$$\begin{aligned}
x(t) &= S(t)x_0 + \int_0^t S(t-s)Q\,dW(s) \\
&= S(t)x_0 + \frac{\sin(\pi\alpha)}{\pi}\int_0^t\int_s^t (t-r)^{\alpha-1}(r-s)^{-\alpha}S(t-s)Q\,dW(s)\,dr \\
&= S(t)x_0 + \frac{\sin(\pi\alpha)}{\pi}\int_0^t S(t-r)\left(\int_0^r (r-s)^{-\alpha}S(r-s)Q\,dW(s)\right)(t-r)^{\alpha-1}\,dr \\
&= S(t)x_0 + \frac{\sin(\pi\alpha)}{\pi}\int_0^t (t-r)^{\alpha-1}S(t-r)y(r)\,dr.
\end{aligned}$$

Consider the map $y \mapsto F_y$ defined by

$$F_y(t) = \int_0^t (t-r)^{\alpha-1} S(t-r) y(r)\, dr.$$

If we can show that there exists $p > 0$ such that F_y maps $L^p([0,T],\mathcal{H})$ into $\mathcal{C}([0,T],\mathcal{H})$, then the proof is complete, since we have shown that $y \in L^p([0,T],\mathcal{H})$ for every $p > 0$.

Pick any $p > \dfrac{1}{\alpha}$ and let q be its dual exponent $\left(\text{so } q \in \left(1, \dfrac{1}{1-\alpha}\right)\right)$. By Hölder's inequality,

$$\begin{aligned}
\|F_y(t)\| &\leqslant M_T \int_0^t (t-r)^{\alpha-1} \|y(r)\|\, dr \\
&\leqslant M_T \left(\int_0^t (t-r)^{q(\alpha-1)}\, dr\right)^{1/q} \left(\int_0^t \|y(r)\|^p\, dr\right)^{1/p} \\
&\leqslant C_T \|y\|_{L^p([0,T],\mathcal{H})}.
\end{aligned}$$

This shows that $F_y : L^p([0,T],\mathcal{H}) \to L^\infty([0,T],\mathcal{H})$ is a bounded operator.

Since continuous functions are dense in $L^p([0,T],\mathcal{H})$, it suffices to prove continuity for continuous $y(t)$ with $y(0) = 0$. Fix such a $y(t)$. To show right continuity, note that for $h > 0$ small enough,

$$\begin{aligned}
\|F_y(t+h) - F_y(t)\| \leqslant &\int_0^t \left\|(t+h-r)^{\alpha-1} S(h) - (t-r)^{\alpha-1}\right\| \, \|S(t-r)\| \, \|y(r)\|\, dr \\
&+ \int_t^{t+h} (t+h-r)^{\alpha-1} \, \|S(t+h-r)\| \, \|y(r)\|\, dr.
\end{aligned}$$

The second integral can be bounded by

$$\begin{aligned}
I_2 &\leqslant M \int_t^{t+h} (t+h-r)^{\alpha-1}\, dr \\
&\leqslant M \int_0^h s^{\alpha-1}\, ds = \frac{M}{\alpha} h^\alpha \to 0, \qquad \text{as } h \to 0.
\end{aligned}$$

To bound the first integral, note that for h small enough,

$$A_h(r) := \left\|(t+h-r)^{\alpha-1} S(h) - (t-r)^{\alpha-1}\right\| \leqslant C'(t-r)^{\alpha-1},$$

which is integrable on $[0, t]$. Moreover,

$$A_h(r) \leqslant (t+h-r)^{\alpha-1}\|S(h)-I\| + \left|(t+h-r)^{\alpha-1} - (t-r)^{\alpha-1}\right|,$$

and the right-hand side tends to 0 as $h \to 0$ for each fixed $r < t$. By the dominated convergence theorem, the first integral also tends to 0 as $h \to 0$. This proves right continuity. Left continuity can be shown similarly. This completes the proof. $\square$

We now give a general result that tells us precisely in which interpolation space one can expect to find the solution to a linear SPDE associated with an analytic semigroup.

Theorem 3.7. *Consider a linear SPDE on a Hilbert space $\mathcal{H}$. Assume that L generates an analytic semigroup, and denote by $\mathcal{H}_\alpha$ the corresponding interpolation spaces. Suppose that there exists $\alpha \geqslant 0$ such that $Q : \mathcal{K} \to \mathcal{H}_\alpha$ is bounded, and that there exists $\beta \in \left(0, \frac{1}{2} + \alpha\right]$ such that $\left\|(-L)^{-\beta}\right\|_{\mathrm{HS}} < \infty$. Then the solution x takes values in $\mathcal{H}_\gamma$ for every $\gamma < \gamma_0 = \frac{1}{2} + \alpha - \beta$.*

Proof. Recall that the solution has the mild form

$$x(t) = S(t)x_0 + \int_0^t S(t-s)Q\,dW(s).$$

Since $S(t)$ maps $\mathcal{H}$ into $\mathcal{H}_\gamma$ for every $\gamma \geqslant 0$, it suffices to show that the stochastic convolution takes values in $\mathcal{H}_\gamma$ for every $\gamma < \gamma_0$. For this purpose, we will show that for every $T > 0$ and every $\gamma < \gamma_0$, one has

$$\mathbb{E}\left\|(-L)^\gamma \int_0^T S(T-s)Q\,dW(s)\right\|^2 < \infty.$$

By the Itô isometry, it is enough to show that

$$\begin{aligned}
I(T) := \mathbb{E}\left\|(-L)^\gamma \int_0^T S(T-s)Q\,dW(s)\right\|^2 \\
&= \int_0^T \|(-L)^\gamma S(T-s)Q\|_{\mathrm{HS}}^2\,ds \\
&= \int_0^T \|(-L)^\gamma S(s)Q\|_{\mathrm{HS}}^2\,ds < \infty.
\end{aligned}$$

Now note that $Q : \mathcal{K} \to \mathcal{H}_\alpha$ is bounded by assumption, so $\|(-L)^\alpha Q\| < \infty$. Hence

$$
\begin{aligned}
I(T) &\leqslant C \int_0^T \|(-L)^{\gamma-\alpha} S(s)\|_{\mathrm{HS}}^2 \, ds \\
&\leqslant C \int_0^T \|(-L)^{\gamma-\alpha+\beta} S(s)\|^2 \, ds.
\end{aligned}
$$

The last inequality follows from the assumption that $(-L)^{-\beta}$ is Hilbert–Schmidt.

By the theory of interpolation spaces, if $\gamma - \alpha + \beta \leqslant 0$, then

$$
\|(-L)^{\gamma-\alpha+\beta} S(s)\| \leqslant C.
$$

On the other hand, if $\gamma - \alpha + \beta > 0$, then

$$
\|(-L)^{\gamma-\alpha+\beta} S(s)\| \leqslant C\, s^{-(\gamma-\alpha+\beta)} = C\, s^{\alpha-\gamma-\beta}.
$$

Therefore, in this case the integrand is bounded by $Cs^{2(\alpha-\gamma-\beta)}$, which is integrable near 0 provided

$$
2(\alpha - \gamma - \beta) > -1 \quad \Longleftrightarrow \quad \gamma < \frac{1}{2} + \alpha - \beta.
$$

This concludes the proof. □

The next example shows that the solution to the stochastic heat equation lies in the Sobolev space H^s for every $s < \frac{1}{2}$.

Example 3.8. Consider the stochastic heat equation on $[0,1]$ with periodic boundary conditions (driven by space-time white noise),

$$
dx = \Delta x \, dt + dW,
$$

where W is a cylindrical Wiener process on $L^2([0,1])$. We claim that the solution lies in the fractional Sobolev space H^s for every $s < \frac{1}{2}$.

We apply Theorem 3.7 with $\mathcal{H} = \mathcal{K} = L^2([0,1])$. In this case, $L = \Delta$, $Q = \mathrm{Id}$, and $\alpha = 0$. Fix any $\beta > \frac{1}{4}$. Consider the orthonormal basis $e_k(x) = e^{i2\pi kx}$ of $L^2([0,1])$. Then for each nonzero $k \in \mathbb{Z}$,

$$
-\Delta e_k = (2\pi|k|)^2 e_k.
$$

Since Δ is self-adjoint, the spectral theorem yields

$$(-\Delta)^{-\beta} e_k = (2\pi|k|)^{-2\beta} e_k.$$

Therefore,

$$\|(-\Delta)^{-\beta}\|_{\mathrm{HS}}^2 = \sum_{k\in\mathbb{Z}} \|(-\Delta)^{-\beta} e_k\|^2 \leqslant C \sum_{k\in\mathbb{Z}\setminus\{0\}} |k|^{-4\beta} < \infty.$$

Thus, by Theorem 3.7, the solution takes values in $\mathcal{H}_\gamma = H^{2\gamma}$ for every $\gamma < \gamma_0 = \frac{1}{2} - \beta$. Letting $\beta \downarrow \frac{1}{4}$ proves the claim.

By adding a smoothing perturbation to the stochastic heat equation, one can obtain solutions with higher regularity.

Example 3.9. Consider the following modified stochastic heat equation on $[0, 1]$ with periodic boundary conditions:

$$dx = \Delta x \, dt + (1 - \Delta)^{-\gamma} \, dW,$$

where W is a cylindrical Wiener process on $L^2([0, 1])$. We claim that x takes values in H^s for every $s < \frac{1}{2} + 2\gamma$.

Let us first show that $Q = (1-\Delta)^{-\gamma}$ is bounded from $L^2([0, 1])$ to $H^{2\gamma}$. For each $k \in \mathbb{Z}$, we have

$$-\Delta e_k(x) = (2\pi k)^2 e_k(x).$$

By the spectral calculus, this implies

$$(1 - \Delta)^{-\gamma} e_k(x) = (1 + 4\pi^2 k^2)^{-\gamma} e_k(x).$$

For any $f \in L^2([0, 1])$, write its Fourier series as

$$f(x) = \sum_{k\in\mathbb{Z}} \hat{f}_k e_k(x).$$

Then

$$(1 - \Delta)^{-\gamma} f(x) = \sum_{k\in\mathbb{Z}} \hat{f}_k (1 + 4\pi^2 k^2)^{-\gamma} e_k(x).$$

Computing the $H^{2\gamma}$ norm of $(1-\Delta)^{-\gamma}f$, we obtain

$$\|(1-\Delta)^{-\gamma}f\|_{H^{2\gamma}}^2 = \sum_{k\in\mathbb{Z}} |\hat{f}_k|^2 \frac{(1+k^2)^{2\gamma}}{(1+4\pi^2k^2)^{2\gamma}} \leqslant C \sum_{k\in\mathbb{Z}} |\hat{f}_k|^2 = C\|f\|_{L^2}^2.$$

Thus Q is bounded from $L^2([0,1])$ to $H^{2\gamma}$, i.e. $\alpha=\gamma$ in Theorem 3.7.

By the same computation as in the Example 3.8, for any $\beta > \dfrac{1}{4}$ we have

$$\|(-\Delta)^{-\beta}\|_{\mathrm{HS}} < \infty.$$

Therefore, Theorem 3.7 implies that the solution takes values in $\mathcal{H}_\rho = H^{2\rho}$ for every

$$\rho < \rho_0 = \frac{1}{2} + \gamma - \beta.$$

Letting $\beta \downarrow \dfrac{1}{4}$ yields that the solution takes values in H^s for every

$$s < \frac{1}{2} + 2\gamma.$$

This completes the proof.

Now, using the spatial regularity, we can deduce time regularity of the solution.

Theorem 3.10. *Consider the same setting as in Theorem 3.7 and fix $\gamma < \gamma_0$. Then, for every $t > 0$, the process x is almost surely δ-Hölder continuous in $\mathcal{H}_\gamma$ for every*

$$\delta < \min\left\{\frac{1}{2}, \gamma_0 - \gamma\right\}.$$

Proof. Let us first consider the case $\gamma_0 - \gamma \leqslant \frac{1}{2}$. For $\gamma \leqslant \tilde{\gamma} < \gamma_0$, if we can show that for any fixed interval $[t_0, T]$ and any s,t in that interval,

$$\mathbf{E}\|x(t)-x(s)\|_\gamma^2 \leqslant C|t-s|^{2(\tilde{\gamma}-\gamma)},$$

then

$$\mathbf{E}\|x(t)-x(s)\|_\gamma \leqslant C^{1/2}|t-s|^{\tilde{\gamma}-\gamma}.$$

Kolmogorov's continuity theorem then implies that x is almost surely δ-Hölder continuous in $\mathcal{H}_\gamma$ for every $\delta < \tilde{\gamma} - \gamma$. Letting $\tilde{\gamma} \uparrow \gamma_0$ yields the desired result.

Now, for $t > s$, we have

$$\begin{aligned}
\int_0^t S(t-r)Q\,dW(r) &= \int_0^s S(t-r)Q\,dW(r) + \int_s^t S(t-r)Q\,dW(r) \\
&= S(t-s)\int_0^s S(s-r)Q\,dW(r) + \int_s^t S(t-r)Q\,dW(r) \\
&= S(t-s)(x(s) - S(s)x_0) + \int_s^t S(t-r)Q\,dW(r) \\
&= S(t-s)x(s) - S(t)x_0 + \int_s^t S(t-r)Q\,dW(r).
\end{aligned}$$

Thus,

$$x(t) = S(t-s)x(s) + \int_s^t S(t-r)Q\,dW(r).$$

Therefore,

$$\begin{aligned}
\mathbf{E}\|x(t) - x(s)\|_\gamma^2 &= \mathbf{E}\|S(t-s)x(s) - x(s)\|_\gamma^2 + \mathbf{E}\left\|\int_s^t (-L)^\gamma S(t-r)Q\,dW(r)\right\|^2 \\
&= \mathbf{E}\|S(t-s)x(s) - x(s)\|_\gamma^2 + \int_0^{t-s} \|(-L)^\gamma S(r)Q\|_{\mathrm{HS}}^2\,dr.
\end{aligned}$$

By the theory of interpolation spaces, we have

$$\begin{aligned}
\|S(t-s)x(s) - x(s)\|_\gamma &= \|(S(t-s) - I)(-L)^\gamma x(s)\| \\
&= |t-s|^{\tilde{\gamma}-\gamma}\,\|(-L)^\gamma x(s)\|_{\tilde{\gamma}-\gamma} \\
&\leqslant C|t-s|^{\tilde{\gamma}-\gamma}\,\|x(s)\|_{\tilde{\gamma}}.
\end{aligned}$$

Hence

$$\mathbf{E}\|S(t-s)x(s) - x(s)\|_\gamma^2 \leqslant C|t-s|^{2(\tilde{\gamma}-\gamma)}\mathbf{E}\|x(s)\|_{\tilde{\gamma}}^2 \leqslant C|t-s|^{2(\tilde{\gamma}-\gamma)}.$$

The last inequality follows from Theorem 3.7.

Next, we estimate the second term by the same trick as in Theorem 3.7:

$$\begin{aligned}
\|(-L)^\gamma S(r)Q\|_{\mathrm{HS}} &= \|(-L)^{\gamma-\alpha}S(r)(-L)^\alpha Q\|_{\mathrm{HS}} \\
&\leqslant C\|(-L)^{\gamma-\alpha}S(r)\|_{\mathrm{HS}} \\
&\leqslant C\|(-L)^{-\beta}(-L)^{\gamma-\alpha+\beta}S(r)\|_{\mathrm{HS}} \\
&\leqslant C\|(-L)^{\gamma-\alpha+\beta}S(r)\|.
\end{aligned}$$

Recall that $\gamma - \alpha + \beta = -(\gamma_0 - \gamma) + \frac{1}{2} \geqslant 0$ in the present case. By the theory of interpolation spaces,

$$\|(-L)^\gamma S(r)Q\|_{\mathrm{HS}} \leqslant C r^{\gamma_0 - \gamma - \frac{1}{2}}.$$

Therefore,

$$\int_0^{t-s} \|(-L)^\gamma S(r)Q\|_{\mathrm{HS}}^2 \, dr \leqslant C \int_0^{t-s} r^{2(\gamma_0-\gamma)-1} \, dr = C|t-s|^{2(\gamma_0-\gamma)}.$$

Combining the two terms, we obtain

$$\mathbf{E}\|x(t) - x(s)\|_\gamma^2 \leqslant C|t-s|^{2(\tilde{\gamma}-\gamma)} + C|t-s|^{2(\gamma_0-\gamma)} \leqslant C|t-s|^{2(\tilde{\gamma}-\gamma)}.$$

The last inequality uses the fact that we work on a compact interval $[t_0, T]$ and $\tilde{\gamma} - \gamma < \gamma_0 - \gamma$.

Now, let us consider the case $\gamma_0 - \gamma > \frac{1}{2}$. If we can show that for any fixed interval $[t_0, T]$ and any s, t in that interval,

$$\mathbf{E}\|x(t) - x(s)\|_\gamma^2 \leqslant C|t-s|,$$

then

$$\mathbf{E}\|x(t) - x(s)\|_\gamma \leqslant C^{1/2}|t-s|^{\frac{1}{2}}.$$

Kolmogorov's continuity theorem then implies that x is almost surely δ-Hölder continuous in $\mathcal{H}_\gamma$ for every $\delta < \frac{1}{2}$.

As before, for $t > s$ we use

$$x(t) = S(t-s)x(s) + \int_s^t S(t-r)Q \, dW(r),$$

so

$$\mathbf{E}\|x(t) - x(s)\|_\gamma^2 = \mathbf{E}\|S(t-s)x(s) - x(s)\|_\gamma^2 + \int_0^{t-s} \|(-L)^\gamma S(r)Q\|_{\mathrm{HS}}^2 \, dr.$$

By the theory of interpolation spaces,

$$\|S(t-s)x(s) - x(s)\|_\gamma = \|(S(t-s) - I)(-L)^\gamma x(s)\| \leqslant C|t-s|^{1/2} \, \|x(s)\|_{\gamma + \frac{1}{2}}.$$

Hence

$$\mathbf{E}\|S(t-s)x(s)-x(s)\|_\gamma^2 \leqslant C|t-s|\,\mathbf{E}\|x(s)\|_{\gamma+\frac{1}{2}}^2 \leqslant C|t-s|,$$

where the last inequality follows from Theorem 3.7.

For the second term, we again write

$$\|(-L)^\gamma S(r)Q\|_{\mathrm{HS}} \leqslant C\|(-L)^{\gamma-\alpha+\beta}S(r)\|.$$

Now $\gamma-\alpha+\beta = -(\gamma_0-\gamma)+\frac{1}{2} \leqslant 0$ in this case, so the interpolation theory yields

$$\|(-L)^{\gamma-\alpha+\beta}S(r)\| \leqslant C.$$

Therefore,

$$\int_0^{t-s} \|(-L)^\gamma S(r)Q\|_{\mathrm{HS}}^2\,dr \leqslant C|t-s|.$$

Combining both estimates gives

$$\mathbf{E}\|x(t)-x(s)\|_\gamma^2 \leqslant C|t-s|.$$

This completes the proof. □

3.2 Markov semigroups and invariant measures

In this subsection, we introduce the notions of Markov semigroups and invariant measures associated with solutions to the linear SPDEs (3.1). Throughout, we denote by $\mathbf{B}_b(\mathcal{B})$ the space of bounded measurable functions from a Banach space $\mathcal{B}$ to $\mathbf{R}$.

Definition 3.11. Consider a family of operators $\mathcal{P}_t : \mathbf{B}_b(\mathcal{B}) \to \mathbf{B}_b(\mathcal{B})$ defined by

$$(\mathcal{P}_t\varphi)(x) = \mathbf{E}\,\varphi\left(S(t)x + \int_0^t S(t-s)Q\,dW(s)\right).$$

The family $\{\mathcal{P}_t\}_{t\geqslant 0}$ is called the **Markov semigroup** associated with the SPDE

$$du = Lu\,dt + Q\,dW(t).$$

The operators $\mathcal{P}_t$ are called **Markov operators**.

If we denote the law of the solution to (3.1) at time t with initial condition

x by $\mathcal{P}_t(x,\cdot)$, then the Markov operator can be written as

$$(\mathcal{P}_t\varphi)\,(x) = \int_{\mathcal{B}} \varphi(y)\,\mathcal{P}_t(x,dy).$$

On the other hand, we can define the dual semigroup $\mathcal{P}_t^*$ acting on probability measures on $\mathcal{B}$ by

$$(\mathcal{P}_t^*\mu)\,(A) = \int_{\mathcal{B}} \mathcal{P}_t(x,A)\,\mu(dx).$$

This is precisely the law of the solution to (3.1) at time t when the initial condition is distributed according to μ.

We have the following dual formulation:

Theorem 3.12 (Duality identity)**.** *For every $t \geqslant 0$, $\varphi \in \mathbf{B}_b(\mathcal{B})$, and every probability measure μ on $\mathcal{B}$, we have*

$$\int_{\mathcal{B}} (\mathcal{P}_t\varphi)(x)\,\mu(dx) = \int_{\mathcal{B}} \varphi(x)\,(\mathcal{P}_t^*\mu)(dx).$$

Proof. Let X_0 be a random variable with law μ, independent of the Wiener process W, and set

$$X_t = S(t)X_0 + \int_0^t S(t-s)Q\,dW(s).$$

By definition of $\mathcal{P}_t^*\mu$, the law of X_t is exactly $\mathcal{P}_t^*\mu$. Hence

$$\int_{\mathcal{B}} \varphi(x)\,(\mathcal{P}_t^*\mu)(dx) = \mathbf{E}\big[\varphi(X_t)\big].$$

By the definition of $\mathcal{P}_t$, we also have

$$\mathbf{E}\big[\varphi(X_t) \mid X_0\big] = (\mathcal{P}_t\varphi)(X_0).$$

Taking expectations on both sides yields

$$\mathbf{E}\big[\varphi(X_t)\big] = \int_{\mathcal{B}} (\mathcal{P}_t\varphi)(x)\,\mu(dx),$$

which proves the desired identity. □

Now, we can define the invariant measure associate to the linear SPDE (3.1)

Definition 3.13. A probability measure μ on $\mathcal{B}$ is called an **invariant mea-**

sure for the linear SPDE (3.1) if for every $t \geqslant 0$, we have

$$\mathcal{P}_t^* \mu = \mu,$$

where $\mathcal{P}_t^*$ is the dual Markov semigroup associated to the linear SPDE (3.1).

Consider (3.1) with solutions in a Hilbert space $\mathcal{H}$. We prove some equivalent conditions for the existence of invariant measures for linear SPDEs. We start with a lemma.

Lemma 3.14. *For every $t \geqslant 0$, define the self-adjoint operator $Q_t : \mathcal{H} \to \mathcal{H}$ by*

$$Q_t = \int_0^t S(s) Q Q^* S^*(s)\, ds.$$

Then for every invariant measure μ on $\mathcal{H}$, we have

$$\widehat{\mu}(x) = \hat{\mu}\left(S^*(t)x\right) e^{-\frac{1}{2}\langle x, Q_t x\rangle}.$$

Proof. First, we compute $\widehat{\mathcal{P}_t^* \mu}(x)$. Let $\varphi_x(y) = e^{i\langle x, y\rangle}$ be the complex exponential function. Then

$$\widehat{\mathcal{P}_t^* \mu}(x) = \int_{\mathcal{H}} e^{i\langle x, y\rangle}\, (\mathcal{P}_t^* \mu)(dy) = \int_{\mathcal{H}} \varphi_x(y)\, (\mathcal{P}_t^* \mu)(dy) = \int_{\mathcal{H}} (\mathcal{P}_t \varphi_x)(y)\, \mu(dy).$$

By definition of $\mathcal{P}_t$, we have

$$\begin{aligned}
(\mathcal{P}_t \varphi_x)(y) &= \mathbb{E}\left[\varphi_x\left(S(t)y + \int_0^t S(t-s) Q\, dW(s)\right)\right] \\
&= \mathbb{E}\left[\exp\Big(i\Big\langle x,\, S(t)y + \int_0^t S(t-s) Q\, dW(s)\Big\rangle\Big)\right] \\
&= \exp\big(i\langle S^*(t)x, y\rangle\big)\, \mathbb{E}\left[\exp\Big(i\Big\langle x, \int_0^t S(t-s) Q\, dW(s)\Big\rangle\Big)\right].
\end{aligned}$$

Since $\left\langle x, \int_0^t S(t-s) Q\, dW(s)\right\rangle$ is a real-valued Gaussian random variable with mean zero and variance

$$\begin{aligned}\mathbb{E}\left|\left\langle x, \int_0^t S(t-s)Q\,dW(s)\right\rangle\right|^2 &= \mathbb{E}\left|\int_0^t \langle Q^* S^*(t-s)x, dW(s)\rangle\right|^2 \\ &= \int_0^t \|Q^* S^*(t-s)x\|^2\,ds \\ &= \int_0^t \|Q^* S^*(s)x\|^2\,ds \\ &= \langle x, Q_t x\rangle,\end{aligned}$$

we obtain

$$(\mathcal{P}_t\varphi_x)(y) = e^{i\langle S^*(t)x,y\rangle}\, e^{-\frac{1}{2}\langle x,Q_t x\rangle}.$$

Therefore,

$$\widehat{\mathcal{P}_t^*\mu}(x) = \int_{\mathcal{H}} e^{i\langle S^*(t)x,y\rangle}\, e^{-\frac{1}{2}\langle x,Q_t x\rangle}\,\mu(dy) = \hat{\mu}\left(S^*(t)x\right) e^{-\frac{1}{2}\langle x,Q_t x\rangle}.$$

Using $\mathcal{P}_t^*\mu = \mu$, this completes the proof. □

The following theorem characterises the invariant measure of the linear SPDE (3.1).

Theorem 3.15. *Consider* (3.1) *with solutions in a Hilbert space $\mathcal{H}$ and define the self-adjoint operator $Q_t : \mathcal{H} \to \mathcal{H}$ by*

$$Q_t = \int_0^t S(s)QQ^*S^*(s)\,ds.$$

Then there exists an invariant measure μ for (3.1) *if and only if one of the following two equivalent conditions holds:*

(i) There exists a positive definite trace class operator $Q_\infty : \mathcal{H} \to \mathcal{H}$ such that

$$2\,\mathrm{Re}\,\langle Q_\infty L^* x, x\rangle + \|Q^* x\|^2 = 0$$

for every $x \in \mathcal{D}(L^*)$.

(ii) One has $\sup_{t>0} \operatorname{tr} Q_t < \infty$.

Proof. We first show that the existence of an invariant measure implies (ii), then that (ii) implies (i), and finally that (i) implies the existence of an invariant measure.

We first show the existence of an invariant measure implies (ii). By Lemma 3.14, we have

$$\hat{\mu}(x) = \hat{\mu}\left(S^*(t)x\right) e^{-\frac{1}{2}\langle x, Q_t x\rangle}.$$

Since $|\hat{\mu}(x)| \leqslant 1$ for every $x \in \mathcal{H}$, we obtain

$$|\hat{\mu}(x)| \leqslant e^{-\frac{1}{2}\langle x, Q_t x\rangle}.$$

Taking logarithms yields

$$\langle x, Q_t x\rangle \leqslant -2\log|\hat{\mu}(x)|.$$

Now we bound $|\hat{\mu}(x)|$ from below. Choose $R > 0$ such that $\mu(\|y\| > R) \leqslant \frac{1}{8}$. Then

$$\begin{aligned}
|1-\hat{\mu}(x)| &= \left|\int_{\mathcal{H}} \left(1 - e^{i\langle x,y\rangle}\right)\mu(dy)\right| \\
&\leqslant \int_{\|y\|\leqslant R} |1 - e^{i\langle x,y\rangle}|\mu(dy) + \int_{\|y\|>R} |1 - e^{i\langle x,y\rangle}|\mu(dy) \\
&\leqslant \int_{\|y\|\leqslant R} |\langle x,y\rangle|\mu(dy) + 2\mu(\|y\| > R) \\
&\leqslant \left(\int_{\|y\|\leqslant R} |\langle x,y\rangle|^2\mu(dy)\right)^{1/2} + \frac{1}{4}.
\end{aligned}$$

Define a symmetric positive definite operator $A_R : \mathcal{H} \to \mathcal{H}$ by

$$\langle x, A_R x\rangle = \int_{\|y\|\leqslant R} |\langle y, x\rangle|^2\mu(dy).$$

Note that A_R is trace class since

$$\operatorname{tr} A_R = \sum_{n=1}^{\infty}\langle e_n, A_R e_n\rangle = \sum_{n=1}^{\infty}\int_{\|y\|\leqslant R} |\langle y e_n\rangle|^2\mu(dy) = \int_{\|y\|\leqslant R} \|y\|^2\mu(dy) \leqslant R^2.$$

If $x \in \mathcal{H}$ satisfies $\langle x, A_R x\rangle \leqslant \dfrac{1}{4}$, then

$$|\hat{\mu}(x)| \geqslant 1 - |1-\hat{\mu}(x)| \geqslant \frac{1}{4},$$

this implies that

$$\langle x, Q_t x\rangle \leqslant -2\log|\hat{\mu}(x)| \leqslant -2\log\frac{1}{4} = 2\log 4.$$

For arbitrary $x \in \mathcal{H}$, set $\tilde{x} = \dfrac{x}{2\sqrt{\langle x, A_R x\rangle}}$. Then $\langle \tilde{x}, A_R\tilde{x}\rangle = \dfrac{1}{4}$, and hence

$$\langle x, Q_t x\rangle \leqslant 4\,\langle \tilde{x}, Q_t\tilde{x}\rangle\,\langle x, A_R x\rangle \leqslant 8\log 4\,\langle x, A_R x\rangle.$$

Applying this with $x = e_n$ and summing over n, we obtain

$$\operatorname{tr} Q_t = \sum_{n=1}^{\infty}\langle e_n, Q_t e_n\rangle \leqslant 8\log 4\sum_{n=1}^{\infty}\langle e_n, A_R e_n\rangle = 8\log 4\operatorname{tr} A_R \leqslant 8R^2\log 4.$$

Therefore $\sup_{t>0}\operatorname{tr} Q_t < \infty$, which proves (ii).

Now, we show that (ii) implies (i). The assumption $\sup_{t>0}\operatorname{tr} Q_t < \infty$ implies that the limit

$$Q_\infty = \lim_{t\to\infty} Q_t = \int_0^\infty S(s)QQ^*S^*(s)\,ds$$

is a well-defined positive definite trace class operator. Moreover, for every $x \in \mathcal{D}(L^*)$,

$$\begin{aligned}\langle x, Q_\infty x\rangle &= \int_0^\infty \|Q^*S^*(s)x\|^2\,ds\\ &= \int_0^t \|Q^*S^*(s)x\|^2\,ds + \int_t^\infty \|Q^*S^*(s)x\|^2\,ds.\end{aligned}$$

Changing variables $s \mapsto s+t$ in the second term yields

$$\begin{aligned}\langle x, Q_\infty x\rangle &= \int_0^t \|Q^*S^*(s)x\|^2\,ds + \int_0^\infty \|Q^*S^*(s)S^*(t)x\|^2\,ds\\ &= \int_0^t \|Q^*S^*(s)x\|^2\,ds + \langle S^*(t)x, Q_\infty S^*(t)x\rangle\,.\end{aligned}$$

Differentiating with respect to t gives

$$0 = \|Q^*S^*(t)x\|^2 + \langle L^*S^*(t)x, Q_\infty S^*(t)x\rangle + \langle S^*(t)x, Q_\infty L^*S^*(t)x\rangle.$$

Evaluating at $t = 0$, we obtain

$$0 = \|Q^*x\|^2 + \langle L^*x, Q_\infty x\rangle + \langle x, Q_\infty L^*x\rangle.$$

Since Q_∞ is self-adjoint, this becomes

$$0 = \|Q^* x\|^2 + 2\,\mathrm{Re}\,\langle Q_\infty L^* x, x\rangle\,,$$

which is (i).

Now, we show (i) implies existence of an invariant measure. For $x \in \mathcal{D}(L^*)$, define

$$F_x(t) = \langle S^*(t)x, Q_\infty S^*(t)x\rangle.$$

Then F_x is differentiable with

$$\frac{d}{dt}F_x(t) = 2\,\mathrm{Re}\,\langle L^* S^*(t)x, Q_\infty S^*(t)x\rangle = -\|Q^* S^*(t)x\|^2\,.$$

Integrating from 0 to t yields

$$\begin{aligned}\langle S^*(t)x, Q_\infty S^*(t)x\rangle &= \langle x, Q_\infty x\rangle - \int_0^t \|Q^* S^*(s)x\|^2\, ds\\ &= \langle x, Q_\infty x\rangle - \langle x, Q_t x\rangle.\end{aligned}$$

Hence

$$\langle x, Q_\infty x\rangle = \langle x, Q_t x\rangle + \langle S^*(t)x, Q_\infty S^*(t)x\rangle = \langle x, \big(Q_t + S(t)Q_\infty S^*(t)\big)x\rangle.$$

Let μ_∞ be the centred Gaussian measure with covariance operator Q_∞, so that

$$\hat{\mu}_\infty(x) = e^{-\frac{1}{2}\langle x, Q_\infty x\rangle}.$$

Then for every $t \geqslant 0$,

$$\begin{aligned}\widehat{\mathcal{P}_t^*\mu_\infty}(x) &= \hat{\mu}_\infty\left(S^*(t)x\right) e^{-\frac{1}{2}\langle x, Q_t x\rangle}\\ &= e^{-\frac{1}{2}\langle S^*(t)x, Q_\infty S^*(t)x\rangle}\, e^{-\frac{1}{2}\langle x, Q_t x\rangle}\\ &= e^{-\frac{1}{2}\langle x, (S(t)Q_\infty S^*(t) + Q_t)x\rangle}\\ &= e^{-\frac{1}{2}\langle x, Q_\infty x\rangle}\\ &= \hat{\mu}_\infty(x).\end{aligned}$$

Therefore $\mathcal{P}_t^*\mu_\infty = \mu_\infty$ for every $t \geqslant 0$, so μ_∞ is an invariant measure. This completes the proof. $\square$

Now, we want to characterise all invariant measures for the linear SPDE

(3.1). We need the following lemmas for that.

Lemma 3.16. *Let μ_t be the centered Gaussian measure with covariance operator*

$$Q_t = \int_0^t S(s)QQ^*S^*(s)\,ds.$$

Then $\{\mu_t\}_{t>0}$ is tight.

Proof. We claim that there exists a sequence of bounded linear operators $\{A_n\}_{n=1}^{\infty}$ on $\mathcal{H}$ such that

(i) One has $\|A_{n+1}x\| \geqslant \|A_n x\|$ for every $x \in \mathcal{H}$ and every $n \in \mathbb{N}$.

(ii) The set $B_R = \{x : \sup_n \|A_n x\| \leqslant R\}$ is compact for every $R > 0$.

(iii) One has $\sup_n \operatorname{tr}(A_n Q_t A_n^*) < \infty$.

Assume the claim for the moment. Let Y_t be a random variable with distribution μ_t. Then

$$\begin{aligned}
\mathbb{E}\|A_n Y_t\|^2 &= \sum_{k=1}^{\infty} \mathbb{E}|\langle A_n Y_t, e_k\rangle|^2 \\
&= \sum_{k=1}^{\infty} \mathbb{E}|\langle Y_t, A_n^* e_k\rangle|^2 \\
&= \sum_{k=1}^{\infty} \langle A_n^* e_k, Q_t A_n^* e_k\rangle \\
&= \operatorname{tr}(A_n Q_t A_n^*) < \infty.
\end{aligned}$$

Thus, by Chebyshev's inequality,

$$\mathbf{P}(\|A_n Y_t\| > R) \leqslant \frac{\mathbb{E}\|A_n Y_t\|^2}{R^2} \leqslant \frac{C}{R^2}.$$

By definition of B_R,

$$\mu_t(\mathcal{H} \setminus B_R) = \mathbf{P}\left(\sup_n \|A_n Y_t\| > R\right) \leqslant \mathbf{P}(\|A_n Y_t\| > R) \leqslant \frac{C}{R^2}.$$

This shows the tightness of the family $\{\mu_t\}_{t>0}$.

It remains to prove the claim. Since Q_t is self-adjoint, positive, and trace class on $\mathcal{H}$, the spectral theorem for compact self-adjoint operators yields an orthonormal basis $\{e_n\}_{n=1}^{\infty}$ of $\mathcal{H}$ consisting of eigenvectors of Q_t, with corresponding eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$, such that:

(i) $\lambda_n \geqslant 0$ for every $n \in \mathbb{N}$,

(ii) $\lambda_n \to 0$ as $n \to \infty$,

(iii) since Q_t is trace class, $\sum_{n=1}^{\infty} \lambda_n = \mathrm{tr}(Q_t) < \infty$.

We choose a sequence of positive numbers $\{a_n\}_{n=1}^{\infty}$ such that

$$\sum_{n=1}^{\infty} a_n \lambda_n < \infty, \qquad \lim_{n\to\infty} a_n = \infty.$$

Let $N_0 = 1$ and, for $m \geqslant 0$, let N_{m+1} be the smallest integer such that

$$\sum_{k=N_{m+1}}^{\infty} \lambda_k \leqslant 2^{-(m+1)}.$$

Define $a_n = m+1$ for every $N_m \leqslant n < N_{m+1}$. Then $a_n \to \infty$ as $n \to \infty$, and

$$\sum_{n=1}^{\infty} a_n \lambda_n = \sum_{m=0}^{\infty} \sum_{k=N_m}^{N_{m+1}-1} a_k \lambda_k \leqslant \sum_{m=0}^{\infty} (m+1) \sum_{k=N_m}^{\infty} \lambda_k \leqslant \sum_{m=0}^{\infty} (m+1) 2^{-m} < \infty.$$

Now define $\{A_n\}_{n=1}^{\infty}$ on $\mathcal{H}$ by

$$A_n e_k = \begin{cases} \sqrt{a_k}\, e_k, & k \leqslant n, \\ 0, & k > n. \end{cases}$$

For $h \in \mathcal{H}$ with expansion $h = \sum_{k=1}^{\infty} h_k e_k$, we have

$$A_n h = \sum_{k=1}^{n} \sqrt{a_k}\, h_k e_k.$$

It is immediate that $\|A_{n+1}h\| \geqslant \|A_n h\|$ for every $h \in \mathcal{H}$ and $n \in \mathbb{N}$. Moreover, for every $R > 0$,

$$B_R = \left\{ h \in \mathcal{H} : \sup_n \|A_n h\| \leqslant R \right\} = \left\{ h \in \mathcal{H} : \sum_{k=1}^{\infty} a_k |h_k|^2 \leqslant R^2 \right\},$$

which is a compact subset of $\mathcal{H}$.

Finally, since $A_n = A_n^*$ and $Q_t e_k = \lambda_k e_k$, we have

$$A_n Q_t A_n^* e_k = \begin{cases} a_k \lambda_k e_k, & k \leqslant n, \\ 0, & k > n. \end{cases}$$

Thus,

$$\operatorname{tr}(A_n Q_t A_n^*) = \sum_{k=1}^{n} \langle e_k, A_n Q_t A_n^* e_k \rangle = \sum_{k=1}^{n} a_k \lambda_k \leqslant \sum_{k=1}^{\infty} a_k \lambda_k < \infty,$$

and taking the supremum over n yields

$$\sup_n \operatorname{tr}(A_n Q_t A_n^*) \leqslant \sum_{k=1}^{\infty} a_k \lambda_k < \infty.$$

This proves the claim and completes the proof. □

Lemma 3.17. *Let μ be an invariant measure for the linear SPDE* (3.1)*, and set*

$$\mu_t = S(t)^{\#} \mu.$$

Then the family of measures $\{\mu_t\}_{t>0}$ is tight.

Proof. Fix $t > 0$. Let x_0 be a random variable with distribution μ. By invariance of μ, the law of $x(t)$ is also μ. By the mild formulation,

$$x(t) = S(t) x_0 + \int_0^t S(t-s) Q \, dW(s).$$

Let X, Y, and Z be random variables with distributions μ_t, ν_t, and μ, respectively, where ν_t denotes the law of

$$\int_0^t S(t-s) Q \, dW(s).$$

Then

$$Z \stackrel{d}{=} X + Y,$$

with X and Y independent.

Pick any $\varepsilon > 0$ and choose a compact set $\widehat{K}$ such that $\mu(\mathcal{H} \setminus \widehat{K}) < \varepsilon$. By Lemma 3.16, the family $\{\nu_t\}_{t>0}$ is tight, so there exists a compact set B_R such that for every $t > 0$,

$$\nu_t(\mathcal{H} \setminus B_R) < \varepsilon.$$

Define

$$K = \{z - y : \ z \in \widehat{K}, \, y \in B_R\}.$$

Then K is compact in $\mathcal{H}$. Since $\{Z \in \widehat{K}\} \cap \{Y \in B_R\} \subset \{X \in K\}$, we obtain

$$\mu_t(\mathcal{H} \setminus K) = \mathbf{P}(X \notin K) \leqslant \mathbf{P}(Z \notin \widehat{K}) + \mathbf{P}(Y \notin B_R) \leqslant 2\varepsilon.$$

This shows that the family $\{\mu_t\}_{t>0}$ is tight. □

The next theorem gives an exact characterisation of invariant measures for the linear SPDE (3.1).

Theorem 3.18. *Any invariant measure of the linear SPDE* (3.1) *is of the form* $\nu \star \mu_\infty$*, where* ν *is a measure on* $\mathcal{H}$ *that is invariant under the action of the semigroup* S*, and* μ_∞ *is the centred Gaussian measure with covariance* Q_∞*.*

Proof. We first prove that any measure of the form $\nu \star \mu_\infty$ is invariant. By Lemma 3.14, we have

$$\begin{aligned}
P_t^* \widehat{(\nu \star \mu_\infty)}(h) &= \widehat{\nu \star \mu_\infty}(S^*(t)h) \exp\left(-\frac{1}{2}\langle h, \, Q_t h\rangle\right) \\
&= \hat{\nu}(S^*(t)h)\hat{\mu}_\infty(S^*(t)h) \exp\left(-\frac{1}{2}\langle h, \, Q_t h\rangle\right) \\
&= \hat{\nu}(S^*(t)h) \, \hat{\mu}_\infty(h).
\end{aligned}$$

By the assumption that ν is invariant under $S(t)$,

$$\hat{\nu}(S^*(t)h) = \int_{\mathcal{H}} e^{i\langle S^*(t)h, \, x\rangle} \, \nu(dx) = \int_{\mathcal{H}} e^{i\langle h, \, S(t)x\rangle} \, \nu(dx) = \hat{\nu}(h).$$

Therefore,

$$P_t^* \widehat{(\nu \star \mu_\infty)}(h) = \hat{\nu}(h)\hat{\mu}_\infty(h) = \widehat{\nu \star \mu_\infty}(h),$$

which implies that $\nu \star \mu_\infty$ is an invariant measure.

Now assume that μ is an invariant measure for the linear SPDE (3.1) and define $\mu_t = S(t)^{\#}\mu$. Then

$$\widehat{\mu_t}(h) = \int_{\mathcal{H}} e^{i\langle h, \, x\rangle} \mu_t(dx) = \int_{\mathcal{H}} e^{i\langle h, \, S(t)x\rangle} \mu(dx) = \hat{\mu}(S^*(t)h).$$

By Lemma 3.14 again,

$$\widehat{\mu_t}(h) = \hat{\mu}(h) \exp\left(\frac{1}{2}\langle h,\, Q_t h\rangle\right).$$

By Lemma 3.17, the family $\{\mu_t\}_{t>0}$ is tight. By Prokhorov's theorem, along a subsequence we have $\mu_t \Rightarrow \nu$ for some probability measure ν. Since $Q_t \to Q_\infty$ in trace norm as $t \to \infty$, it follows that

$$\hat{\nu}(h) = \lim_{t\to\infty} \widehat{\mu_t}(h) = \hat{\mu}(h) \exp\left(\frac{1}{2}\langle h,\, Q_\infty h\rangle\right).$$

In particular,

$$\hat{\mu}(h) = \hat{\nu}(h) \exp\left(-\frac{1}{2}\langle h, Q_\infty h\rangle\right) = \hat{\nu}(h)\hat{\mu}_\infty(h) = \widehat{\nu \star \mu_\infty}(h),$$

so $\mu = \nu \star \mu_\infty$.

It remains to show that ν is invariant under the action of the semigroup $S(t)$. Note that

$$\mu_{t+s} = S(t)^{\#}\mu_s.$$

Taking characteristic functions on both sides yields

$$\widehat{\mu}_{t+s}(h) = \widehat{\mu_s}(S^*(t)h).$$

Letting $s \to \infty$ and using $\mu_s \Rightarrow \nu$, we obtain

$$\widehat{\nu}(h) = \widehat{\nu}(S^*(t)h) = \widehat{S(t)^{\#}\nu}(h),$$

so ν is invariant under $S(t)$. This completes the proof. □

The last corollary gives a sufficient condition for the uniqueness of invariant measures for the linear SPDE (3.1).

Corollary 3.19. *If* $\lim_{t\to\infty} \|S(t)x\| = 0$ *for every* $x \in \mathcal{H}$*, then the linear SPDE* (3.1) *can have at most one invariant measure. Furthermore, if an invariant measure* μ_∞ *exists in this situation, then one has* $\mathcal{P}_t^*\nu \to \mu_\infty$ *weakly for every probability measure* ν *on* $\mathcal{H}$*.*

Proof. By Theorem 3.18, any invariant measure μ has the form $\mu = \nu \star \mu_\infty$, where ν is invariant under the action of the semigroup $S(t)$. If we can show that the only measure invariant under $S(t)$ is the Dirac measure δ_0, then the first part of the corollary follows immediately.

Let ν be a measure invariant under $S(t)$ and pick any $\varphi \in \mathbf{B}_b(\mathcal{H})$. Then

$$\int_{\mathcal{H}} \varphi(x)\,\nu(dx) = \int_{\mathcal{H}} \varphi(x)\,(S(t)^{\#}\nu)(dx) = \int_{\mathcal{H}} \varphi(S(t)x)\,\nu(dx).$$

Taking $t \to \infty$ on both sides and using the dominated convergence theorem, we obtain

$$\int_{\mathcal{H}} \varphi(x)\,\nu(dx) = \int_{\mathcal{H}} \lim_{t\to\infty} \varphi(S(t)x)\,\nu(dx) = \varphi(0),$$

so $\nu = \delta_0$.

Now assume that an invariant measure μ_∞ exists and fix any probability measure ν on $\mathcal{H}$. Let μ_t be the centered Gaussian measure with covariance operator Q_t. Then μ_t is the law of

$$Y_t = \int_0^t S(t-s)Q\,dW(s).$$

Since $Q_t \to Q_\infty$ in trace norm as $t \to \infty$, we have that Y_t converges in $L^2(\Omega;\mathcal{H})$ to the random variable

$$Y_\infty = \int_0^\infty S(s)Q\,dW(s),$$

so $\mu_t \Rightarrow \mu_\infty$ as $t \to \infty$. By the same argument as above, we also have $S(t)^{\#}\nu \Rightarrow \delta_0$ as $t \to \infty$.

Moreover, $\mathcal{P}_t^*\nu$ is the law of $X_t + Y_t$, where X_t and Y_t are independent random variables with laws $S(t)^{\#}\nu$ and μ_t, respectively. Therefore,

$$\mathcal{P}_t^*\nu = (S(t)^{\#}\nu) \star \mu_t.$$

Letting $t \to \infty$ gives

$$\mathcal{P}_t^*\nu \Rightarrow \delta_0 \star \mu_\infty = \mu_\infty.$$

This completes the proof. $\square$

4 Semilinear SPDEs

In this section, we consider semilinear SPDEs of the form

$$dx = Lx\,dt + F(x)\,dt + Q\,dW(t), \quad x(0) = x_0 \in \mathcal{B}, \tag{4.1}$$

where $x(t)$ is a $\mathcal{B}$-valued stochastic process; L is the generator of a strongly continuous semigroup $S(t)$ on $\mathcal{B}$; F is a measurable map from some linear subspace $\mathcal{D}(F) \subset \mathcal{B}$ to $\mathcal{B}$; W is a cylindrical Wiener process on some Hilbert space $\mathcal{K}$; and $Q : \mathcal{K} \to \mathcal{B}$ is a bounded linear operator.

The solution of (4.1) is defined in the mild sense as follows:

Definition 4.1 (Mild solution)**.** A stochastic process $t \mapsto x(t) \in \mathcal{D}(F)$ is called a **mild solution** to (4.1) if the identity

$$x(t) = S(t)x_0 + \int_0^t S(t-s)F(x(s))\,ds + \int_0^t S(t-s)Q\,dW(s) \tag{4.2}$$

holds for every $t > 0$ almost surely.

Remark. For every $\lambda \in \mathbb{R}$, if we replace (L, F) by $(L - \lambda I, F + \lambda I)$, then the equation (4.1) does not change. Thus, the solution to (4.1) is invariant under this transformation.

4.1 Local solutions

For convenience, throughout this subsection, we assume that the stochastic integral

$$W_L(t) \stackrel{\text{def}}{=} \int_0^t S(t-s)Q\,dW(s)$$

has continuous sample paths in $\mathcal{B}$. It is possible that the solution to the SPDE (4.1) blows up in finite time. In the case of SPDEs, the situation can be worse: the blow-up time can even be random. To deal with these issues, we introduce the notion of local solutions.

Let us denote the natural filtration associated to the Wiener process W by

$$\{\mathcal{F}_t\}_{t\geqslant 0} = \{\sigma(W(s) - W(r) : 0 \leqslant r, s \leqslant t)\}_{t\geqslant 0}.$$

Let us recall the definition of a stopping time with respect to the natural filtration of W.

Definition 4.2 (Stopping time)**.** A positive random variable $\tau : \Omega \to (0, \infty]$ is called a **stopping time** with respect to $\{\mathcal{F}_t\}$ if for every fixed $T > 0$, the event $\{\tau \leqslant T\}$ is $\mathcal{F}_T$-measurable.

Definition 4.3 (Local mild solution)**.** A $\mathcal{D}(F)$-valued stochastic process $x(t)$ together with a stopping time τ such that $\tau > 0$ almost surely is called a **local mild solution** to (4.1) if the identity

$$x(t) = S(t)x_0 + \int_0^t S(t-s)F(x(s))\,ds + \int_0^t S(t-s)Q\,dW(s) \tag{4.3}$$

holds for every stopping time t such that $t < \tau$ almost surely.

It is possible that (4.1) has multiple local mild solutions with different stopping times. To deal with this issue, we introduce the notion of a maximal mild solution.

Definition 4.4. A local mild solution (x, τ) to (4.1) is called a **maximal mild solution** if for any other local mild solution $(\tilde{x}, \tilde{\tau})$, we have $\tilde{\tau} \leqslant \tau$ almost surely.

We will need Banach's fixed point theorem to prove the existence and uniqueness of local mild solutions to (4.1).

Theorem 4.5 (Banach fixed point theorem)**.** *Let (X, d) be a complete metric space and let $T : X \to X$ be a contraction map, i.e. there exists a constant $0 < c < 1$ such that*

$$d(Tx, Ty) \leqslant c\, d(x, y)$$

for every $x, y \in X$. Then there exists a unique fixed point $x^* \in X$ such that $Tx^* = x^*$.

Our first result on the existence and uniqueness of mild solutions to (4.1) makes the rather restrictive assumption that the nonlinearity F is defined on the whole space $\mathcal{B}$ and that it is locally Lipschitz there:

Theorem 4.6. *Consider* (4.1) *on a Banach space $\mathcal{B}$ and assume that W_L is a continuous $\mathcal{B}$-valued process. Assume furthermore that $F : \mathcal{B} \to \mathcal{B}$ is such that its restriction to every bounded set is Lipschitz continuous. Then there exists a unique maximal mild solution (x, τ) to* (4.1)*. Furthermore, this solution has continuous sample paths and one has $\lim_{t\uparrow\tau} \|x(t)\| = \infty$ almost surely on the set $\{\tau < \infty\}$.*

If F is globally Lipschitz continuous, then $\tau = \infty$ almost surely.

Proof. Without loss of generality, we may assume that the semigroup satisfies

$$\|S(t)\| \leqslant M \quad \text{for all } t \geqslant 0.$$

We will use the Banach fixed point theorem to prove local existence of mild solutions. Given a fixed terminal time $T > 0$ and a continuous function $g : \mathbb{R}_+ \to \mathcal{B}$, define the map $M_{g,T} : \mathcal{C}([0,T], \mathcal{B}) \to \mathcal{C}([0,T], \mathcal{B})$ by

$$(M_{g,T}u)(t) = \int_0^t S(t-s)F(u(s))\, ds + g(t).$$

Fix $R > 0$. We claim that for sufficiently small $T > 0$, the map $M_{g,T}$ is a contraction on the closed ball

$$B_{R,T}(g) = \left\{ u \in \mathcal{C}([0,T], \mathcal{B}) : \sup_{t \in [0,T]} \|u(t) - g(t)\| \leqslant R \right\}.$$

First note that for any $u \in B_{R,T}(g)$ and any $t \in [0,T]$, we have

$$\|u(t)\| \leqslant \sup_{t \in [0,T]} \|g(t)\|_t cp + R.$$

Let $L_{R'}$ be the Lipschitz constant of F on the ball of radius $R' = \sup_{t \in [0,T]} \|g(t)\| + R$. For any $u, v \in B_{R,T}(g)$, one has

$$\begin{aligned} \|M_{g,T}u(t) - M_{g,T}v(t)\| &\leqslant \int_0^t \|S(t-s)\| \, \|F(u(s)) - F(v(s))\| \, ds \\ &\leqslant M \int_0^t \|F(u(s)) - F(v(s))\| \, ds, \end{aligned}$$

so taking the supremum over $t \in [0,T]$ yields

$$\begin{aligned} \sup_{t \in [0,T]} \|M_{g,T}u(t) - M_{g,T}v(t)\| &\leqslant MT \sup_{t \in [0,T]} \|F(u(t)) - F(v(t))\| \\ &\leqslant MTL_{R'} \sup_{t \in [0,T]} \|u(t) - v(t)\|. \end{aligned}$$

Similarly, one has

$$\begin{aligned}\sup_{t\in[0,T]}\|M_{g,T}u(t)-g(t)\| &\leqslant \sup_{t\in[0,T]}\int_0^t \|S(t-s)\|\,\|F(u(s))\|\,ds\\ &\leqslant MT\sup_{t\in[0,T]}\|F(u(t))\|\\ &\leqslant MT\left(L_{R'}R'+\|F(0)\|\right).\end{aligned}$$

Now pick $T>0$ small enough such that

$$\begin{aligned}MTL_{R'} &= \frac{1}{2},\\ MT\left(L_{R'}R'+\|F(0)\|\right) &\leqslant R.\end{aligned}$$

Then $M_{g,T}$ is a contraction map on $B_{R,T}(g)$, so by Banach's fixed point theorem, there exists a unique fixed point $x^*\in\mathcal{C}([0,T],\mathcal{B})$ such that $M_{g,T}x^*=x^*$. Now take

$$g(t)=S(t)x_0+W_L(t).$$

Then

$$x^*(t)=\int_0^t S(t-s)F(x^*(s))\,ds+S(t)x_0+W_L(t),$$

which means (x^*,T) is a local mild solution to (4.1).

To construct a maximal mild solution, we iterate the local existence argument. First, using the fixed-point map $M_{g,T}$, we construct a mild solution on $[0,\tau_1)$. Then we use $x(\tau_1)$ as a new initial condition and apply the same construction on $[\tau_1,\tau_1+\tau_2)$. This procedure can be continued as long as the solution remains bounded.

If the solution blows up at some time $\tau(\omega)$, then $\|x(t,\omega)\|$ becomes unbounded as $t\uparrow\tau(\omega)$. In particular, the solution cannot be contained in any fixed bounded ball near $\tau(\omega)$, and we can no longer apply the local fixed-point argument to extend it beyond $\tau(\omega)$.

On the other hand, if F is globally Lipschitz continuous, then $M_{g,T}$ is a contraction on $\mathcal{C}([0,T],\mathcal{B})$ for all sufficiently small T, independently of the initial condition. In this case, we can repeat the above argument on consecutive time intervals to extend the solution for all times, so the maximal lifetime satisfies $\tau=\infty$ almost surely. $\square$

By assuming that L generates an analytic semigroup on $\mathcal{B}$ and that F has better regularity properties, we can obtain solutions taking values in more regular spaces.

Theorem 4.7. *Let L generate an analytic semigroup on $\mathcal{B}$ (denote by $\mathcal{B}_\alpha$, $\alpha \in \mathbf{R}$, the corresponding interpolation spaces) and assume that Q is such that the stochastic convolution W_L has almost surely continuous sample paths in $\mathcal{B}_\alpha$ for some $\alpha \geqslant 0$. Assume furthermore that there exist $\gamma \geqslant 0$ and $\delta \in [0,1)$ such that, for every $\beta \in [0,\gamma]$, the map F extends to a locally Lipschitz continuous map from $\mathcal{B}_\beta$ to $\mathcal{B}_{\beta-\delta}$ that, together with its local Lipschitz constant, grows at most polynomially.*

Then (4.1) *has a unique maximal mild solution (x,τ) with x taking values in $\mathcal{B}_\beta$ for every $\beta < \beta_\star \stackrel{def}{=} \alpha \wedge (\gamma + 1 - \delta)$.*

Proof. To prove existence and uniqueness of a mild solution, we apply the same strategy as in the previous theorem. The only difference is in the estimates involving the semigroup $S(t)$.

$$\begin{aligned}\|M_{g,T}u(t) - M_{g,T}v(t)\| &\leqslant \int_0^t \|S(t-s)\big(F(u(s)) - F(v(s))\big)\|\,ds \\ &\leqslant C\int_0^t (t-s)^{-\delta}\,\|F(u(s)) - F(v(s))\|_{-\delta}\,ds \\ &\leqslant CT^{1-\delta}\sup_{t\in[0,T]}\|F(u(t)) - F(v(t))\|_{-\delta}.\end{aligned}$$

Taking the supremum over $t \in [0,T]$ on both sides, we obtain

$$\sup_{t\in[0,T]}\|M_{g,T}u(t) - M_{g,T}v(t)\| \leqslant CT^{1-\delta}\sup_{t\in[0,T]}\|F(u(t)) - F(v(t))\|_{-\delta}.$$

Similarly, one has

$$\begin{aligned}\sup_{t\in[0,T]}\|M_{g,T}u(t) - g(t)\| &\leqslant \sup_{t\in[0,T]}\int_0^t \|S(t-s)F(u(s))\|\,ds \\ &\leqslant C\int_0^T r^{-\delta}\,dr\,\sup_{t\in[0,T]}\|F(u(t))\|_{-\delta} \\ &\leqslant CT^{1-\delta}\sup_{t\in[0,T]}\|F(u(t))\|_{-\delta}.\end{aligned}$$

We then apply the local Lipschitz continuity of F again to obtain a unique maximal local mild solution.

In order to show that the solution $x(t)$ takes values in $\mathcal{B}_\beta$ for every $t < \tau$ and every $\beta < \beta_\star$, we apply a bootstrapping argument, i.e. an induction on β.

For convenience, for $a \in [0,1)$, we denote

$$W_L^a(t) = \int_{at}^t S(t-s)Q\,dW(s).$$

Then one has the identity

$$\begin{aligned} W_L(t) &= \int_0^{at} S(t-s)Q\,dW(s) + W_L^a(t) \\ &= \int_0^{at} S\big(t(1-a)+at-s\big)Q\,dW(s) + W_L^a(t) \\ &= S\big(t(1-a)\big)\int_0^{at} S(at-s)Q\,dW(s) + W_L^a(t) \\ &= S\big(t(1-a)\big)W_L(at) + W_L^a(t). \end{aligned}$$

Thus, if $W_L(t)$ is continuous with values in $\mathcal{B}_\alpha$, then $W_L^a(t)$ is also continuous with values in $\mathcal{B}_\alpha$.

To continue, we need the following claim. Fix $T > 0$. For every $\beta \in [0, \beta_\star)$, there exist exponents $p_\beta \geqslant 1$, $q_\beta \geqslant 0$ and constants $a \in (0,1)$, $C > 0$ such that the bound

$$\|x(t)\|_\beta \leqslant Ct^{-q_\beta}\left(1 + \sup_{s\in[at,t]}\|x(s)\| + \sup_{0\leqslant s\leqslant t}\|W_L^a(s)\|_\beta\right)^{p_\beta} \tag{4.4}$$

holds almost surely for all $t \in (0,T]$.

If we can prove the claim, then since $\sup_{s\in[at,t]}\|x(s)\| < \infty$ almost surely for every $t < \tau$, and $\sup_{0\leqslant s\leqslant t}\|W_L^a(s)\|_\beta < \infty$ almost surely for every $\beta < \alpha$, the proof is complete.

When $\beta = 0$, we can take $p_0 = 1$, $q_0 = 0$, $C = 1$ and any $a \in (0,1)$, and we have

$$\|x(t)\| \leqslant 1 + \sup_{s\in[at,t]}\|x(s)\| + \sup_{0\leqslant s\leqslant t}\|W_L^a(s)\|.$$

Now we perform the induction step. Assume that the claim is true for some

$\beta = \beta_0 \in [0, \gamma]$. We show that it is also true for $\beta = \beta_0 + \varepsilon$ for any $\varepsilon \in (0, 1-\delta)$, provided we adjust the constants appearing in the bound (4.4) accordingly. Then the induction implies that the bound holds for all $\beta < \gamma + 1 - \delta$, and the claim follows.

By the mild formulation, it is easy to check that

$$x(t) = S\big(t(1-a)\big)x(at) + \int_{at}^{t} S(t-s)F(x(s))\,ds + W_L^a(t).$$

We want to bound $\|x(t)\|_{\beta_0+\varepsilon}$. For the first term, by the smoothing property of analytic semigroups, we have

$$\|S\big(t(1-a)\big)x(at)\|_{\beta_0+\varepsilon} \leqslant C\big(t(1-a)\big)^{-\varepsilon}\|x(at)\|_{\beta_0} \leqslant Ct^{-\varepsilon}\|x(at)\|_{\beta_0},$$

where the constant C depends on ε and a. For the second term, we use the growth assumption on F to obtain a constant $C > 0$ and an integer $n \geqslant 1$ such that

$$\begin{aligned}
\left\|\int_{at}^{t} S(t-s)F(x(s))\,ds\right\|_{\beta_0+\varepsilon} &\leqslant \int_{at}^{t} \|S(t-s)F(x(s))\|_{\beta_0+\varepsilon}\,ds \\
&\leqslant C\int_{at}^{t} (t-s)^{-(\delta+\varepsilon)}\|F(x(s))\|_{\beta_0-\delta}\,ds \\
&\leqslant C\int_{at}^{t} (t-s)^{-(\delta+\varepsilon)}\left(1 + \|x(s)\|_{\beta_0}^{n}\right)\,ds \\
&\leqslant Ct^{1-\delta-\varepsilon}\left(1 + \sup_{s\in[at,t]}\|x(s)\|_{\beta_0}^{n}\right) \\
&\leqslant Ct^{-\varepsilon}\left(1 + \sup_{s\in[at,t]}\|x(s)\|_{\beta_0}^{n}\right).
\end{aligned}$$

Combining the bounds, we obtain

$$\|x(t)\|_{\beta_0+\varepsilon} \leqslant Ct^{-\varepsilon}\left(1 + \sup_{s\in[at,t]}\|x(s)\|_{\beta_0}^{n}\right) + \|W_L^a(t)\|_{\beta_0+\varepsilon}.$$

Now, by the induction hypothesis,

$$1+\|x(s)\|_{\beta_0}^n \leqslant 1+\left(Cs^{-q_{\beta_0}}\left(1+\sup_{r\in[as,s]}\|x(r)\|+\sup_{0\leqslant r\leqslant s}\|W_L^a(r)\|_{\beta_0}\right)^{p_{\beta_0}}\right)^n$$
$$\leqslant C's^{-nq_{\beta_0}}\left(1+\sup_{r\in[as,s]}\|x(r)\|+\sup_{0\leqslant r\leqslant s}\|W_L^a(r)\|_{\beta_0}\right)^{np_{\beta_0}}.$$

Plugging this bound into the previous inequality, we obtain

$$\|x(t)\|_{\beta_0+\varepsilon} \leqslant Ct^{-\varepsilon-nq_{\beta_0}}\left(1+\sup_{r\in[a^2t,t]}\|x(r)\|+\sup_{0\leqslant r\leqslant t}\|W_L^a(r)\|_{\beta_0}\right)^{np_{\beta_0}}+\|W_L^a(t)\|_{\beta_0+\varepsilon}.$$

Since for every $\beta<\beta_\star\leqslant\alpha$, the assumption implies that $W_L^a(t)$ belongs to $\mathcal{B}_\beta$ continuously, there exists a constant $C>0$ such that

$$\|W_L^a(t)\|_{\beta_0}\leqslant C\|W_L^a(t)\|_{\beta_0+\varepsilon}.$$

Adjusting the constant C accordingly, we obtain

$$\|x(t)\|_{\beta_0+\varepsilon} \leqslant Ct^{-\varepsilon-nq_{\beta_0}}\left(1+\sup_{r\in[a^2t,t]}\|x(r)\|+\sup_{0\leqslant r\leqslant t}\|W_L^a(r)\|_{\beta_0+\varepsilon}\right)^{np_{\beta_0}}.$$

By setting

$$p_{\beta_0+\varepsilon}=np_{\beta_0},\quad q_{\beta_0+\varepsilon}=\varepsilon+nq_{\beta_0},$$

and replacing a by a^2, we complete the induction step. Thus, the claim holds for all $\beta<\beta_\star$, which concludes the proof. □

4.2 Sobelev embedding theorem

Sobolev spaces play an important role in the study of semilinear SPDEs, since they are interpolation spaces associated with many differential operators. In this subsection, we state and prove some important properties of these spaces. We conclude this section with the Sobolev embedding theorem, which states that Sobolev spaces are continuously embedded into certain L^p spaces and spaces of continuous functions.

To be specific, we work on the d-dimensional torus

$$\mathbb{T}^d=(\mathbb{R}/2\pi\mathbb{Z})^d=[0,2\pi]^d.$$

Let $L^2(\mathbb{T}^d)$ be the space of square-integrable functions on $\mathbb{T}^d$ with inner product

$$\langle u,\, v\rangle = \int_{\mathbb{T}^d} u(x)\overline{v(x)}\, dx.$$

This is a Hilbert space with orthonormal basis given by

$$e_k(x) = \frac{1}{(2\pi)^{d/2}} e^{i\langle k,x\rangle}, \quad k \in \mathbb{Z}^d.$$

For every $u \in L^2(\mathbb{T}^d)$, we can write

$$u(x) = \sum_{k\in\mathbb{Z}^d} u_k e_k(x),$$

where $u_k = \langle u,\, e_k\rangle$.We also have the Plancherel identity

$$\|u\|^2_{L^2(\mathbb{T}^d)} = \sum_{k\in\mathbb{Z}^d} |u_k|^2.$$

The fractional Sobolev spaces are defined as follows.

Definition 4.8 (Fractional Sobolev spaces on $\mathbb{T}^d$)**.** For $s \geqslant 0$, the **fractional Sobolev space** $H^s(\mathbb{T}^d)$ is the subspace of functions $u \in L^2(\mathbb{T}^d)$ such that

$$\|u\|^2_{H^s(\mathbb{T}^d)} := \sum_{k\in\mathbb{Z}^d} (1+|k|^2)^s |u_k|^2 < \infty.$$

For $s < 0$, we define $H^s(\mathbb{T}^d)$ as the completion of $L^2(\mathbb{T}^d)$ under the norm above.

Theorem 4.9. *First, for every $s, t \geqslant 0$ and every $u \in H^{s+t}$, we have the identity*

$$\|u\|_{H^{s+t}} = \|(1-\Delta)^{s/2} u\|_{H^t}.$$

In particular, for every $s \geqslant 0$ and every $u \in H^s(\mathbb{T}^d)$, we have

$$\|u\|_{H^s(\mathbb{T}^d)} = \|(1-\Delta)^{s/2} u\|_{L^2(\mathbb{T}^d)}.$$

Proof. This follows from the fact that

$$\begin{aligned}
\|(1-\Delta)^{s/2}u\|_{H^t}^2 &= \left\| \sum_k (1+|k|^2)^{s/2} u_k e_k(x) \right\|_{H^t}^2 \\
&= \sum_k (1+|k|^2)^{s+t} |u_k|^2 \\
&= \|u\|_{H^{s+t}}^2 .
\end{aligned}$$

Taking square roots on both sides completes the proof. □

Now, let us apply Hölder's inequality to obtain some useful results.

Theorem 4.10. *Let A be a self-adjoint, positive definite operator on a Hilbert space $\mathcal{H}$. Then for every $\alpha \in [0,1]$, we have*

$$\|A^\alpha u\| \leqslant \|Au\|^\alpha \|u\|^{1-\alpha},$$

for every $u \in \mathcal{D}(A^\alpha) \subset \mathcal{H}$.

Proof. The cases $\alpha = 0$ and $\alpha = 1$ are trivial. Now consider $\alpha \in (0,1)$. By the spectral theorem, we may assume that $\mathcal{H} = L^2(M,\mu)$ for some measure space (M,μ) and that

$$(Au)(x) = \varphi(x)u(x),$$

for some non-negative measurable function $\varphi : M \to [0,\infty)$. Thus, for every $u \in \mathcal{D}(A^\alpha)$, we have

$$\|A^\alpha u\|^2 = \int_M \varphi(x)^{2\alpha} |u(x)|^2 \, d\mu(x) = \int_M |\varphi(x)u(x)|^{2\alpha} |u(x)|^{2(1-\alpha)} \, d\mu(x).$$

Applying Hölder's inequality with $p = \frac{1}{\alpha}$ and $q = \frac{1}{1-\alpha}$, we get

$$\begin{aligned}
\|A^\alpha u\|^2 &\leqslant \left(\int_M |\varphi(x)u(x)|^2 \, d\mu(x) \right)^\alpha \left(\int_M |u(x)|^2 \, d\mu(x) \right)^{1-\alpha} \\
&= \|\varphi u\|^{2\alpha} \|u\|^{2(1-\alpha)} \\
&= \|Au\|^{2\alpha} \|u\|^{2(1-\alpha)}.
\end{aligned}$$

Taking square roots on both sides completes the proof. □

Now, we apply Theorem 4.10 to the operator $A = (1-\Delta)^{\frac{t-s}{2}}$.

Corollary 4.11. *For any $t > s$ and $r \in [s,t]$, the bound*

$$\|u\|_{H^r}^{t-s} \leqslant \|u\|_{H^t}^{r-s} \|u\|_{H^s}^{t-r}$$

holds for every $u \in H^t(\mathbb{T}^d)$.

Proof. We prove that

$$\|u\|_{H^r} \leqslant \|u\|_{H^t}^{\frac{r-s}{t-s}} \|u\|_{H^s}^{\frac{t-r}{t-s}}.$$

Apply the Theorem 4.10 with

$$A = (1-\Delta)^{\frac{t-s}{2}}, \quad \alpha = \frac{r-s}{t-s}.$$

Then

$$\begin{aligned} \|u\|_{H^r} &= \|(1-\Delta)^{(r-s)/2} u\|_{H^s} \\ &= \|A^\alpha u\|_{H^s} \\ &\leqslant \|Au\|_{H^s}^\alpha \|u\|_{H^s}^{1-\alpha} \\ &= \|u\|_{H^t}^\alpha \|u\|_{H^s}^{1-\alpha}. \end{aligned}$$

This completes the proof. □

Next, we embed the fractional Sobolev spaces into the space of bounded functions.

Theorem 4.12. *For every $s > \frac{d}{2}$, we have the continuous embedding*

$$H^s(\mathbb{T}^d) \hookrightarrow L^\infty(\mathbb{T}^d).$$

In particular, there exists a constant $C > 0$ such that

$$\|u\|_{L^\infty(\mathbb{T}^d)} \leqslant C \|u\|_{H^s(\mathbb{T}^d)},$$

for every $u \in H^s(\mathbb{T}^d)$.

Proof. By the Fourier series representation and the triangle inequality,

$$\|u\|_{L^\infty(\mathbb{T}^d)} \leqslant \sum_{k \in \mathbb{Z}^d} |u_k|.$$

Applying the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}\sum_{k\in\mathbb{Z}^d}|u_k| &\leqslant \left(\sum_{k\in\mathbb{Z}^d}(1+|k|^2)^{-s}\right)^{1/2}\left(\sum_{k\in\mathbb{Z}^d}(1+|k|^2)^{s}|u_k|^2\right)^{1/2}\\ &= \left(\sum_{k\in\mathbb{Z}^d}(1+|k|^2)^{-s}\right)^{1/2}\|u\|_{H^s(\mathbb{T}^d)}.\end{aligned}$$

Since $s>\dfrac{d}{2}$, the series $\sum_{k\in\mathbb{Z}^d}(1+|k|^2)^{-s}$ converges, which yields the desired bound. □

Now, we state the general Sobolev embeddings.

Theorem 4.13 (Sobolev embeddings)**.** *Let $p\in[2,\infty]$. Then, for every $s>\dfrac{d}{2}-\dfrac{d}{p}$, the space $H^s\left(\mathbf{T}^d\right)$ is contained in $L^p\left(\mathbf{T}^d\right)$ and there exists a constant C such that $\|u\|_{L^p}\leqslant C\|u\|_{H^s}$.*

Proof. The case $p=\infty$ was proved above. Now consider $p=2$. Recall that $H^0(\mathbb{T}^d)=L^2(\mathbb{T}^d)$, so the statement is trivial. Now, consider $p\in(2,\infty)$. Given $u\in H^s(\mathbb{T}^d)$ with $s>\dfrac{d}{2}-\dfrac{d}{p}$, we divide the Fourier modes into blocks of dyadic size. More concretely, let $u^{(-1)}=u_0$ and for every $n\geqslant 0$, let

$$u^{(n)}(x)=\sum_{2^n\leqslant|k|<2^{n+1}}u_k e_k(x),$$

so that $u=\sum_{n\geqslant-1}u^{(n)}$. For every $f\in L^2(\mathbb{T}^d)$, we have the bound

$$\|f\|_{L^p}^p=\int_{\mathbb{T}^d}|f(x)|^p\,dx\leqslant\|f\|_{L^\infty}^{p-2}\|f\|_{L^2}^2.$$

Therefore, for every $n\geqslant 0$, we obtain

$$\|u^{(n)}\|_{L^p}^p\leqslant\|u^{(n)}\|_{L^\infty}^{p-2}\|u^{(n)}\|_{L^2}^2.$$

By the definition of $u^{(n)}$, we have

$$\begin{aligned}
\|u^{(n)}\|_{H^s}^2 &= \sum_{2^n \leqslant |k| < 2^{n+1}} (1+|k|^2)^s |u_k|^2 \\
&\geqslant C 2^{2ns} \sum_{2^n \leqslant |k| < 2^{n+1}} |u_k|^2 \\
&= C 2^{2ns} \|u^{(n)}\|_{L^2}^2,
\end{aligned}$$

for some constant $C > 0$. Thus,

$$\|u^{(n)}\|_{L^2}^2 \leqslant C 2^{-2ns} \|u^{(n)}\|_{H^s}^2.$$

Now, let $s' = \frac{d}{2} + \varepsilon$ for some $\varepsilon > 0$. By the Sobolev embedding for $p = \infty$, we have

$$\begin{aligned}
\|u^{(n)}\|_{L^\infty} &\leqslant C \|u^{(n)}\|_{H^{s'}} \\
&= C \left(\sum_{2^n \leqslant |k| < 2^{n+1}} (1+|k|^2)^{s'} |u_k|^2 \right)^{1/2} \\
&\leqslant C 2^{n(s'-s)} \left(\sum_{2^n \leqslant |k| < 2^{n+1}} (1+|k|^2)^{s} |u_k|^2 \right)^{1/2} \\
&= C 2^{n(s'-s)} \|u^{(n)}\|_{H^s}.
\end{aligned}$$

Combining the bounds above, we get

$$\|u^{(n)}\|_{L^p} \leqslant C \|u^{(n)}\|_{L^\infty}^{\frac{p-2}{p}} \|u^{(n)}\|_{L^2}^{\frac{2}{p}} \leqslant C 2^{n\left((s'-s)\frac{p-2}{p} - \frac{2s}{p}\right)} \|u^{(n)}\|_{H^s}.$$

Now, note that the exponent of 2^n equals

$$\begin{aligned}
E = (s'-s)\frac{p-2}{p} - \frac{2s}{p} &= \left(\frac{d}{2} + \varepsilon - s\right)\frac{p-2}{p} - \frac{2s}{p} \\
= \frac{d}{2} - \frac{d}{p} - s + O(\varepsilon).
\end{aligned}$$

Since ε can be chosen arbitrarily small and $s > \frac{d}{2} - \frac{d}{p}$, we can choose ε small enough so that $E < 0$. Therefore,

$$\sum_{n \geqslant -1} \|u^{(n)}\|_{L^p} \leqslant C \sum_{n \geqslant -1} 2^{nE} \|u^{(n)}\|_{H^s}.$$

This yields $\|u\|_{L^p} \leqslant C\|u\|_{H^s}$ and completes the proof. $\square$

Finally, we show that fractional Sobolev spaces are included in spaces of Hölder continuous functions.

Theorem 4.14. *For every $s > \dfrac{d}{2}$, $H^s(\mathbb{T}^d)$ is included in the space $\mathcal{C}^\alpha(\mathbb{T}^d)$ of Hölder continuous functions for every exponent $\alpha < s - \dfrac{d}{2}$.*

Proof. We will show that for every $u \in H^s(\mathbb{T}^d)$, one has

$$\|u\|_{\mathcal{C}^\alpha(\mathbb{T}^d)} = \sup_{x \neq y} \frac{|u(x) - u(y)|}{|x-y|^\alpha} < \infty.$$

First, for every $x, y \in \mathbb{T}^d$, we have

$$|u(x) - u(y)| = \left| \sum_{k \in \mathbb{Z}^d} u_k (e_k(x) - e_k(y)) \right| \leqslant \sum_{k \in \mathbb{Z}^d} |u_k| \, |e_k(x) - e_k(y)|.$$

By the mean value theorem, for any $\theta, \varphi \in \mathbb{R}$,

$$|e^{i\theta} - e^{i\varphi}| \leqslant |\theta - \varphi|.$$

Thus,

$$|e_k(x) - e_k(y)| \leqslant \min\bigl(2, |\langle k, x\rangle - \langle k, y\rangle|\bigr) \leqslant \min\bigl(2, |k| \, |x-y|\bigr).$$

For any $r \geqslant 0$ and $\alpha \in (0,1]$, since we have $\min(2,r) \leqslant 2r^\alpha$, so

$$|u(x) - u(y)| \leqslant \sum_{k \in \mathbb{Z}^d} |u_k| \min\bigl(2, |k| \, |x-y|\bigr) \leqslant 2|x-y|^\alpha \sum_{k \in \mathbb{Z}^d} |u_k| \, |k|^\alpha.$$

It therefore suffices to show that $\displaystyle\sum_{k \in \mathbb{Z}^d} |u_k| \, |k|^\alpha < \infty$.

By the Cauchy–Schwarz inequality,

$$\begin{aligned}\sum_{k\in\mathbb{Z}^d} |u_k|\,|k|^\alpha &\leqslant \left(\sum_{k\in\mathbb{Z}^d}(1+|k|^2)^s|u_k|^2\right)^{1/2}\left(\sum_{k\in\mathbb{Z}^d}(1+|k|^2)^{-s}|k|^{2\alpha}\right)^{1/2}\\ &= \|u\|_{H^s(\mathbb{T}^d)}\left(\sum_{k\in\mathbb{Z}^d}(1+|k|^2)^{-s}|k|^{2\alpha}\right)^{1/2}.\end{aligned}$$

Since $2s-2\alpha>d$, the last series converges. This completes the proof. □

References

[Hairer(2009)] Martin Hairer. An introduction to stochastic pdes. Lecture notes, 2009. Available online.