An Introduction to Stochastic PDEs

January 26, 2009

Martin Hairer

The University of Warwick / Courant Institute

Contents

1 Introduction

These notes are based on a series of lectures given first at the University of Warwick in spring 2008 and then at the Courant Institute in spring 2009. It is an attempt to give a reasonably self-contained presentation of the basic theory of stochastic partial differential equations, taking for granted basic measure theory, functional analysis and probability theory, but nothing else. Since the aim was to present most of the material covered in these notes during a 30-hours series of postgraduate lecture, such an attempt is doomed to failure unless drastic choices are made. This is why many important facets of the theory of stochastic PDEs are missing from these notes. In particular, we do *not* treat equations with multiplicative noise, we do *not* treat equations driven Lévy noise, we do *not* consider equations with 'rough' (that is not locally Lipschitz, even in a suitable space) nonlinearities, we do *not* treat measure-valued processes, we do *not* consider hyperbolic or elliptic problems, we do *not* cover Malliavin calculus and densities of solutions, etc. The reader who is interested in a more detailed exposition of these more technically subtle parts of the theory might be advised to read the excellent works [DPZ92b, DPZ96, PZ07, PR07, SS05].

Instead, the approach taken in these notes is to focus on semilinear *parabolic* problems driven by *additive* noise. These can be treated as stochastic evolution equations in some infinite-dimensional Banach or Hilbert space that usually have nice regularising properties and they already form (in my humble opinion) a very rich class of problems with many interesting properties. Furthermore, this class of problems has the advantage of allowing to completely pass under silence many subtle problems arising from stochastic integration in infinite-dimensional spaces.

1.1 Acknowledgements

These notes would never have been completed, were it not for the enthusiasm of the attendants of the course. Hundreds of typos and mistakes were spotted and corrected. I am particularly indebted to David Epstein and Jochen Voß who carefully worked their way through these notes when they were still in a state of wilderness. Special thanks are also due to Pavel Bubak who was running the tutorials for the course given in Warwick.

2 Some motivating examples

2.1 A model for a random string (polymer)

Take $N+1$ particles with positions u_n immersed in a fluid and assume that nearest-neighbours are connected by harmonic springs. If the particles are furthermore subject to an external forcing F, the equations of motion (in the overdamped regime where the forces acting on the particle are more important than inertia, which can also formally be seen as the limit where the masses of the particles go to zero) would be given by

$$\begin{aligned}
\frac{du_0}{dt} &= k(u_1 - u_0) + F(u_0)\;, \\
\frac{du_n}{dt} &= k(u_{n+1} + u_{n-1} - 2u_n) + F(u_n)\;, \quad n = 1, \ldots, N-1\;, \\
\frac{du_N}{dt} &= k(u_{N-1} - u_N) + F(u_N)\;.
\end{aligned}$$

This is a primitive model for a polymer chain consisting of $N+1$ monomers and without self-interaction. It does however not take into account the effect of the molecules of water that would randomly 'kick' the particles that make up our string. Assuming that these kicks occur randomly and independently at high rate, this effect can be modelled in first instance by independent white noises acting on all degrees of freedom of our model. We thus obtain a system of coupled stochastic differential equations:

$$\begin{aligned}
du_0 &= k(u_1 - u_0)\,dt + F(u_0)\,dt + \sigma\,dw_0(t)\;, \\
du_n &= k(u_{n+1} + u_{n-1} - 2u_n)\,dt + F(u_n)\,dt + \sigma\,dw_n(t)\;, \quad n = 1, \ldots, N-1\;, \\
du_N &= k(u_{N-1} - u_N)\,dt + F(u_N)\,dt + \sigma\,dw_N(t)\;.
\end{aligned}$$

Formally taking the continuum limit (with the scalings $k \approx \nu N^2$ and $\sigma \approx \sqrt{N}$), we can infer that if N is very large, this system is well-described by the solution to a stochastic *partial differential equation*

$$du(x,t) = \nu\partial_x^2 u(x,t)\,dt + F(u(x,t))\,dt + dW(x,t)\;,$$

endowed with the boundary conditions $\partial_x u(0,t) = \partial_x u(1,t) = 0$. It is not so clear *a priori* what the meaning of the term $dW(x,t)$ should be. We will see in the next section that, at least on a formal level, it is reasonable to assume that $\mathbf{E}\frac{dW(x,t)}{dt}\frac{dW(y,s)}{ds} = \delta(x-y)\delta(t-s)$. The precise meaning of this formula will be discussed later.

2.2 The stochastic Navier-Stokes equations

The Navier-Stokes equations describing the evolution of the velocity field $u(x,t)$ of an incompressible viscous fluid are given by

$$\frac{du}{dt} = \nu\Delta u - (u\cdot\nabla)u - \nabla p + f\;, \tag{2.1}$$

complemented with the (algebraic) incompressibility condition $\operatorname{div} u = 0$. Here, f denotes some external force acting on the fluid, whereas the pressure p is given implicitly by the requirement that $\operatorname{div} u = 0$ at all times.

While it is not too difficult in general to show that solutions to (2.1) exist in some weak sense, in the case where $x \in \mathbf{R}^d$ with $d \geq 3$, their *uniqueness* is an open problem with a \$1,000,000 prize. We will of course not attempt to solve this long-standing problem, so we are going to restrict ourselves to the case $d = 2$. (The case $d = 1$ makes no sense since there the condition $\operatorname{div} u = 0$ would imply that u is constant. However, one could also consider the Burger's equation which has similar features to the Navier-Stokes equations.)

For simplicity, we consider solutions that are periodic in space, so that we view u as a function from $\mathbf{T}^2 \times \mathbf{R}_+$ to $\mathbf{R}^2$. In the absence of external forcing f, one can use the incompressibility assumption to see that

$$\frac{d}{dt}\int_{\mathbf{T}^2} |u(x,t)|^2\,dx = -2\nu\int_{\mathbf{T}^2} \operatorname{tr} Du(x,t)^* Du(x,t)\,dx \leq -2\nu\int_{\mathbf{T}^2} |u(x,t)|^2\,dx\;,$$

where we used the Poincaré inequality in the last line (assuming that $\int_{\mathbf{T}^2} u(x,t)\,dx = 0$). Therefore, by Gronwall's inequality, the solutions decay to 0 exponentially fast. This shows that energy needs to be pumped into the system continuously if one wishes to maintain an interesting regime.

One way to achieve this from a mathematical point of view is to add a force f that is randomly fluctuating. We are going to show that if one takes a random force that is Gaussian and such that

$$\mathbf{E}f(x,t)f(y,s) = \delta(t-s)C(x-y)\;,$$

for some correlation function C then, provided that C is sufficiently regular, one can show that (2.1) has solutions for all times. Furthermore, these solutions do not blow up in the sense that one can find a constant K such that, for any solution to (2.1), one has

$$\limsup_{t\to\infty} \mathbf{E}\|u(t)\|^2 \leq K\;,$$

for some suitable norm $\|\cdot\|$. This allows to provide a construction of a model for homogeneous turbulence which is amenable to mathematical analysis.

2.3 The stochastic heat equation

In this section, we focus on the particular example of the stochastic heat equation. We will perform a number of calculations that give us a feeling for what the solutions to this equation look like. These calculations will not be completely rigorous but could be made so with some extra effort. Most tools required to make them rigorous will be introduced later in the course.

2.3.1 Setup

Recall that the *heat equation* is the partial differential equation:

$$\partial_t u = \Delta u \;, \quad u\colon \mathbf{R}_+ \times \mathbf{R}^n \to \mathbf{R} \;. \tag{2.2}$$

Given any bounded continuous initial condition $u_0\colon \mathbf{R}^n \to \mathbf{R}$, there exists a unique solution u to (2.2) which is continuous on $\mathbf{R}_+ \times \mathbf{R}^n$ and such that $u(0,x) = u_0(x)$ for every $x \in \mathbf{R}^n$.

This solution is given by the formula

$$u(t,x) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4t}} u_0(y)\, dy \;.$$

We will denote this by the shorthand $u(t,\cdot\,) = e^{\Delta t} u_0$ by analogy with the solution to an $\mathbf{R}^d$-valued linear equation of the type $\partial_t u = Au$.

Let us now go one level up in difficulty by considering (2.2) with an additional 'forcing term' f:

$$\partial_t u = \Delta u + f \;, \quad u\colon \mathbf{R}_+ \times \mathbf{R}^n \to \mathbf{R} \;. \tag{2.3}$$

From the variations of constants formula, we obtain that the solution to (2.3) is given by

$$u(t,\cdot\,) = e^{\Delta t} u_0 + \int_0^t e^{\Delta(t-s)} f(s,\cdot\,)\, ds \;. \tag{2.4}$$

Since the kernel defining $e^{\Delta t}$ is very smooth, this expression actually makes sense for a large class of distributions f. Suppose now that f is 'space-time white noise'. We do not define this rigorously for the moment, but characterise it as a (distribution-valued) centred Gaussian process ξ such that $\mathbf{E}\xi(s,x)\xi(t,y) = \delta(t-s)\delta(x-y)$.

The stochastic heat equation is then the stochastic partial differential equation

$$\partial_t u = \Delta u + \xi \;, \quad u\colon \mathbf{R}_+ \times \mathbf{R}^n \to \mathbf{R} \;. \tag{2.5}$$

Consider the simplest case $u_0 = 0$, so that its solution is given by

$$u(t,x) = \int_0^t \frac{1}{(4\pi|t-s|)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4(t-s)}} \xi(s,y)\, dy\, ds \tag{2.6}$$

This is again a centred Gaussian process, but its covariance function is more complicated. The aim of this section is to get some idea about the space-time regularity properties of (2.6). While the solutions to ordinary stochastic differential equations are in general α-Hölder continuous (in time) for every $\alpha < 1/2$ but not for $\alpha = 1/2$, we will see that in dimension $n = 1$, u as given by (2.6) is only 'almost' $1/4$-Hölder continuous in time and 'almost' $1/2$-Hölder continuous in space. In higher dimensions, it is not even function-valued... The reason for this lower time-regularity is that the driving space-time white noise is not only very singular as a function of time, but also as a function of space. Therefore, some of the regularising effect of the heat equation is required to turn it into a continuous function in space.

Heuristically, the appearance of the Hölder exponents $1/2$ for space and $1/4$ for time in dimension $n = 1$ can be understood by the following argument. If we were to remove the term $\partial_t u$ in (2.5), then u would have the same time-regularity as ξ, but two more derivatives of space regularity. If on the other hand we were to remove the term Δu, then u would have the sample space regularity as ξ, but one more derivative of time regularity. The consequence of keeping both terms is that we can 'trade' space-regularity against time-regularity at a cost of one time derivative for two space derivatives. Now we know that white noise (that is the centred Gaussian process η with $\mathbf{E}\eta(t)\eta(s) = \delta(t-s)$) is the time derivative of Brownian motion, which itself is 'almost' $1/2$-Hölder continuous. Therefore, the regularity of η requires 'a bit more than half a derivative' of improvement if we wish to obtain a continuous function.

Turning back to ξ, we see that it is expected to behave like η both in the space direction and in the time direction. So, in order to turn it into a continuous function of time, roughly half of a time derivative is required. This leaves over half of a time derivative, which we trade against one spatial derivative, thus concluding that for fixed time, u will be almost $1/2$-Hölder continuous in space. Concerning the time regularity, we note that half of a space derivative is required to turn ξ into a continuous function of space, thus leaving one and a half space derivative. These can be traded against $3/4$ of a time derivative, thus explaining the $1/4$-Hölder continuity in time.

In Section 5.1, we are going to see more precisely how the space-regularity and the time-regularity interplay in the solutions to linear SPDEs, thus allowing us to justify rigorously this type of heuristic arguments. For the moment, let us justify it by a calculation in the particular case of the stochastic heat equation.

2.3.2 A formal calculation

Define the covariance for the solution to the stochastic heat equation by

$$C(s,t,x,y) = \mathbf{E}u(s,x)u(t,y)\;, \tag{2.7}$$

where u is given by (2.6).

By (statistical) translation invariance, it is clear that $C(s,t,x,y) = C(s,t,0,x-y)$. Using (2.6) and the expression for the covariance of ξ, one has

$$\begin{aligned}
&C(s,t,0,x)\\
&\quad= \frac{1}{(4\pi)^n}\mathbf{E}\int_0^t\int_0^s\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}\frac{1}{|s-r'|^{n/2}|t-r|^{n/2}}e^{-\frac{|x-y|^2}{4(t-r)}-\frac{|y'|^2}{4(s-r')}}\xi(r,y)\xi(r',y')\,dy\,dy'\,dr'\,dr\\
&\quad= \frac{1}{(4\pi)^n}\int_0^{s\wedge t}\int_{\mathbf{R}^n}\frac{1}{|s-r|^{n/2}|t-r|^{n/2}}e^{-\frac{|x-y|^2}{4(t-r)}-\frac{|y|^2}{4(s-r)}}\,dy\,dr\\
&\quad= \frac{1}{(4\pi)^n}\int_0^{s\wedge t}\int_{\mathbf{R}^n}\frac{1}{|s-r|^{n/2}|t-r|^{n/2}}\\
&\qquad\quad\times\exp\Big(-\frac{|x|^2}{4(t-r)}-\frac{\langle x,y\rangle}{2(t-r)}-\frac{|y|^2}{4(s-r)}-\frac{|y|^2}{4(t-r)}\Big)\,dy\,dr\;.
\end{aligned}$$

The integral over y can be performed explicitly by 'completing the square' and one obtains

$$\begin{aligned}
C(s,t,0,x) &= 2^{-n}\int_0^{s\wedge t}(s+t-2r)^{-n/2}\exp\Big(-\frac{|x|^2}{4(s+t-2r)}\Big)\,dr\\
&= 2^{-(n+1)}\int_{|s-t|}^{s+t}\ell^{-n/2}\exp\Big(-\frac{|x|^2}{4\ell}\Big)\,d\ell\;.
\end{aligned}\tag{2.8}$$

We notice that the singularity at $\ell = 0$ is integrable if and only if $n < 2$, so that $C(t,t,0,0)$ is finite only in the one-dimensional case. We therefore limit ourselves to this case in the sequel.

Remark 2.1 Even though the random variable u defined by (2.6) is not function-valued in dimension 2, it is 'almost' the case since the singularity in (2.8) diverges only logarithmically. The stationary solution to (2.5) is called the *Gaussian free field* and has been the object of intense studies over the last few years, especially in dimension 2. Its interest stems from the fact that many of its features are conformally invariant (as a consequence of the conformal invariance of the Laplacian), thus linking probability theory to quantum field theory on one hand and to complex geometry on the other hand. The Gaussian free field also relates directly to the Schramm-Loewner evolutions (SLEs) for the study of which W. Werner was awarded the Fields medal in 2006, see [Law04, SS06]. For more information on the Gaussian free field, see for example [She07].

The regularity of u is determined by the behaviour of C near the 'diagonal' $s = t$, $x = y$. We first consider the time regularity. We therefore set $x = 0$ and compute

$$C(s,t,0,0) = \frac{1}{4}\int_{|s-t|}^{s+t} \ell^{-1/2}\,d\ell = \tfrac{1}{2}\big(|s+t|^{\frac{1}{2}} - |s-t|^{\frac{1}{2}}\big)\;.$$

This shows that, in the case $n = 1$ and for $s \approx t$, one has the asymptotic behaviour

$$\mathbf{E}|u(s,0) - u(t,0)|^2 \approx |t-s|^{\frac{1}{2}}\;.$$

Comparing this with the standard Brownian motion for which $\mathbf{E}|B(s) - B(t)|^2 = |t-s|$, we conclude that the process $t \mapsto u(t,x)$ is, for fixed x, almost surely α-Hölder continuous for any exponent $\alpha < 1/4$ but *not* for $\alpha = 1/4$. This argument is a special case of Kolmogorov's celebrated continuity test, of which we will see a version adapted to Gaussian measures in Section 3.1.

If, on the other hand, we fix $s = t$, we obtain (still in the case $n = 1$) via the change of variables $z = |x|^2/4\ell$, the expression

$$C(t,t,0,x) = \frac{|x|}{8}\int_{\frac{|x|^2}{8t}}^{\infty} z^{-\frac{3}{2}} e^{-z}\,dz\;.$$

Integrating by parts, we get

$$C(t,t,0,x) = \frac{\sqrt{t}}{4} e^{-\frac{|x|^2}{8t}} + \frac{|x|}{4}\int_{\frac{|x|^2}{8t}}^{\infty} z^{-\frac{1}{2}} e^{-z}\,dz\;,$$

So that to leading order we have for small values of x:

$$C(t,t,0,x) \approx \frac{\sqrt{t}}{4} + \frac{|x|}{4}\int_0^{\infty} z^{-\frac{1}{2}} e^{-z}\,dz = \sqrt{t} + \frac{\sqrt{\pi}|x|}{4} + \mathcal{O}(|x|^2/8\sqrt{t})\;.$$

This shows that, at any fixed instant t, the solution to (2.5) looks like a Brownian motion in space over lengthscales of order $t^{1/2}$. Note that over such a lengthscale the Brownian motion fluctuates by about $t^{1/4}$, which is precisely the order of magnitude of $\mathbf{E}|u(t,x)|$.

2.4 What have we learned?

1. At a 'hand-waving' level, we have forced our equation with a term that has a temporal evolution resembling white noise, so that one would naively expect its solutions to have a temporal regularity resembling Brownian motion. However, for any fixed location in space, the solution to the stochastic heat equation has a time-regularity which is only almost Hölder-$\frac{1}{4}$, as opposed to the almost Hölder-$\frac{1}{2}$ time-regularity of Brownian motion.

2. Unlike the solutions to an ordinary parabolic PDE, the solutions to a stochastic PDE tend to be spatially 'rough'. It is therefore not obvious *a priori* how the formal expression that we obtained is to be related to the original equation (2.5), since even for positive times, the map $x \mapsto u(t,x)$ is certainly not twice differentiable.

3. Even though the deterministic heat equation has the property that $e^{\Delta t}u \to 0$ for every $u \in L^2$, the solution to the stochastic heat equation has the property that $\mathbf{E}|u(x,t)|^2 \to \infty$ for fixed x as $t \to \infty$. This shows that in this particular case, the stochastic forcing term pumps energy into the system faster than the deterministic evolution can dissipate it.

Exercise 2.2 Perform the same calculation, but for the equation

$$\partial_t u = \Delta u - au + \xi\ , \quad u\colon \mathbf{R}_+ \times \mathbf{R} \to \mathbf{R}\ .$$

Show that as $t \to \infty$, the law of its solution converges to the law of an Ornstein-Uhlenbeck process (if the space variable is viewed as 'time'):

$$\lim_{t\to\infty} \mathbf{E}u(t,x)u(t,y) = Ce^{-c|x-y|}\ .$$

Compute the constants C and c as functions of the parameter a.

3 Gaussian Measure Theory

This section is devoted to the study of Gaussian measures on general Banach spaces. Throughout this section and throughout most of the remainder of these notes, we will denote by $\mathcal{B}$ an arbitrary separable Banach space. Recall that a space is separable if it contains a countable dense subset. This separability assumption turns out to be crucial for measures on $\mathcal{B}$ to behave in a non-pathological way. It turns out that this assumption can be circumvented by trickery in most natural situations where non-separable spaces arise, but we choose not to complicate our lives by considering overly general cases in these notes.

One additional assumption that would appear to be natural in the context of Gaussian measure theory is that $\mathcal{B}$ be reflexive (that is $\mathcal{B}^{**} = \mathcal{B}$). This is because the mean of a measure μ appears in general to be an element of $\mathcal{B}^{**}$ rather than of $\mathcal{B}$, since the natural way of defining the mean m of μ is to set $m(\ell) = \int_{\mathcal{B}} \ell(x)\,\mu(dx)$ for any $\ell \in \mathcal{B}^*$. This turns out not to be a problem, since the mean of a Gaussian measure on a separable Banach space $\mathcal{B}$ turns out to always be an element of $\mathcal{B}$ itself, see the monograph [Bog98]. However this result is not straightforward to prove, so we will take here the more pragmatic approach that whenever we consider Gaussian measures with non-zero mean, we simply consider the mean $m \in \mathcal{B}$ as given.

Before we proceed, let us just mention a few examples of Banach spaces. The spaces $L^p(\mathcal{M},\nu)$ (with $(\mathcal{M},\mu)$ a countably generated measure space like $\mathbf{R}^n$) for $p \in [1,\infty)$ are both reflexive and separable. However, both properties fail to hold for L^∞ spaces. The space of bounded continuous functions on a compact space is separable, but not reflexive. The space of bounded continuous functions from $\mathbf{R}^n$ to $\mathbf{R}$ is neither separable nor reflexive, but the space of continuous functions from $\mathbf{R}^n$ to $\mathbf{R}$ vanishing at infinity is separable. (The last two statements are still true if we replace $\mathbf{R}^n$ by any locally compact complete separable metric space.) Hilbert spaces are obviously reflexive since $\mathcal{H}^* = \mathcal{H}$ for every Hilbert space $\mathcal{H}$ by the Riesz representation theorem [Yos95]. There exist non-separable Hilbert spaces, but they have rather pathological properties and do not appear very often in practice.

We start with the definition of a Gaussian measure on a Banach space. Since there is no equivalent to Lebesgue measure in infinite dimensions (one could never expect it to be σ-additive), we cannot define it by prescribing the form of its density. However, it turns out that Gaussian measures on $\mathbf{R}^n$ can be characterised by prescribing that the projections of the measure onto any one-dimensional subspace of $\mathbf{R}^n$ are all Gaussian. This is a property that can readily be generalised to infinite-dimensional spaces:

Definition 3.1 A *Gaussian probability measure* μ on a Banach space $\mathcal{B}$ is a Borel measure such that $\ell^*\mu$ is a real Gaussian probability measure on $\mathbf{R}$ for every linear functional $\ell\colon \mathcal{B} \to \mathbf{R}$. (Dirac measures are considered to be Gaussian measures with zero covariance.) We call it *centred* if $\ell^*\mu$ is centred for every ℓ.

Remark 3.2 We used here the notation $f^*\mu$ for the push-forward of a measure μ under a map f. This is defined by $(f^*\mu)(A) = \mu(f^{-1}(A))$.

One first question that one may ask is whether this is a reasonable definition. After all, it only makes a statement about the one-dimensional projections of the measure μ, which itself lives on an infinite-dimensional space. However, this turns out to reasonable since, provided that $\mathcal{B}$ is separable, the finite-dimensional projections of any probability measure contain sufficiently information to characterise it:

Proposition 3.3 *Let $\mathcal{B}$ be a separable Banach space and let μ and ν be two probability Borel measures on $\mathcal{B}$. If $\ell^*\mu = \ell^*\nu$ for every $\ell \in \mathcal{B}^*$, then $\mu = \nu$.*

Proof. Denote by $\mathrm{Cyl}(\mathcal{B})$ the algebra of cylindrical sets on $\mathcal{B}$, that is $A \in \mathrm{Cyl}(\mathcal{B})$ if and only if there exists $n > 0$, a continuous linear map $T\colon \mathcal{B} \to \mathbf{R}^n$, and a Borel set $\tilde{A} \subset \mathbf{R}^n$ such that $A = T^{-1}\tilde{A}$. It follows from the assumption that $\mu(A) = \nu(A)$ for every $A \in \mathrm{Cyl}(\mathcal{B})$ and therefore, by a basic uniqueness result in measure theory (see Lemma II.4.6 in [RW94] for example), for every A in the σ-algebra $\mathcal{E}(\mathcal{B})$ generated by $\mathrm{Cyl}(\mathcal{B})$. It thus remains to show that $\mathcal{E}(\mathcal{B})$ coincides with the Borel σ-algebra of $\mathcal{B}$.

Since $\mathcal{B}$ is separable, every open set U can be written as a countable union of closed balls. (Fix any dense countable subset $\{x_n\}$ of $\mathcal{B}$ and check that one has for example $U = \bigcup_{x_n \in U} \bar{B}(x_n, r_n)$, where $r_n = \frac{1}{2}\sup\{r > 0 : \bar{B}(x_n, r) \subset U\}$ and $\bar{B}(x, r)$ denotes the closed ball of radius r centred at x.) Since $\mathcal{E}(\mathcal{B})$ is invariant under translations and dilations, it remains to check that $\bar{B}(0,1) \in \mathcal{E}(\mathcal{B})$. Let $\{x_n\}$ be a countable dense subset of $\{x \in \mathcal{B} : \|x\| = 1\}$ and let ℓ_n by any sequence in $\mathcal{B}^*$ such that $\|\ell_n\| = 1$ and $\ell_n(x_n) = 1$ (such elements exist by the Hahn-Banach extension theorem [Yos95]). Define now $K = \bigcap_{n\geq 0}\{x \in \mathcal{B} : |\ell_n(x)| \leq 1\}$. It is clear that $K \in \mathcal{E}(\mathcal{B})$, so that the proof is complete if we can show that $K = \bar{B}(0,1)$.

Since obviously $\bar{B}(0,1) \subset K$, it suffices to show that the reverse inclusion holds. Let $y \in \mathcal{B}$ with $\|y\| > 1$ be arbitrary and set $\hat{y} = y/\|y\|$. By the density of the x_n's, there exists a subsequence x_{k_n} such that $\|x_{k_n} - \hat{y}\| \leq \frac{1}{n}$, say, so that $\ell_{k_n}(\hat{y}) \geq 1 - \frac{1}{n}$. By linearity, this implies that $\ell_{k_n}(y) \geq \|y\|(1 - \frac{1}{n})$, so that there exists a sufficiently large n so that $\ell_{k_n}(y) > 1$. This shows that $y \notin K$ and we conclude that $K \subset \bar{B}(0,1)$ as required. □

From now on, we will mostly consider centred Gaussian measures, since one can always reduce oneself to the general case by a simple translation. Given a centred Gaussian measure μ, we define a map $C_\mu\colon \mathcal{B}^* \times \mathcal{B}^* \to \mathbf{R}$ by

$$C_\mu(\ell, \ell') = \int_{\mathcal{B}} \ell(x)\ell'(x)\,\mu(dx)\;. \tag{3.1}$$

Remark 3.4 In the case $\mathcal{B} = \mathbf{R}^n$, this is just the covariance matrix, provided that we perform the usual identification of $\mathbf{R}^n$ with its dual.

The map C_μ will be called the *Covariance operator* of μ. It follows immediately from the definitions that the operator C_μ is bilinear and definite positive. Furthermore, the Fourier transform of μ is given by

$$\hat{\mu}(\ell) \stackrel{\text{def}}{=} \int_{\mathcal{B}} e^{i\ell(x)}\,\mu(dx) = \exp(-\tfrac{1}{2}C_\mu(\ell,\ell))\;, \tag{3.2}$$

where $\ell \in \mathcal{B}^*$. This can be checked by using the explicit form of the one-dimensional Gaussian measure. Conversely, (3.2) characterises Gaussian measures in the sense that if μ is a measure such that there exists C_μ satisfying (3.2) for every $\ell \in \mathcal{B}^*$, then μ must be centred Gaussian. The reason why this is so is that two distinct probability measures necessarily have distinct Fourier transforms:

Proposition 3.5 *Let μ and ν be any two probability measures on a separable Banach space $\mathcal{B}$. If $\hat{\mu}(\ell) = \hat{\nu}(\ell)$ for every $\ell \in \mathcal{B}^*$, then $\mu = \nu$.*

Proof. In the particular case $\mathcal{B} = \mathbf{R}^n$, if φ is a smooth function with compact support, it follows from Fubini's theorem and the invertibility of the Fourier transform that one has the identity

$$\int_{\mathbf{R}^n} \varphi(x)\,\mu(dx) = \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} \hat{\varphi}(k)e^{-ikx}\,dk\,\mu(dx) = \frac{1}{(2\pi)^n}\int_{\mathbf{R}^n}\hat{\varphi}(k)\,\hat{\mu}(-k)\,dk\;,$$

so that, since bounded continuous functions can be approximated by smooth functions, μ is indeed determined by $\hat{\mu}$. The general case then follows immediately from Proposition 3.3. □

Remark 3.6 We could also have defined Gaussian measures by imposing that $\ell^*\mu$ is Gaussian for every bounded linear map $\ell\colon \mathcal{B} \to \mathbf{R}^n$ and every n. These two definitions are equivalent because measures on $\mathbf{R}^n$ are characterised by their Fourier transforms and these can be constructed from one-dimensional marginals.

As a simple consequence, we have the following trivial but useful property:

Proposition 3.7 *Let μ be a Gaussian measure on $\mathcal{B}$ and, for every $\varphi \in \mathbf{R}$, define the 'rotation' $R_\varphi\colon \mathcal{B}^2 \to \mathcal{B}^2$ by*

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

Then, one has $R_\varphi^*(\mu\otimes\mu) = \mu\otimes\mu$.

Proof. Since a measure is characterised by its Fourier transform [Bog98, Prop A.3.18], it suffices to check that $\widehat{\mu\otimes\mu}\circ R_\varphi = \widehat{\mu\otimes\mu}$, which is an easy exercise. □

3.1 A-priori bounds on Gaussian measures

We are going to show now that the operator C_μ has to also be bounded, as a straightforward consequence of the fact that $x \mapsto \|x\|^2$ is integrable. Actually, we are going to show much more, namely that there always exists a constant $\alpha > 0$ such that $\exp(\alpha\|x\|^2)$ is integrable! In other words, the norm of any Banach-space valued Gaussian random variable has Gaussian tails, just like in the finite-dimensional case. This is the content of a celebrated theorem:

Theorem 3.8 (Fernique, 1970) *Let μ be a centred Gaussian probability measure on a separable Banach space $\mathcal{B}$. Then, there exists $\alpha > 0$ such that $\int_{\mathcal{B}} \exp(\alpha\|x\|^2)\,\mu(dx) < \infty$.*

Proof. Note first that, from Proposition 3.7, one has for any two positive numbers t and τ the bound

$$\begin{aligned}\mu(\|x\| \le \tau)\,\mu(\|x\| > t) &= \int_{\|x\|\le\tau}\int_{\|y\|>t}\mu(dx)\,\mu(dy) = \int_{\|\frac{x-y}{\sqrt{2}}\|\le\tau}\int_{\|\frac{x+y}{\sqrt{2}}\|>t}\mu(dx)\,\mu(dy)\\ &\le \int_{\|x\|>\frac{t-\tau}{\sqrt{2}}}\int_{\|y\|>\frac{t-\tau}{\sqrt{2}}}\mu(dx)\,\mu(dy) = \mu\Big(\|x\| > \tfrac{t-\tau}{\sqrt{2}}\Big)^2 . \end{aligned}\tag{3.3}$$

In order to go from the first to the second line, we have used the fact that the triangle inequality implies

$$\min\{\|x\|,\|y\|\} \ge \tfrac{1}{2}(\|x+y\| - \|x-y\|) ,$$

so that $\|x+y\| > \sqrt{2}t$ and $\|x-y\| \le \sqrt{2}\tau$ do indeed imply that both $\|x\|$ and $\|y\|$ are greater than $\frac{t-\tau}{\sqrt{2}}$. Since $\|x\|$ is μ-almost surely finite, there exists some $\tau > 0$ such that $\mu(\|x\| \le \tau) \ge \frac{1}{2}$. Set now $t_0 = \tau$ and define t_n for $n > 0$ recursively by the relation $t_n = \frac{t_{n+1}-\tau}{\sqrt{2}}$. It follows from (3.3) that

$$\tfrac{1}{2}\mu(\|x\| > t_{n+1}) \le \mu(\|x\| \le \tau)\mu(\|x\| > t_{n+1}) \le \mu\Big(\|x\| > \tfrac{t_{n+1}-\tau}{\sqrt{2}}\Big)^2 \le \mu(\|x\| > t_n)^2 ,$$

so that $\mu(\|x\| > t_{n+1}) \le 2\mu(\|x\| > t_n)^2$. Applying this inequality repeatedly, we obtain

$$\mu(\|x\| > t_n) \le 2^{2^n-1}\mu(\|x\| > t_0)^{2^n} \le 2^{2^n-1-2^{n+1}} \le 2^{-2^n} .$$

On the other hand, one can check explicitly that $t_n = \frac{\sqrt{2}^{n+1}-1}{\sqrt{2}-1}\tau \le 2^{n/2}\cdot(2+\sqrt{2})\tau$, so that in particular $t_{n+1} \le 2^{n/2}\cdot 5\tau$. This shows that one has the bound

$$\mu(\|x\| > t_n) \le 2^{-\frac{t_{n+1}^2}{25\tau^2}} ,$$

implying that there exists $\alpha > 0$ such that $\mu(\|x\| > t) \le \exp(-2\alpha t^2)$ for every $t \ge \tau$. Integrating by parts, we finally obtain

$$\int_{\mathcal{B}} \exp(\alpha\|x\|^2)\,\mu(dx) \le e^{\alpha\tau^2} + 2\alpha\int_\tau^\infty t e^{\alpha t^2}\mu(\|x\| > t)\,dt \le e^{\alpha\tau^2} + 2\alpha\int_\tau^\infty t e^{-\alpha t^2}\,dt < \infty ,$$

which is the desired result. □

As an immediate corollary, we have

Corollary 3.9 *There exists a constant $\|C_\mu\| < \infty$ such that $C_\mu(\ell,\ell') \le \|C_\mu\|\|\ell\|\|\ell'\|$ for any $\ell,\ell' \in \mathcal{B}^*$. In particular, C_μ can be interpreted as a continuous operator from $\mathcal{B}^*$ to $\mathcal{B}$ when $\mathcal{B}$ is reflexive.*

Actually, C_μ is more than just bounded. If $\mathcal{B}$ happens to be a Hilbert space, one has indeed the following result, which allows us to characterise in a very precise way the set of all centred Gaussian measures on a Hilbert space:

Proposition 3.10 *If $\mathcal{B} = \mathcal{H}$ is a Hilbert space, then the operator $\hat{C}_\mu\colon\mathcal{H} \to \mathcal{H}$ defined by the identity $\langle\hat{C}_\mu h, k\rangle = C_\mu(h,k)$ is trace class and one has the identity*

$$\int_{\mathcal{H}} \|x\|^2\,\mu(dx) = \operatorname{tr}\hat{C}_\mu . \tag{3.4}$$

(Here, we used Riesz's representation theorem to identify $\mathcal{H}$ with its dual.)

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

Proof. Fix an arbitrary orthonormal basis $\{e_n\}$ of $\mathcal{H}$. We know from Theorem 3.8 that the second moment of the norm is finite: $\int_{\mathcal{H}} \|h\|^2 \mu(dh) < \infty$. On the other hand, one has

$$\int_{\mathcal{H}} \|h\|^2 \mu(dh) = \sum_{n=1}^{\infty} \int_{\mathcal{H}} \langle h, e_n\rangle^2 \mu(dh) = \sum_{n=1}^{\infty} \langle e_n, \hat{C}_\mu e_n\rangle = \operatorname{tr} \hat{C}_\mu ,$$

which is precisely (3.4). To pull the sum out of the integral in the first equality, we used Lebegue's dominated convergence theorem.

In order to prove the converse statement, since K is compact, we can find an orthonormal basis $\{e_n\}$ of $\mathcal{H}$ such that $Ke_n = \lambda_n e_n$ and $\lambda_n \geq 0$, $\sum_n \lambda_n < \infty$. Let furthermore $\{\xi_n\}$ be a collection of i.i.d. $\mathcal{N}(0,1)$ Gaussian random variables (such a family exists by Kolmogorov's extension theorem). Then, since $\sum_n \lambda_n \mathbf{E}\xi_n^2 = \operatorname{tr} K < \infty$, the series $\sum_n \sqrt{\lambda_n}\xi_n e_n$ converges almost surely in $\mathcal{H}$. One can easily check that the law of the limiting random variable is Gaussian and has the requested covariance. □

In many situations, it is furthermore helpful to find out whether a given covariance structure can be realised as a Gaussian measure on some space of Hölder continuous functions. This can be achieved through the following version of Kolmogorov's continuity criterion, which can be found for example in [RY94, p. 26]:

Theorem 3.11 (Kolmogorov) *For $d > 0$, let $C\colon [0,1]^d \times [0,1]^d \to \mathbf{R}$ 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 $C(x,x) + C(y,y) - 2C(x,y) \leq K|x-y|^{2\alpha}$ for any two points $x, y \in [0,1]^d$ then there exists a unique centred Gaussian measure μ on $\mathcal{C}([0,1]^d, \mathbf{R})$ such that*

$$\int_{\mathcal{C}([0,1]^d,\mathbf{R})} f(x) f(y)\, \mu(df) = C(x,y)\,, \tag{3.5}$$

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

Before we turn to the proof of Theorem 3.11, we make a few preparations.

Proof of Theorem 3.11. Set $\mathcal{B} = \mathcal{C}([0,1]^d, \mathbf{R})$ and $\mathcal{B}^*$ its dual, which consists of the set of Borel measures with finite total variation [Yos95, p. 119]. Since convex combinations of Dirac measures are dense (in the topology of weak convergence) in the set of probability measures, it follows that the set of linear combinations of point evaluations is weakly dense in $\mathcal{B}^*$. Therefore, the claim follows if we are able to construct a measure μ on $\mathcal{B}$ such that (3.5) holds and such that, if f is distributed according to μ, then the law of $f(x)$ is Gaussian for every $x \in [0,1]^d$.

By Kolmogorov's extension theorem, we can construct a measure μ_0 on $\mathbf{R}^{[0,1]^d}$ endowed with the product σ-algebra such that the laws of all finite-dimensional marginals are Gaussian and satisfy (3.5). We denote by X a $\mathbf{R}^{[0,1]^d}$-valued random variable with law μ_0.

Denote now by $\mathcal{D} \subset [0,1]^d$ the subset of dyadic numbers and define the event Ω_β by

$$\Big\{X \;:\; \hat{X}(x) \stackrel{\text{def}}{=} \lim_{\substack{y\to x\\ y\in\mathcal{D}}} X(y) \text{ exists for every } x \in [0,1]^d \text{ and } \hat{X} \text{ belongs to } \mathcal{C}^\beta([0,1]^d, \mathbf{R})\Big\}\,.$$

Since the event Ω_β can be constructed from evaluating X at only countably many points, it is a measurable set. For the same reason, the map $\iota\colon \mathbf{R}^{[0,1]^d} \to \mathcal{C}^\beta([0,1]^d, \mathbf{R})$ given by

$$\iota(X) = \begin{cases} \hat{X} & \text{if } X \in \Omega_\beta, \\ 0 & \text{otherwise} \end{cases}$$

is measurable with respect to the product σ-algebra on $\mathbf{R}^{[0,1]^d}$, so that the claim follows if we can show that $\mu_0(\Omega_\beta) = 1$ for every $\beta < \alpha$. (Take $\mu = \iota^*\mu_0$.) Denoting the β-Hölder norm of X by $M_\beta(X) = \sup_{x\neq y\,:\,x,y\in\mathcal{D}}\{|X(x) - X(y)|/|x-y|^\beta\}$, we see that Ω_β can alternatively be characterised as $\Omega_\beta = \{X \,:\, M_\beta(X) < \infty\}$.

Denote by $\mathcal{D}_m \subset \mathcal{D}$ the set of those numbers whose coordinates are integer multiples of 2^{-m} and denote by Δ_m the set of pairs $x, y \in \mathcal{D}_m$ such that $|x-y| = 2^{-m}$. In particular, note that Δ_m has at most $2^{(m+2)d}$ such pairs. We are now going to make use of our simplifying assumption that we are dealing with Gaussian random variables, so that pth moments can be bounded in terms of second moments. More precisely, for every $p \geq 1$ there exists a constant C_p such that if X is a Gaussian random variable, then one has the bound $\mathbf{E}|X|^p \leq (\mathbf{E}|X|^2)^{p/2}$.

Setting $K_m(X) = \sup_{x,y\in\Delta_m}|X(x) - X(y)|$ and fixing some arbitrary $\beta' \in (\beta, \alpha)$, we see that for $p \geq 1$ large enough, there exists a constant K_p such that

$$\begin{aligned}
\mathbf{E}K_m^p(X) &\leq \sum_{x,y\in\Delta_m}\mathbf{E}|X(x) - X(y)|^p \leq C_p \sum_{x,y\in\Delta_m}(\mathbf{E}|X(x) - X(y)|^2)^{p/2} \\
&= C_p \sum_{x,y\in\Delta_m}(C(x,x) + C(y,y) - 2C(x,y))^{p/2} \leq K_p 2^{(m+2)d - \alpha m p} \\
&\leq K_p 2^{-\beta' m p}\;.
\end{aligned}$$

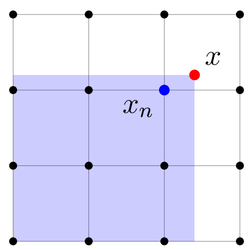
(In order to obtain the last inequality, we had to assume that $p \geq \frac{d}{\alpha-\beta'}\frac{m+2}{m}$ which can always be achieved since we assumed that $\beta' < \alpha$.) Using Jensen's inequality, this shows that there exists a constant K' such that the bound

$$\mathbf{E}K_m(X) \leq K' 2^{-\beta' m} \tag{3.6}$$

holds uniformly in m. Fix now any two points $x, y \in \mathcal{D}$ with $x \neq y$ and denote by m_0 the largest m such that $|x - y| \leq 2^{-m}$. One can then find sequences $\{x_n\}_{n\geq m_0}$ and $\{y_n\}_{n\geq m_0}$ with the following properties:

1. One has $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} y_n = y$.
2. Either $x_{m_0} = y_{m_0}$ or $(x_{m_0}, y_{m_0}) \in \Delta_{m_0}$.
3. For every $n \geq m_0$, x_n and x_{n+1} can be connected by at most d 'bonds' in Δ_{n+1} and similarly for (y_n, y_{n+1}).

One way of constructing this sequence is to order elements in $\mathcal{D}_m$ by lexicographic order and to choose $x_n = \max\{\bar{x} \in \mathcal{D}_n \,:\, \bar{x}_j \leq x_j\ \forall j\}$, as illustrated in the picture to the right. This shows that one has the bound



$$\begin{aligned}
|X(x) - X(y)| &\leq |X(x_{m_0}) - Y(x_{m_0})| + \sum_{n=m_0}^\infty |X(x_{n+1}) - X(x_n)| \\
&\quad + \sum_{n=m_0}^\infty |X(y_{n+1}) - X(y_n)| \\
&\leq K_{m_0}(X) + 2d\sum_{n=m_0}^\infty K_{n+1}(X) \leq 2d \sum_{n=m_0}^\infty K_n(X)\;.
\end{aligned}$$

Since m_0 was chosen in such a way that $|x - y| \geq 2^{-m_0-1}$, one has the bound

$$M_\beta(X) \leq 2d \sup_{m\geq 0} 2^{\beta(m+1)} \sum_{n=m}^\infty K_n(X) \leq 2^{\beta+1} d \sum_{n=0}^\infty 2^{\beta n} K_n(X)\;.$$

It follows from this and from the bound (3.6) that

$$\mathbf{E}|M_\beta(X)| \le 2^{\beta+1} d \sum_{n=0}^\infty 2^{\beta n} \mathbf{E} K_n(X) \le 2^{\beta+1} d K' \sum_{n=0}^\infty 2^{(\beta-\beta')n} < \infty \;,$$

since β' was chosen strictly larger than β. □

Remark 3.12 The space $\mathcal{C}^\beta([0,1]^d,\mathbf{R})$ is not separable. However, the space $\mathcal{C}_0^\beta([0,1]^d,\mathbf{R})$ of Hölder continuous functions that furthermore satisfy $\lim_{y\to x} \frac{|f(x)-f(y)|}{|x-y|^\beta} = 0$ uniformly in x is separable (polynomials with rational coefficients are dense in it). This is in complete analogy with the fact that the space of bounded measurable functions is not separable, while the space of continuous functions is.

It is furthermore possible to check that $\mathcal{C}^{\beta'} \subset \mathcal{C}_0^\beta$ for every $\beta' > \beta$, so that Exercise 3.29 below shows that μ can actually be realised as a Gaussian measure on $\mathcal{C}_0^\beta([0,1]^d,\mathbf{R})$.

Exercise 3.13 Try to find conditions on $G \subset \mathbf{R}^d$ that are as weak as possible and such that Kolmogorov's continuity theorem still holds if the cube $[0,1]^d$ is replaced by G. **Hint:** One possible strategy is to embed G into a cube and then to extend $C(x,y)$ to that cube.

Exercise 3.14 Show that if G is as in the previous exercise, $\mathcal{H}$ is a Hilbert space, and $C\colon G\times G \to \mathcal{L}(\mathcal{H},\mathcal{H})$ is such that $C(x,y)$ positive definite, symmetric, and trace class for any two $x,y \in G$, then Kolmogorov's continuity theorem still holds if its condition is replaced by $\operatorname{tr} C(x,x) + \operatorname{tr} C(y,y) - 2\operatorname{tr} C(x,y) \le K|x-y|^\alpha$. More precisely, one can construct a measure μ on the space $\mathcal{C}^\beta([0,1]^d,\mathcal{H})$ such that

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

for any $x,y \in G$ and $h,k \in \mathcal{H}$.

Corollary 3.15 *Let $\{\eta_k\}_{k\ge0}$ be countably many i.i.d. standard Gaussian random variables (real or complex). Moreover let $\{f_k\}_{k\ge0} \subset \mathrm{Lip}(G,\mathbf{C})$ where the domain $G \subset \mathbf{R}^d$ is sufficiently regular for Kolomgorov's continuity theorem to hold. Suppose there is some $\delta \in (0,2)$ such that*

$$S_1^2 = \sum_{k\in I} \|f_k\|_{L^\infty}^2 < \infty \quad\text{and}\quad S_2^2 = \sum_{k\in I} \|f_k\|_{L^\infty}^{2-\delta} \mathrm{Lip}(f_k)^\delta < \infty \;,$$

and define $f = \sum_{k\in I} \eta_k f_k$. Then f is almost surely bounded and continuous.

Proof. From the assumptions we immediately derive that $f(x)$ and $f(x) - f(y)$ are a centred Gaussian for any $x,y \in G$. Moreover, the corresponding series converge absolutely. Using that the η_k are i.i.d., we obtain

$$\begin{aligned}\mathbf{E}|f(x)-f(y)|^2 &= \sum_{k\in I} |f_k(x) - f_k(y)|^2 \le \sum_{k\in I} \min\{2\|f_k\|_{L^\infty}^2, \mathrm{Lip}(f_k)^2|x-y|^2\} \\ &\le 2\sum_{k\in I} \|f_k\|_{L^\infty}^{2-\delta} \mathrm{Lip}(f_k)^\delta |x-y|^\delta = 2S_2^2|x-y|^\delta \;,\end{aligned}$$

where we used that $\min\{a, bx^2\} \le a^{1-\delta/2} b^{\delta/2} |x|^\delta$ for any $a,b \ge 0$. The claim now follows from Kolmogorov's continuity theorem. □

3.2 The Cameron-Martin space

Given a Gaussian measure μ on a separable Banach space $\mathcal{B}$, it is possible to associate to it in a canonical way a Hilbert space $\mathcal{H}_\mu \subset \mathcal{B}$, called the Cameron-Martin space of μ. The main importance of the Cameron-Martin space is that it characterises precisely those directions in $\mathcal{B}$ in which translations leave the measure μ 'quasi-invariant' in the sense that the translated measure has the same null sets as the original measure. In general, the space $\mathcal{H}_\mu$ will turn out to be strictly smaller than $\mathcal{B}$. Actually, this is always the case as soon as $\dim \mathcal{H}_\mu = \infty$. Contrast this to the case of finite-dimensional Lebesgue measure which is invariant under translations in any direction! This is a striking illustration of the fact that measures in infinite-dimensional spaces have a strong tendency of being mutually singular.

The definition of the Cameron-Martin space is the following, where we postpone to Remark 3.17 and Proposition 3.20 the verification that $\|h\|_\mu$ is well-defined and that $\|h\|_\mu > 0$ for $h \neq 0$:

Definition 3.16 The Cameron-Martin space $\mathcal{H}_\mu$ of μ is the completion of the linear subspace $\hat{\mathcal{H}}_\mu \subset \mathcal{B}$ defined by

$$\hat{\mathcal{H}}_\mu = \{h \in \mathcal{B} \,:\, \exists\, h^* \in \mathcal{B}^* \text{ with } C_\mu(h^*, \ell) = \ell(h)\ \forall \ell \in \mathcal{B}^*\}\;,$$

under the norm $\|h\|_\mu^2 = \langle h, h\rangle_\mu = C_\mu(h^*, h^*)$. It is a Hilbert space when endowed with the scalar product $\langle h, k\rangle_\mu = C_\mu(h^*, k^*)$.

Remark 3.17 Even though the map $h \mapsto h^*$ may not be one to one, the norm $\|h\|_\mu$ is well-defined. To see this, assume that for a given $h \in \hat{\mathcal{H}}_\mu$, there are two corresponding elements h_1^* and h_2^* in $\mathcal{B}^*$. Then, defining $k = h_1^* + h_2^*$, one has

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

showing that $\|h\|_\mu$ does indeed not depend on the choice of h^*.

Exercise 3.18 The Wiener measure μ is defined on $\mathcal{B} = \mathcal{C}([0,1], \mathbf{R})$ as the centred Gaussian measure with covariance operator given by $C_\mu(\delta_s, \delta_t) = s \wedge t$. Show that the Cameron-Martin space for the Wiener measure on $\mathcal{B} = \mathcal{C}([0,1], \mathbf{R})$ is given by the set of all absolutely continuous functions h such that $\int_0^1 \dot{h}^2(t)\, dt < \infty$.

Exercise 3.19 Show that in the case $\mathcal{B} = \mathbf{R}^n$, the Cameron-Martin space is given by the range of the covariance matrix. Write an expression for $\|h\|_\mu$ in this case.

Let us discuss a few properties of the Cameron-Martin space. First of all, we show that it is a subspace of $\mathcal{B}$ despite the completion procedure and that all non-zero elements of $\mathcal{H}_\mu$ have strictly positive norm:

Proposition 3.20 *One has $\mathcal{H}_\mu \subset \mathcal{B}$. Furthermore, one has the bound*

$$\langle h, h\rangle_\mu \geq \|C_\mu\|^{-1} \|h\|^2\;, \tag{3.7}$$

where the norms on the right hand side are understood to be taken in $\mathcal{B}$.

Proof. One has the chain of inequalities

$$\|h\|^2 = \sup_{\ell \in \mathcal{B}^* \setminus \{0\}} \frac{\ell(h)^2}{\|\ell\|^2} = \sup_{\ell \in \mathcal{B}^* \setminus \{0\}} \frac{C_\mu(h^*, \ell)^2}{\|\ell\|^2} \le \sup_{\ell \in \mathcal{B}^* \setminus \{0\}} \frac{C_\mu(h^*, h^*)\, C_\mu(\ell, \ell)}{\|\ell\|^2} \le \|C_\mu\| \langle h, h \rangle_\mu ,$$

which yields the bound on the norms. The fact that $\mathcal{H}_\mu$ is a subset of $\mathcal{B}$ then follows from the fact that $\mathcal{B}$ is complete and that Cauchy sequences in $\hat{\mathcal{H}}_\mu$ are also Cauchy sequences in $\mathcal{B}$ by (3.7). □

Another remark is that the correspondence $h \mapsto h^*$ in the definition of $\hat{\mathcal{H}}_\mu$ is not necessarily unique. Consider for example the case $\mu = \delta_0$, so that $C_\mu = 0$. If one chooses $h = 0$, any $h^* \in \mathcal{B}$ has the required property that $C_\mu(h^*, \ell) = \ell(h)$. However, if we view $\mathcal{B}^*$ as a subset of $L^2(\mathcal{B}, \mu)$ (by identifying linear functionals that agree μ-almost surely), then the correspondence $h \mapsto h^*$ is an isomorphism. One has indeed $\int_{\mathcal{B}} h^*(x)^2 \,\mu(dx) = C_\mu(h^*, h^*) = \|h\|_\mu^2$. In particular, if h_1^* and h_2^* are two distinct elements of $\mathcal{B}^*$ associated to the same element $h \in \mathcal{B}$, then $h_1^* - h_2^*$ is associated to the element 0 and therefore $\int_{\mathcal{B}} (h_1^*(x) - h_2^*(x))^2 \,\mu(dx) = 0$, showing that $h_1 = h_2$ as elements of $L^2(\mathcal{B}, \mu)$. We have:

Proposition 3.21 *There is a canonical isomorphism $\iota\colon h \mapsto h^*$ between $\mathcal{H}_\mu$ and the closure R_μ of $\mathcal{B}^*$ in $L^2(\mathcal{B}, \mu)$.*

Proof. We have already shown that $\iota\colon \mathcal{H}_\mu \to L^2(\mathcal{B}, \mu)$ is an isomorphism onto its image, so it remains to show that all of $\mathcal{B}^*$ belongs to the image of ι. For $h \in \mathcal{B}^*$, define $h_* \in \mathcal{B}$ by

$$h_* = \int_{\mathcal{B}} x h(x) \mu(dx) .$$

(This integral converges since $\|x\|^2$ is integrable by Fernique's theorem.) Since one has the identity $\ell(h_*) = C_\mu(\ell, h)$, it follows that $h_* \in \hat{\mathcal{H}}_\mu$ and $h = \iota(h_*)$, as required to conclude the proof. □

Remark 3.22 The space R_μ is called the *reproducing kernel Hilbert space* for μ (or just *reproducing kernel* for short). However, since it is isomorphic to the Cameron-Martin space in a natural way, there is considerable confusion between the two in the literature. We retain in these notes the terminology from [Bog98].

Remark 3.23 Since $L^2(\mathcal{B}, \mu)$ is separable if $\mathcal{B}$ is separable, the same is true for $\mathcal{H}_\mu$ and R_μ. In general, there do however exist Gaussian measures with non-separable Cameron-Martin space.

Exercise 3.24 Let μ be a Gaussian measure on a Hilbert space $\mathcal{H}$ with covariance K and consider the spectral decomposition of K: $Ke_n = \lambda_n e_n$ with $\sum_{n \ge 1} \lambda_n < \infty$. Assume that $\lambda_n > 0$ for every n. Show that $\hat{\mathcal{H}}_\mu$ is given by the range of K and that the correspondence $h \mapsto h^*$ is given by $h^* = K^{-1} h$. Show furthermore that the Cameron-Martin space $\mathcal{H}_\mu$ consists of those elements h of $\mathcal{H}$ such that $\sum_{n \ge 1} \lambda_n^{-1} \langle h, e_n \rangle^2 < \infty$ and that $\langle h, k \rangle_\mu = \langle K^{-1/2} h, K^{-1/2} k \rangle$.

Exercise 3.25 Show that one has the alternative characterisation

$$\|h\|_\mu = \sup\{\ell(h) \,:\, C_\mu(\ell, \ell) \le 1\} , \tag{3.8}$$

and $\mathcal{H}_\mu = \{h \in \mathcal{B} \,:\, \|h\|_\mu < \infty\}$. **Hint:** Use the fact that in any Hilbert space $\mathcal{H}$, one has $\|h\| = \sup\{\langle k, h \rangle \,:\, \|k\| \le 1\}$.

Note also that

Proposition 3.26 *The law of any element of R_μ is a centred Gaussian.*

Proof. We already know from the definition of a Gaussian measure that the law of any element of $\mathcal{B}^*$ is a centred Gaussian. Let now h be any element of R_μ and let h_n be a sequence in $R_\mu \cap \mathcal{B}^*$ such that $h_n \to h$ in R_μ. We see that, if h and g are any two centred random variables with variances σ_h^2 and σ_g^2 respectively, then

$$\mathbf{E}(h-g)^2 = \sigma_h^2 + \sigma_g^2 - 2\mathbf{E}hg \geq \sigma_h^2 + \sigma_g^2 - 2|\sigma_h \sigma_g| = (|\sigma_h| - |\sigma_g|)^2 ,$$

thus showing that the variances σ_n^2 of h_n form a Cauchy sequence in $\mathbf{R}$ and therefore have a limit σ^2. Since L^2-convergence implies convergence in law and since we know that if $\sigma_n \to \sigma$, then $\mathcal{N}(0, \sigma_n^2) \to \mathcal{N}(0, \sigma^2)$ in law, we conclude that the law of h is given by $\mathcal{N}(0, \sigma^2)$. □

Furthermore, elements in R_μ are 'almost' linear functionals on $\mathcal{B}$ in the following sense:

Proposition 3.27 *For every $\ell \in R_\mu$ there exists a subspace V_ℓ of $\mathcal{B}$ such that $\mu(V_\ell) = 1$ and a linear map $\hat{\ell}\colon V_\ell \to \mathbf{R}$ such that $\ell = \hat{\ell}$ μ-almost surely.*

Proof. Fix $\ell \in R_\mu$. By the definition of R_μ and Borel-Cantelli, we can find a sequence $\ell_n \in \mathcal{B}^*$ such that $\lim_{n\to\infty} \ell_n(x) = \ell(x)$ for μ-almost every $x \in \mathcal{B}$. (Take for example ℓ_n such that $\|\ell_n - \ell\|_\mu^2 \leq n^{-4}$.) It then suffices to define

$$V_\ell = \{x \,:\, \lim_{n\to\infty} \ell_n(x) \text{ exists}\} ,$$

and to set $\hat{\ell}(x) = \lim_{n\to\infty} \ell_n(x)$ on V_ℓ. □

Remark 3.28 Actually, the converse of Proposition 3.27 is also true: if $\ell\colon \mathcal{B} \to \mathbf{R}$ is measurable and linear on a subspace of full measure, then ℓ belongs to R_μ. This is not an obvious statement. It uses the highly non-trivial fact that every Borel measurable linear map between two 'nice' topological vector spaces is bounded (see for example [Sch66, Kat82]), but we will not give its proof in these notes.

Exercise 3.29 Show that if $\tilde{\mathcal{B}} \subset \mathcal{B}$ is a continuously embedded Banach space with $\mu(\tilde{\mathcal{B}}) = 1$, then the embedding $\mathcal{B}^* \hookrightarrow R_\mu$ extends to an embedding $\tilde{\mathcal{B}}^* \hookrightarrow R_\mu$. Deduce from this that the restriction of μ to $\tilde{\mathcal{B}}$ is again a Gaussian measure. In particular, Kolmogorov's continuity criterion yields a Gaussian measure on $\mathcal{C}_0^\beta([0,1]^d, \mathbf{R})$.

The properties of the reproducing kernel space of a Gaussian measure allow us to give another illustration of the fact that measures on infinite-dimensional spaces behave in a rather different way from measures on $\mathbf{R}^n$:

Proposition 3.30 *Let μ be a centred Gaussian measure on a separable Banach space $\mathcal{B}$ such that* $\dim \mathcal{H}_\mu = \infty$. *Denote by D_c the dilatation by a real number c on $\mathcal{B}$, that is $D_c(x) = cx$. Then, μ and $D_c^*\mu$ are mutually singular for every $c \neq \pm 1$.*

Proof. Since the reproducing Kernel space R_μ is a separable Hilbert space, we can find an orthonormal basis $\{e_n\}_{n\geq 0}$. Consider the sequence of random variables $X_N(x) = \frac{1}{N}\sum_{n=1}^N |e_n(x)|^2$ over $\mathcal{B}$. If $\mathcal{B}$ is equipped with the measure μ then, since the e_n are independent under μ, we can apply the law of large numbers and deduce that

$$\lim_{N\to\infty} X_N(x) = 1 , \tag{3.9}$$

for μ-almost every x. On the other hand, it follows from the linearity of the e_n that when we equip $\mathcal{B}$ with the measure $D_c^*\mu$, the e_n are still independent, but have variance c^2, so that

$$\lim_{N\to\infty} X_N(x) = c^2 \;,$$

for $D_c^*\mu$-almost every x. This shows that if $c \neq \pm 1$, the set on which the convergence (3.9) takes place must be of $D_c^*\mu$-measure 0, which implies that μ and $D_c^*\mu$ are mutually singular. □

As already mentioned earlier, the importance of the Cameron-Martin space is that it represents precisely those directions in which one can translate the measure μ without changing its null sets:

Theorem 3.31 (Cameron-Martin) *For $h \in \mathcal{B}$, define the map $T_h\colon \mathcal{B} \to \mathcal{B}$ by $T_h(x) = x + h$. Then, the measure $T_h^*\mu$ is absolutely continuous with respect to μ if and only if $h \in \mathcal{H}_\mu$.*

Proof. Fix $h \in \mathcal{H}_\mu$ and let $h^* \in L^2(\mathcal{B}, \mu)$ be the corresponding element of the reproducing kernel. Since the law of h^* is Gaussian by Proposition 3.26, the map $x \mapsto \exp(h^*(x))$ is integrable. Since furthermore the variance of h^* is given by $\|h\|_\mu^2$, the function

$$\mathcal{D}_h(x) = \exp(h^*(x) - \tfrac{1}{2}\|h\|_\mu^2) \tag{3.10}$$

is strictly positive, belongs to $L^1(\mathcal{B}, \mu)$, and integrates to 1. It is therefore the Radon-Nikodym derivative of a measure μ_h that is absolutely continuous with respect to μ. To check that one has indeed $\mu_h = T_h^*\mu$, it suffices to show that their Fourier transforms coincide. Assuming that $h^* \in \mathcal{B}^*$, one has

$$\begin{aligned}\hat\mu_h(\ell) &= \int_{\mathcal{B}} \exp(i\ell(x) + h^*(x) - \tfrac{1}{2}\|h\|_\mu^2)\,\mu(dx) = \exp(\tfrac{1}{2}C_\mu(i\ell + h^*, i\ell + h^*) - \tfrac{1}{2}\|h\|_\mu^2)\\ &= \exp(-\tfrac{1}{2}C_\mu(\ell,\ell) - iC_\mu(\ell, h^*)) = \exp(-\tfrac{1}{2}C_\mu(\ell,\ell) + i\ell(h)) \;.\end{aligned}$$

Using Lebegue's dominated convergence theorem, it is an easy exercise to check that this equality still holds for arbitrary $h \in \mathcal{H}_\mu$.

On the other hand, we have

$$\begin{aligned}\widehat{T_h^*\mu}(\ell) &= \int_{\mathcal{B}} \exp(i\ell(x))\,T_h^*\mu(dx) = \int_{\mathcal{B}} \exp(i\ell(x+h))\,\mu(dx) = e^{i\ell(h)}\int_{\mathcal{B}} \exp(i\ell(x))\,\mu(dx)\\ &= \exp(-\tfrac{1}{2}C_\mu(\ell,\ell) + i\ell(h)) \;,\end{aligned}$$

showing that $\mu_h = T_h^*\mu$.

To show the converse, note first that one can check by an explicit calculation that $\|\mathcal{N}(0,1) - \mathcal{N}(h,1)\|_{\mathrm{TV}} \geq 2 - 2\exp(-\frac{h^2}{8})$. Fix now some arbitrary $n > 0$. If $h \notin \mathcal{H}_\mu$ then, by Exercise 3.25, there exists $\ell \in \mathcal{B}^*$ with $C_\mu(\ell,\ell) = 1$ such that $\ell(h) \geq n$. Since the image $\ell^*\mu$ of μ under ℓ is $\mathcal{N}(0,1)$ and the image of $T_h^*\mu$ under ℓ is $\mathcal{N}(-\ell(h), 1)$, this shows that

$$\|\mu - T_h^*\mu\|_{\mathrm{TV}} \geq \|\ell^*\mu - \ell^*T_h^*\mu\|_{\mathrm{TV}} = \|\mathcal{N}(0,1) - \mathcal{N}(-\ell(h),1)\|_{\mathrm{TV}} \geq 2 - 2\exp(-\frac{n^2}{8}) \;.$$

Since this is true for every n, we conclude that $\|\mu - T_h^*\mu\|_{\mathrm{TV}} = 2$, thus showing that they are mutually singular. □

As a consequence, we have the following characterisation of the Cameron-Martin space

Proposition 3.32 *The space $\mathcal{H}_\mu \subset \mathcal{B}$ is the intersection of all linear subspaces of full measure. However, if $\mathcal{H}_\mu$ is infinite-dimensional, then one has $\mu(\mathcal{H}_\mu) = 0$.*

Proof. Take an arbitrary linear subspace $V \subset \mathcal{B}$ of full measure and take an arbitrary $h \in \mathcal{H}_\mu$. It follows from Theorem 3.31 that the affine space $V-h$ also has full measure. Since $(V-h)\cap V = \phi$ unless $h \in V$, one must have $h \in V$, so that $\mathcal{H}_\mu \in \bigcap\{V \subset \mathcal{B} \,:\, \mu(V) = 1\}$.

Conversely, take an arbitrary $x \notin \mathcal{H}_\mu$ and let us construct a linear space $V \subset \mathcal{B}$ of full measure, but not containing x. Since $x \notin \mathcal{H}_\mu$, one has $\|x\|_\mu = \infty$ with $\|\cdot\|_\mu$ extended to $\mathcal{B}$ as in (3.8). Therefore, we can find a sequence $\ell_n \in \mathcal{B}^*$ such that $C_\mu(\ell_n, \ell_n) \leq 1$ and $\ell_n(x) \geq n$. Defining the norm $|y|^2 = \sum_n n^{-2}(\ell_n(y))^2$, we see that

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

so that the linear space $V = \{y \,:\, |y| < \infty\}$ has full measure. However, $|x| = \infty$ by construction, so that $x \notin V$.

To show that $\mu(\mathcal{H}_\mu) = 0$ if $\dim \mathcal{H}_\mu = \infty$, consider an orthonormal sequence $e_n \in R_\mu$ so that the random variables $\{e_n(x)\}$ are i.i.d. $\mathcal{N}(0,1)$ distributed. By the second Borel-Cantelli lemma, it follows that $\sup_n |e_n(x)| = \infty$ for μ-almost every x, so that in particular $\|x\|_\mu^2 \geq \sum_n e_n^2(x) = \infty$ almost surely. □

Exercise 3.33 Recall that the (topological) support $\operatorname{supp}\mu$ of a Borel measure on a complete separable metric space consists of those points x such that $\mu(U) > 0$ for every neighbourhood U of x. Show that, if μ is a Gaussian measure, then $\operatorname{supp}\mu$ is the closure $\bar{\mathcal{H}}_\mu$ of $\mathcal{H}_\mu$ in $\mathcal{B}$.

3.3 Images of Gaussian measures

It follows immediately from the definition of a Gaussian measure and the expression for its Fourier transform that if μ is a Gaussian measure on some Banach space $\mathcal{B}$ and $A\colon \mathcal{B} \to \mathcal{B}_2$ is a bounded linear map for $\mathcal{B}_2$ some other Banach space, then $\nu = A^*\mu$ is a Gaussian measure on $\mathcal{B}_2$ with covariance

$$C_\nu(\ell, \ell') = C_\mu(A^*\ell, A^*\ell') \;,$$

where $A^*\colon \mathcal{B}_2^* \to \mathcal{B}^*$ is the adjoint to A.

Recall now that $\mathcal{H}_\mu$ is the intersection over all linear subspaces of $\mathcal{B}$ that have full measure under μ. This suggests that in order to determine the image of μ under a linear map, it is sufficient to know how that map acts on elements of $\mathcal{H}_\mu$. This intuition is made precise by the following theorem:

Theorem 3.34 *Let μ be a centred Gaussian probability measure on a separable Banach space $\mathcal{B}$. Let furthermore $\mathcal{H}$ be a separable Hilbert space and let $A\colon \mathcal{H}_\mu \to \mathcal{H}$ be a Hilbert-Schmidt operator. (That is $AA^*\colon \mathcal{H} \to \mathcal{H}$ is trace class.) Then, there exists a measurable map $\hat{A}\colon \mathcal{B} \to \mathcal{H}$ such that $\nu = \hat{A}^*\mu$ is Gaussian with covariance $C_\nu(h,k) = \langle A^*h, A^*k\rangle_\mu$. Furthermore, there is a subspace $V \subset \mathcal{B}$ of full μ-measure such that $\hat{A}$ restricted to V is linear and $\hat{A}$ restricted to $\mathcal{H}_\mu \subset V$ agrees with A.*

Proof. Let $\{e_n\}_{n\geq 1}$ be an orthonormal basis for $\mathcal{H}_\mu$ and denote by e_n^* the corresponding elements in $R_\mu \subset L^2(\mathcal{B}, \mu)$. Recall from Proposition 3.27 that we can find subspaces V_{e_n} of full measure such that e_n^* is linear on V_{e_n}. Define now a linear subspace $V \subset \mathcal{B}$ by

$$V = \Big\{x \in \bigcap_{n\geq 0} V_{e_n} \,:\, \sum_{n\geq 0} e_n^*(x) A e_n \text{ converges in } \mathcal{H}\Big\} \;,$$

(the fact that V is linear follows from the linearity of each of the e_n^*) and set

$$\hat{A}(x) = \begin{cases} \sum_{n\geq 0} e_n^*(x) A e_n & \text{for } x \in V, \\ 0 & \text{otherwise.} \end{cases}$$

Since the random variables $\{e_n^*\}$ are i.i.d. $\mathcal{N}(0,1)$-distributed under μ, one has

$$\mathbf{E}_\mu \Big\| \sum_{n\geq 1} e_n^* A e_n \Big\|^2 = \sum_{n\geq 1} \|A e_n\|^2 = \operatorname{tr} A^* A < \infty$$

as a consequence of A being Hilbert-Schmidt. Therefore, $\mu(V) = 1$.

To see that $\nu = \hat{A}^*\mu$ has the stated property, fix an arbitrary $h \in \mathcal{H}$ and note that the series $\sum_{n\geq 1} e_n^* \langle A e_n, h\rangle$ converges in R_μ to an element with covariance $\|A^* h\|^2$. The statement then follows from Proposition 3.26 and the fact that $C_\nu(h,h)$ determines C_ν by polarisation. To check that ν is Gaussian, we can compute its Fourier transform in a similar way. □

Remark 3.35 Similarly to Proposition 3.27, the converse is again true: if $\hat{A}\colon \mathcal{B} \to \mathcal{H}$ is a measurable map which is linear on a subspace of full measure and agrees with A on $\mathcal{H}_\mu$, then it agrees μ-almost surely with the extension constructed in Theorem 3.34.

3.4 Cylindrical Wiener processes and stochastic integration

Central to the theory of stochastic PDEs is the notion of a *cylindrical Wiener process*. Recall that the usual (one-dimensional) Wiener process is a real-valued Gaussian process $B(t)$ such that $B(0) = 0$ and $\mathbf{E}|B(t) - B(s)|^2 = |t - s|$ for any pair of times s, t. From our Gaussian measure point of view, the Wiener process can always be realised as the canonical process for the Gaussian measure on $\mathcal{C}(\mathbf{R}, \mathbf{R})$ with covariance function $C(s,t) = s \wedge t = \min\{s,t\}$. (See Kolmogorov's continuity theorem.)

Let us now fix a (separable) Hilbert space $\mathcal{H}$, as well as a larger Hilbert space $\mathcal{H}'$ containing $\mathcal{H}$ as a dense subset and such that the inclusion map $\iota\colon \mathcal{H} \to \mathcal{H}'$ is Hilbert-Schmidt. Given $\mathcal{H}$, it is always possible to construct a space $\mathcal{H}'$ with this property: choose an orthonormal basis $\{e_n\}$ of $\mathcal{H}$ and take $\mathcal{H}'$ to be the closure of $\mathcal{H}$ under the norm

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

One can check that the map $\iota\iota^*$ is then given by $\iota\iota^* e_n = \frac{1}{n^2} e_n$, so that it is indeed trace class.

Definition 3.36 Let $\mathcal{H}$ and $\mathcal{H}'$ be as above. We then call a *cylindrical Wiener process on* $\mathcal{H}$ an $\mathcal{H}'$-valued Gaussian process W such that

$$\mathbf{E}\langle h, W(s)\rangle_{\mathcal{H}'} \langle W(t), k\rangle_{\mathcal{H}'} = (s\wedge t)\langle \iota^* h, \iota^* k\rangle = (s \wedge t)\langle \iota\iota^* h, k\rangle_{\mathcal{H}'} , \tag{3.11}$$

for any two times s and t and any two elements $h, k \in \mathcal{H}'$. By Kolmogorov's continuity theorem, this can be realised as the canonical process for some Gaussian measure on $\mathcal{C}(\mathbf{R}, \mathcal{H}')$.

Proposition 3.37 *In the same setting as above, the Gaussian measure μ on $\mathcal{H}'$ with covariance $\iota\iota^*$ has $\mathcal{H}$ as its Cameron-Martin space. Furthermore, $\|h\|_\mu^2 = \|h\|^2$ for every $h \in \mathcal{H}$.*

Proof. It follows from the definition of $\hat{\mathcal{H}}_\mu$ that this is precisely the range of $\iota\iota^*$ and that the map $h \mapsto h^*$ is given by $h^* = (\iota\iota^*)^{-1}h$. In particular, $\hat{\mathcal{H}}_\mu$ is contained in the range of ι. Therefore, for any $h, k \in \hat{\mathcal{H}}_\mu$, there exist $\hat{h}, \hat{g} \in \mathcal{H}$ such that $h = \iota\hat{h}$ and $k = \iota\hat{k}$. Using this, we have

$$\langle h, k\rangle_\mu = \langle (\iota\iota^*)h^*, k^*\rangle_{\mathcal{H}'} = \langle h, (\iota\iota^*)^{-1}k\rangle_{\mathcal{H}'} = \langle \iota\hat{h}, (\iota\iota^*)^{-1}\iota\hat{k}\rangle_{\mathcal{H}'} = \langle \hat{h}, \iota^*(\iota\iota^*)^{-1}\iota\hat{k}\rangle = \langle \hat{h}, \hat{k}\rangle\ ,$$

from which the claim follows. □

The name 'cylindrical Wiener process on $\mathcal{H}$' may sound confusing at first, since it is actually *not* an $\mathcal{H}$-valued process. (A better terminology may have been 'cylindrical Wiener process over $\mathcal{H}$', but we choose to follow the convention that is found in the literature.) Note however that if h is an element in $\mathcal{H}$ that is in the range of ι^* (so that ιh belongs to the range of $\iota\iota^*$ and $\iota^*(\iota\iota^*)^{-1}\iota h = h$), then

$$\langle h, k\rangle = \langle \iota^*(\iota\iota^*)^{-1}\iota h, k\rangle = \langle (\iota\iota^*)^{-1}\iota h, \iota k\rangle_{\mathcal{H}'}\ .$$

In particular, if we *pretend* that $W(t)$ belongs to $\mathcal{H}$ (which is of course not true), then we get

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

Here we used (3.11) to go from the first to the second line. This shows that $W(t)$ should be thought of as an $\mathcal{H}$-valued random variable with covariance t times the identity operator (which is of course not trace class if $\mathcal{H}$ is infinite-dimensional, so that such an object does not exist in general). Combining Proposition 3.37 with Theorem 3.34, we see however that if $\mathcal{K}$ is some Hilbert space and $A\colon \mathcal{H} \to \mathcal{K}$ is a Hilbert-Schmidt operator, then the $\mathcal{K}$-valued random variable $AW(t)$ is well-defined. (Here we made an abuse of notation and also used the symbol A for the measurable extension of A to $\mathcal{H}'$.) Furthermore, its law does not depend on the choice of the larger space $\mathcal{H}'$.

Example 3.38 (White noise) Recall that we informally defined 'white noise' as a Gaussian process ξ with covariance $\mathbf{E}\xi(s)\xi(t) = \delta(t-s)$. In particular, if we denote by $\langle\cdot,\cdot\rangle$ the scalar product in L^2, this suggests that

$$\mathbf{E}\langle g, \xi\rangle\langle h, \xi\rangle = \mathbf{E}\iint g(s)h(t)\xi(s)\xi(t)\,ds\,dt = \iint g(s)h(t)\delta(t-s)\,ds\,dt = \langle g, h\rangle\ . \tag{3.12}$$

This calculation shows that white noise can be constructed as a Gaussian random variable on any space of distributions containing L^2 and such the embedding is Hilbert-Schmidt. Furthermore, by Theorem 3.34, Integrals of the form $\int g(t)\xi(t)\,dt$ can be well-defined. Taking for g the indicator function of the interval $[0,s]$, we can check that the process $B(t) = \int_0^t \xi(s)\,ds$ is a Brownian motion, thus justifying the statement that 'white noise is the derivative of Brownian motion'.

This will allow us to define a Hilbert space-valued stochastic integral against a cylindrical Wiener process in pretty much the same way as what is usually done in finite dimensions. In the sequel, we fix a cylindrical Wiener process W on some Hilbert space $\mathcal{H} \subset \mathcal{H}'$, which we realise as the canonical coordinate process on $\Omega = \mathcal{C}(\mathbf{R}_+, \mathcal{H}')$ equipped with the measure constructed above. We also denote by $\mathscr{F}_s$ the σ-field on Ω generated by $\{W_r : r \le s\}$.

Consider now a finite collection of *disjoint* intervals $(s_n, t_n] \subset \mathbf{R}_+$ with $n = 1, \ldots, N$ and a corresponding finite collection of $\mathscr{F}_{s_n}$-measurable random variables Φ_n taking values in the space $\mathcal{L}_2(\mathcal{H}, \mathcal{K})$ of Hilbert-Schmidt operators from $\mathcal{H}$ into some other Hilbert space $\mathcal{K}$. Let furthermore Φ be the $L^2(\mathbf{R}_+ \times \Omega, \mathcal{L}_2(\mathcal{H}, \mathcal{K}))$-valued function defined by

$$\Phi(t, \omega) = \sum_{n=1}^{N} \Phi_n(\omega)\, \mathbf{1}_{(s_n, t_n]}(t) ,$$

where we denoted by $\mathbf{1}_A$ the indicator function of a set A. We call such a Φ an *elementary process on* $\mathcal{H}$.

Definition 3.39 Given an elementary process Φ and a cylindrical Wiener process W on $\mathcal{H}$, we define the $\mathcal{K}$-valued stochastic integral

$$\int_0^\infty \Phi(t)\, dW(t) \stackrel{\text{def}}{=} \sum_{n=1}^{N} \Phi_n(W)\, (W(t_n) - W(s_n)) .$$

Note that since Φ_n is $\mathscr{F}_{s_n}$-measurable, $\Phi_n(W)$ is independent of $W(t_n) - W(s_n)$, therefore each term on the right hand side can be interpreted in the sense of the construction of Theorem 3.34.

It follows from Theorem 3.34 and (3.4) that one has the identity

$$\mathbf{E}\Big\| \int_0^\infty \Phi(t)\, dW(t) \Big\|_{\mathcal{K}}^2 = \sum_{n=1}^{N} \mathbf{E}\, \mathrm{tr}(\Phi_n(W)\Phi_n^*(W))(t_n - s_n) = \mathbf{E} \int_0^\infty \mathrm{tr}\, \Phi(t)\Phi^*(t)\, dt , \qquad (3.13)$$

which is an extension of the usual Itô isometry to the Hilbert space setting. It follows that the stochastic integral is an isometry from the subset of elementary processes in $L^2(\mathbf{R}_+ \times \Omega, \mathcal{L}_2(\mathcal{H}, \mathcal{K}))$ to $L^2(\Omega, \mathcal{K})$.

Let now $\mathscr{F}_{\mathrm{pr}}$ be the 'predictable' σ-field, that is the σ-field over $\mathbf{R}_+ \times \Omega$ generated by all subsets of the form $(s, t] \times A$ with $t > s$ and $A \in \mathscr{F}_s$. This is the smallest σ-algebra with respect to which all elementary processes are $\mathscr{F}_{\mathrm{pr}}$-measurable. One furthermore has:

Proposition 3.40 *The set of elementary processes is dense in the space* $L^2_{\mathrm{pr}}(\mathbf{R}_+ \times \Omega, \mathcal{L}_2(\mathcal{H}, \mathcal{K}))$ *of all predictable* $\mathcal{L}_2(\mathcal{H}, \mathcal{K})$*-valued processes.*

Proof. Denote by $\hat{\mathscr{F}}_{\mathrm{pr}}$ the set of all sets of the form $(s, t] \times A$ with $A \in \mathscr{F}_s$. Denote furthermore by $\hat{L}^2_{\mathrm{pr}}$ the closure of the set of elementary processes in L^2. One can check that $\hat{\mathscr{F}}_{\mathrm{pr}}$ is closed under intersections, so that $\mathbf{1}_G \in \hat{L}^2_{\mathrm{pr}}$ for every set G in the algebra generated by $\hat{\mathscr{F}}_{\mathrm{pr}}$. It follows from the monotone class theorem that $\mathbf{1}_G \in \hat{L}^2_{\mathrm{pr}}$ for every set $G \in \mathscr{F}_{\mathrm{pr}}$. The claim then follows from the definition of the Lebesgue integral, just as for the corresponding statement in $\mathbf{R}$. □

By using the Itô isometry (3.13) and the completeness of $L^2(\Omega, \mathcal{K})$, it follows that:

Corollary 3.41 *The stochastic integral* $\int_0^\infty \Phi(t)\, dW(t)$ *can be defined for every process* $\Phi \in L^2_{\mathrm{pr}}(\mathbf{R}_+ \times \Omega, \mathcal{L}_2(\mathcal{H}, \mathcal{K}))$.

This concludes our presentation of the basic properties of Gaussian measures on infinite-dimensional spaces. The next section deals with the other main ingredient to solving stochastic PDEs, which is the behaviour of deterministic linear PDEs.

4 Some semigroup theory

This section is strongly based on the monograph [Dav80] for the first part on strongly continuous semigroups and very loosely follows [Yos95] and [Lun95] for the second part on analytic semigroups. Its aim is to give a rigorous meaning to solutions to linear equations of the type

$$\partial_t x = Lx \;, \quad x(0) = x_0 \;, \tag{4.1}$$

where x takes values in some Banach space and L is a possibly unbounded operator on that Banach space. From a formal point of view, if such a solution exists, one expects the existence of a linear operator $S(t)$ that maps the initial condition x_0 onto the solution $x(t)$ of (4.1) at time t. If such a solution is unique, then the family of operators $S(t)$ should have the property that $S(t) \circ S(s) = S(t+s)$. This motivates the following definition:

Definition 4.1 A semigroup $S(t)$ on a Banach space $\mathcal{B}$ is a family of bounded linear operators $\{S(t)\}_{t\geq 0}$ with the properties that $S(t) \circ S(s) = S(t+s)$ for any $s,t \geq 0$ and that $S(0) = \mathrm{Id}$. A semigroup is furthermore called

- *strongly continuous* if the map $(x,t) \mapsto S(t)x$ is strongly continuous.
- *analytic* if there exists $\theta > 0$ such that the map $t \mapsto S(t)$ has an analytic extension to $\{\lambda \in \mathbf{C} \,:\, |\arg \lambda| < \theta\}$, satisfies the semigroup property there, and is such that $t \mapsto S(e^{i\varphi}t)$ is a strongly continuous semigroup for every $|\varphi| < \theta$.

A strongly continuous semigroup is also called a $\mathcal{C}_0$-semigroup.

Exercise 4.2 Show that being strongly continuous is equivalent to $t \mapsto S(t)x$ being continuous at $t = 0$ for every $x \in \mathcal{B}$ and the operator norm of $S(t)$ being bounded by Me^{at} for some constants M and a. Show then that the first condition can be relaxed to $t \mapsto S(t)x$ being continuous for all x in some dense subset of $\mathcal{B}$. (However, the second condition cannot be relaxed in general. See Exercise 5.20 on how to construct a semigroup of bounded operators such that $\|S(t)\|$ is unbounded near $t = 0$.)

Remark 4.3 Some authors, like [Lun95], do not impose strong continuity in the definition of an analytic semigroup. This can result in additional technical complications due to the fact that the generator may then not have dense domain.

This section is going to assume some familiarity with functional analysis. All the necessary results can be found for example in the classical monograph by Yosida [Yos95]. Recall that an unbounded operator L on a Banach space $\mathcal{B}$ consists of a linear subspace $\mathcal{D}(L) \subset \mathcal{B}$ called the *domain* of L and a linear map $L\colon \mathcal{D}(L) \to \mathcal{B}$. The *graph* of an operator is the subset of $\mathcal{B} \times \mathcal{B}$ consisting of all elements of the form (x, Lx) with $x \in \mathcal{D}(L)$. An operator is *closed* if its graph is a closed subspace of $\mathcal{B} \times \mathcal{B}$. It is *closable* if the closure of its graph is again the graph of a linear operator and that operator is called the *closure* of L.

Exercise 4.4 Show that L being closed is equivalent to the fact that if $\{x_n\} \subset \mathcal{D}(L)$ is Cauchy in $\mathcal{B}$ and $\{Lx_n\}$ is also Cauchy, then $x = \lim_{n\to\infty} x_n$ belongs to $\mathcal{D}(L)$ and $Lx = \lim_{n\to\infty} Lx_n$.

The *resolvent set* $\varrho(L)$ of an operator L is defined by

$$\varrho(L) = \{\lambda \in \mathbf{C} \,:\, \mathrm{range}(\lambda - L) \text{ is dense in } \mathcal{B} \text{ and } \lambda - L \text{ has a continuous inverse.}\}\;,$$

and the *resolvent* R_λ is given for $\lambda \in \varrho(L)$ by $R_\lambda = (\lambda - L)^{-1}$. (Here and in the sequel we view $\mathcal{B}$ as a complex Banach space. If an operator is defined on a real Banach space, it can always be extended to its complexification in a canonical way and we will identify the two without further notice in the sequel.) The spectrum of L is the complement of the resolvent set.

The important results that we are going to use are that any closed operator with non-empty resolvent set is defined in a unique way by its resolvent. Furthermore, the resolvent is an analytic function from $\varrho(L)$ to the space $\mathcal{L}(\mathcal{B})$ of bounded linear operators on $\mathcal{B}$ which all commute and satisfy the resolvent identity

$$R_\lambda - R_\mu = (\mu - \lambda) R_\mu R_\lambda \;,$$

for any two $\lambda, \mu \in \varrho(L)$.

4.1 Strongly continuous semigroups

We start our investigation of semigroup theory with a discussion of the main results that can be obtained for strongly continuous semigroups. Given a $\mathcal{C}_0$-semigroup, one can associate to it a 'generator', which is essentially the derivative of $S(t)$ at $t = 0$:

Definition 4.5 The *generator* L of a $\mathcal{C}_0$-semigroup is given by

$$Lx = \lim_{t\to 0} t^{-1}(S(t)x - x) \;, \tag{4.2}$$

on the set $\mathcal{D}(L)$ of all elements $x \in \mathcal{B}$ such that this limit exists.

The following result shows that if L is the generator of a $\mathcal{C}_0$-semigroup $S(t)$, then $x(t) = S(t)x_0$ is indeed the solution to (4.1) in a weak sense.

Proposition 4.6 *The domain $\mathcal{D}(L)$ of L is dense in $\mathcal{B}$, invariant under S, and $\partial_t S(t)x = LS(t)x = S(t)Lx$ for every $x \in \mathcal{D}(L)$ and every $t \geq 0$. Furthermore, for every $\ell \in \mathcal{D}(L^*)$ and every $x \in \mathcal{B}$, the map $t \mapsto \langle \ell, S(t)x\rangle$ is differentiable and one has $\partial_t \langle \ell, S(t)x\rangle = \langle L^*\ell, S(t)x\rangle$.*

Proof. Fix some arbitrary $x \in \mathcal{B}$ and set $x_t = \int_0^t S(s)x\,ds$. One then has

$$\begin{aligned}
\lim_{h\to 0} h^{-1}(S(h)x_t - x_t) &= \lim_{h\to 0} h^{-1}\Big(\int_h^{t+h} S(s)x\,ds - \int_0^t S(s)x\,ds\Big) \\
&= \lim_{h\to 0} h^{-1}\Big(\int_t^{t+h} S(s)x\,ds - \int_0^h S(s)x\,ds\Big) = S(t)x - x \;,
\end{aligned}$$

where the last equality follows from the strong continuity of S. This shows that $x_t \in \mathcal{D}(L)$. Since $t^{-1}x_t \to x$ as $t \to 0$ and since x was arbitrary, it follows that $\mathcal{D}(L)$ is dense in $\mathcal{B}$. To show that it is invariant under S, note that for $x \in \mathcal{D}(L)$ one has

$$\lim_{h\to 0} h^{-1}(S(h)S(t)x - S(t)x) = S(t) \lim_{h\to 0} h^{-1}(S(h)x - x) = S(t)Lx \;,$$

so that $S(t)x \in \mathcal{D}(L)$ and $LS(t)x = S(t)Lx$. To show that it this is equal to $\partial_t S(t)x$, it suffices to check that the left derivative of this expression exists and is equal to the right derivative. This is left as an exercise.

To show that the second claim holds, it is sufficient (using the strong continuity of S) to check that it holds for $x \in \mathcal{D}(L)$. Since one then has $S(t)x \in \mathcal{D}(L)$ for every t, it follows from the definition (4.2) of $\mathcal{D}(L)$ that $t \mapsto S(t)x$ is differentiable and that its derivative is equal to $LS(t)x$. □

It follows as a corollary that no two semigroups can have the same generator, which justifies the notation $S(t) = e^{Lt}$ that we are occasionally going to use in the sequel.

Corollary 4.7 *If a function* $x\colon [0,1] \to \mathcal{D}(L)$ *satisfies* $\partial_t x_t = L x_t$ *for every* $t \in [0,1]$*, then* $x_t = S(t)x_0$*. In particular, no two distinct* $\mathcal{C}_0$*-semigroups can have the same generator.*

Proof. Fix an arbitrary $\ell \in \mathcal{D}(L^*)$ and $T \in (0,1)$. It follows from Proposition 4.6 that the map $t \mapsto \langle \ell, S(t)x_{T-t}\rangle$ is continuous on $[0,T]$ and differentiable on $(0,T)$ and that $\partial_t S(t)x_{T-t} = LS(t)x_{T-t} - S(t)Lx_{T-t} = 0$, so that $x_T = S(T)x_0$. □

Exercise 4.8 Show that the semigroup $S(t)$ on $L^2(\mathbf{R})$ given by

$$(S(t)f)(\xi) = f(\xi + t)\;,$$

is strongly continuous and that its generator is given by $L = \partial_\xi$ with $\mathcal{D}(L) = H^1$. Similarly, show that the heat semigroup on $L^2(\mathbf{R})$ given by

$$(S(t)f)(\xi) = \frac{1}{\sqrt{4\pi t}}\int \exp\Bigl(-\frac{|\xi-\eta|^2}{4t}\Bigr) f(\eta)\,d\eta\;,$$

is strongly continuous and that its generator is given by $L = \partial_\xi^2$ with $\mathcal{D}(L) = H^2$. **Hint:** Use Exercise 4.2 to show strong continuity.

Remark 4.9 We did not make any assumption on the structure of the Banach space $\mathcal{B}$. *However*, it is a general rule of thumb (although this is *not* a theorem) that semigroups on non-separable Banach spaces tend not to be strongly continuous. For example, neither the heat semigroup nor the translation semigroup from the previous exercise are strongly continuous on $L^\infty(\mathbf{R})$ or even on $\mathcal{C}_b(\mathbf{R})$, the space of all bounded continuous functions.

Recall now that the resolvent set for an operator L consists of those $\lambda \in \mathcal{C}$ such that the operator $\lambda - L$ is one to one. For λ in the resolvent set, we denote by $R_\lambda = (\lambda - L)^{-1}$ the resolvent of L. It turns out that the resolvent of the generator of a $\mathcal{C}_0$-semigroup can easily be computed:

Proposition 4.10 *Let* $S(t)$ *be a* $\mathcal{C}_0$*-semigroup such that* $\|S(t)\| \le Me^{at}$ *for some constants* M *and* a*. If* $\mathrm{Re}\lambda > a$*, then* λ *belongs to the resolvent set of* L *and one has the identity* $R_\lambda x = \int_0^\infty e^{-\lambda t} S(t)x\,dt$*.*

Proof. By the assumption on the bound on S, the expression $Z_\lambda = \int_0^\infty e^{-\lambda t}S(t)x\,dt$ is well-defined for every λ with $\mathrm{Re}\lambda > a$. In order to show that $Z_\lambda = R_\lambda$, we first show that $Z_\lambda x \in \mathcal{D}(L)$ for every $x \in \mathcal{B}$ and that $(\lambda - L)Z_\lambda x = x$. We have

$$\begin{aligned}
LZ_\lambda x &= \lim_{h\to 0} h^{-1}(S(h)Z_\lambda x - Z_\lambda x) = \lim_{h\to 0} h^{-1}\int_0^\infty e^{-\lambda t}(S(t+h)x - S(t)x)\,dt \\
&= \lim_{h\to 0}\Bigl(\frac{e^{\lambda h}-1}{h}\int_0^\infty e^{-\lambda t}S(t)x\,dt - h^{-1}\int_0^h e^{-\lambda t}S(t)x\,dt\Bigr) \\
&= \lambda Z_\lambda x - x\;,
\end{aligned}$$

which is the required identity. To conclude, it remains to show that $\lambda - L$ is an injection on $\mathcal{D}(L)$. If it was not, we could find $x \in \mathcal{D}(L)\setminus\{0\}$ such that $Lx = \lambda x$. Setting $x_t = e^{\lambda t}x$ and applying Corollary 4.7, this yields $S(t)x = e^{\lambda t}x$, thus contradicting the bound $\|S(t)\| \le Me^{at}$ if $\mathrm{Re}\lambda > a$. □

We can deduce from this that:

Proposition 4.11 *The generator L of a $\mathcal{C}_0$-semigroup is a closed operator.*

Proof. We are going to use the characterisation of closed operators given in Exercise 4.4. Shifting L by a constant if necessary (which does not affect it being closed or not), we can assume that $a = 0$. Take now a sequence $x_n \in \mathcal{D}(L)$ such that $\{x_n\}$ and $\{Lx_n\}$ are both Cauchy in $\mathcal{B}$ and set $x = \lim_{n\to\infty} x_n$ and $y = \lim_{n\to\infty} Lx_n$. Setting $z_n = (1 - L)x_n$, we have $\lim_{n\to\infty} z_n = x - y$.

On the other hand, we know that 1 belongs to the resolvent set, so that

$$x = \lim_{n\to\infty} x_n = \lim_{n\to\infty} R_1 z_n = R_1(x - y) \;.$$

By the definition of the resolvent, this implies that $x \in \mathcal{D}(L)$ and that $x - Lx = x - y$, so that $Lx = y$ as required. □

We are now ready to give a full characterisation of the generators of $\mathcal{C}_0$-semigroups. This is the content of the following theorem:

Theorem 4.12 (Hille-Yosida) *A closed densely defined operator L on the Banach space $\mathcal{B}$ is the generator of a $\mathcal{C}_0$-semigroup $S(t)$ with $\|S(t)\| \le Me^{at}$ if and only if all λ with $\mathrm{Re}\lambda > a$ lie in its resolvent set and the bound $\|R_\lambda^n\| \le M\,(\mathrm{Re}\lambda - a)^{-n}$ holds there for every $n \ge 1$.*

Remark 4.13 Such operators are also called m-dissipative, the reason being that on a Hilbert space, they are precisely those operators such that $\langle h, Lh\rangle \le a\|h\|^2$ and having no extension with the same property. There also exists a version of this statement that holds for Banach spaces. This is the content of the Lumer-Phillips theorem.

Proof. The generator L of a $\mathcal{C}_0$-semigroup is closed by Proposition 4.11. The fact that its resolvent satisfies the stated bound follows immediately from the fact that

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

by Proposition 4.10.

To show that the converse also holds, we are going to construct the semigroup $S(t)$ by using the so-called 'Yosida approximations' $L_\lambda = \lambda L R_\lambda$ for L. Note first that $\lim_{\lambda\to\infty} LR_\lambda x = 0$ for every $x \in \mathcal{B}$: it obviously holds for $x \in \mathcal{D}(L)$ since then $\|LR_\lambda x\| = \|R_\lambda Lx\| \le \|R_\lambda\|\|Lx\| \le M\,(\mathrm{Re}\lambda - a)^{-1}\|Lx\|$. Furthermore, $\|LR_\lambda x\| = \|\lambda R_\lambda x - x\| \le M\lambda(\lambda - a)^{-1} + 1 \le M + 2$ for λ large enough, so that $\lim_{\lambda\to\infty} LR_\lambda x = 0$ for every x by a standard density argument.

Using this fact, we can show that the Yosida approximation of L does indeed approximate L in the sense that $\lim_{\lambda\to\infty} L_\lambda x = Lx$ for every $x \in \mathcal{D}(L)$. Fixing an arbitrary $x \in \mathcal{D}(L)$, we have

$$\lim_{\lambda\to\infty} \|L_\lambda x - Lx\| = \lim_{\lambda\to\infty} \|(\lambda R_\lambda - 1)Lx\| = \lim_{\lambda\to\infty} \|LR_\lambda Lx\| = 0 \;.$$

Define now a family of bounded operators $S_\lambda(t)$ by $S_\lambda(t) = e^{L_\lambda t} = \sum_{n\ge 0} \frac{t^n L_\lambda^n}{n!}$. This series converges in the operator norm since L_λ is bounded and one can easily check that S_λ is indeed a $\mathcal{C}_0$-semigroup (actually a group) with generator L_λ. Since $L_\lambda = -\lambda + \lambda^2 R_\lambda$, one has for $\lambda > a$ the bound

$$\|S_\lambda(t)\| = e^{-\lambda t} \sum_{n\ge 0} \frac{t^n \lambda^{2n} \|R_\lambda^n\|}{n!} = M \exp\Bigl(-\lambda t + \frac{\lambda^2}{\lambda - a} t\Bigr) = M \exp\Bigl(\frac{\lambda a t}{\lambda - a}\Bigr) \;, \tag{4.3}$$

so that $\limsup_{\lambda\to\infty} \|S_\lambda(t)\| \le Me^{at}$. Let us show next that the limit $\lim_{\lambda\to\infty} S_\lambda(t)x$ exists for every $t \ge 0$ and every $x \in \mathcal{B}$. Fixing λ and μ large enough so that $\max\{\|S_\lambda(t)\|, \|S_\mu(t)\|\} \le Me^{2at}$, and fixing some arbitrary $t > 0$, we have for $s \in [0,t]$

$$\begin{aligned}\|\partial_s S_\lambda(t-s)S_\mu(s)x\| &= \|S_\lambda(t-s)(L_\mu - L_\lambda)S_\mu(s)x\| = \|S_\lambda(t-s)S_\mu(s)(L_\mu - L_\lambda)x\| \\ &\le M^2 e^{2at}\|(L_\mu - L_\lambda)x\| \;.\end{aligned}$$

Integrating this bound between 0 and t, we obtain

$$\|S_\lambda(t)x - S_\mu(t)x\| \le M^2 t e^{2at}\|L_\mu x - L_\lambda x\| \;, \tag{4.4}$$

which converges to 0 for every $x \in \mathcal{D}(L)$ as $\lambda, \mu \to \infty$ since one then has $L_\lambda x \to Lx$. We can therefore *define* a family of linear operators $S(t)$ by $S(t)x = \lim_{\lambda\to\infty} S_\lambda(t)x$.

It is clear from (4.3) that $\|S(t)\| \le Me^{at}$ and it follows from the semigroup property of S_λ that $S(s)S(t) = S(s+t)$. Furthermore, it follows from (4.4) that the convergence $S_\lambda(t)x \to S(t)x$ is uniform in bounded intervals of t, so that $(x,t) \mapsto S(t)x$ is jointly continuous, showing that S is a $\mathcal{C}_0$-semigroup. It remains to show that the generator $\hat{L}$ of S coincides with L. Taking first the limit $\lambda \to \infty$ and then the limit $t \to 0$ in the identity

$$t^{-1}(S_\lambda(t)x - x) = t^{-1}\int_0^t S_\lambda(s)L_\lambda x\,ds \;,$$

we see that $x \in \mathcal{D}(L)$ implies $x \in \mathcal{D}(\hat{L})$ and $\hat{L}x = Lx$, so that $\hat{L}$ is an extension of L. However, for $\lambda > a$, both $\lambda - L$ and $\lambda - \hat{L}$ are one-to-one between their domain and $\mathcal{B}$, so that they must coincide. □

4.2 Semigroups with selfadjoint generators

In this section, we consider the particular case of selfadjoint semigroups on a Hilbert space $\mathcal{H}$. Let L be a selfadjoint operator on $\mathcal{H}$ which is bounded from above. Without loss of generality, we are going to assume that it is actually negative definite, so that $\langle x, Lx\rangle \le 0$ for any $x \in \mathcal{H}$. In this case, we can use functional calculus (see for example [RS80], in particular chapter VIII in volume I) to define selfadjoint operators $f(L)$ for any measurable map $f\colon \mathbf{R} \to \mathbf{R}$. This is because the spectral decomposition theorem can be formulated as:

Theorem 4.14 (Spectral decomposition) *Let L be a selfadjoint operator on a separable Hilbert space $\mathcal{H}$. Then, there exists a measure space $(\mathcal{M},\mu)$, an isomorphism $K\colon \mathcal{H} \to L^2(\mathcal{M},\mu)$, and a function $f_L\colon \mathcal{M} \to \mathbf{R}$ such that via K, L is equivalent to the multiplication operator by f_L on $L^2(\mathcal{M},\mu)$. In other words, one has $L = K^{-1}f_L K$ and $K\mathcal{D}(L) = \{g \,:\, f_L g \in L^2(\mathcal{M},\mu)\}$.*

In particular, this allows one to define $f(L) = K^{-1}(f \circ f_L)K$, which has all the nice properties that one would expect from functional calculus, like for example $(fg)(L) = f(L)g(L)$, $\|f(L)\| = \|f\|_{L^\infty(\mathcal{M},\mu)}$, etc. Defining $S(t) = e^{Lt}$, it is an exercise to check that S is indeed a $\mathcal{C}_0$-semigroup with generator L (either use the Hille-Yosida theorem and make sure that the semigroup constructed there coincides with S or check 'by hand' that $S(t)$ is indeed $\mathcal{C}_0$ with generator L).

The important property of semigroups generated by selfadjoint operators is that they do not only leave $\mathcal{D}(L)$ invariant, but they have a regularising effect in that they map $\mathcal{H}$ into the domain of any arbitrarily high power of L. More precisely, one has:

Proposition 4.15 *Let L be self-adjoint and negative definite and let $S(t)$ be the semigroup on $\mathcal{H}$ generated by L. Then, $S(t)$ maps $\mathcal{H}$ into the domain of $(1-L)^\alpha$ for any $\alpha, t > 0$ and there exist constants C_α such that $\|(1-L)^\alpha S(t)\| \le C_\alpha(1 + t^{-\alpha})$.*

Proof. By functional calculus, it suffices to show that $\sup_{\lambda\le 0}(1-\lambda)^\alpha e^{\lambda t} \le C_\alpha(1+t^{-\alpha})$. One has

$$\sup_{\lambda\ge 0} \lambda^\alpha e^{-\lambda t} = t^{-\alpha} \sup_{\lambda\ge 0} (\lambda t)^\alpha e^{-\lambda t} = t^{-\alpha} \sup_{\lambda \ge 0} \lambda^\alpha e^{-\lambda} = \alpha^\alpha e^{-\alpha} t^{-\alpha} .$$

The claim now follows from the fact that there exists a constant C_α' such that $(1-\lambda)^\alpha \le C_\alpha'(1 + (-\lambda)^\alpha)$ for every $\lambda \le 0$. □

4.3 Analytic semigroups

Obviously, the conclusion of Proposition 4.15 does not hold for arbitrary $\mathcal{C}_0$-semigroups since the group of translations from Example 4.8 does not have any smoothing properties. It does however hold for analytic semigroups, which is going to be one of the two main results of this subsection. The other result is a characterisation of generators for analytic semigroups that is analogous to the Hille-Yosida theorem for $\mathcal{C}_0$-semigroups. The difference will be that the role of the half-plane $\mathrm{Re}\lambda > a$ will be played by the complement of a sector of the complex plane with an opening angle strictly smaller than π.

We are going to take it for granted in this section that the resolvent $\lambda \mapsto R_\lambda = (\lambda - L)^{-1}$ is analytic in $\mathcal{L}(\mathcal{B})$ for λ in the resolvent set of L. For a proof of this fact and several other properties of the resolvent, see for example [Yos95].

Recall that a semigroup is analytic if there exists $\theta > 0$ such that the map $t \mapsto S(t)$ has an analytic extension to the sector $S_\theta = \{\lambda \in \mathbf{C} \,:\, |\arg \lambda| < \theta\}$, satisfies the semigroup property there, and is such that $t \mapsto S_\varphi(t) = S(e^{i\varphi}t)$ is a strongly continuous semigroup for every $|\varphi| < \theta$. If θ is the largest angle such that the above property holds, we call S *analytic with angle* θ. The strong continuity of $t \mapsto S(e^{i\varphi}t)$ implies that there exist constants $M(\varphi)$ and $a(\varphi)$ such that

$$\|S_\varphi(t)\| \le M(\varphi) e^{a(\varphi)t} .$$

Using the semigroup property, it is not difficult to show that M and a can be chosen bounded over compact intervals:

Proposition 4.16 *Let S be an analytic semigroup with angle θ. Then, for every $\theta' < \theta$, there exist M and a such that $\|S_\varphi(t)\| \le M e^{a|t|}$ for every $t > 0$ and every $|\varphi| \le \theta'$.*

Proof. Fix $\theta' > 0$. Then, for every $t > 0$ and every φ with $|\varphi| < \theta$, there exist numbers $t_+, t_- \in [0,t]$ such that $te^{i\varphi} = t_+ e^{i\theta'} + t_- e^{-i\theta'}$. It follows that one has the bound $\|S_\varphi(t)\| \le M(\theta')M(-\theta')e^{a(\theta')t + a(-\theta')t}$, thus proving the claim. □

We next compute the generators of the semigroups S_φ obtained by evaluating S along a 'ray' extending out of the origin into the complex plane:

Proposition 4.17 *Let S be an analytic semigroup with angle θ. Then, for $|\varphi| < \theta$, the generator L_φ of S_φ is given by $L_\varphi = e^{i\varphi}L$, where L is the generator of S.*

Proof. Recall Proposition 4.10 showing that for $\mathrm{Re}\lambda$ large enough the resolvent R_λ for L is given by

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

Since the map $t \mapsto e^{-\lambda t} S(t)$ is analytic in S_θ by assumption and since, provided again that $\mathrm{Re}\lambda$ is large enough, it decays exponentially to 0 as $|t| \to \infty$, we can deform the contour of integration to obtain

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

Denoting by R^φ_λ the resolvent for the generator L_φ of S_φ, we thus have the identity $R_\lambda = e^{i\varphi} R^\varphi_{\lambda e^{i\varphi}}$, which is equivalent to $(\lambda - L)^{-1} = (\lambda - e^{-i\varphi} L_\varphi)^{-1}$, thus showing that $L_\varphi = e^{i\varphi} L$ as stated. □

We now use this to show that if S is an analytic semigroup, then the resolvent set of its generator L not only contains the right half plane, but it contains a larger sector of the complex plane. Furthermore, this characterises the generators of analytic semigroups, providing a statement similar to the Hille-Yosida theorem:

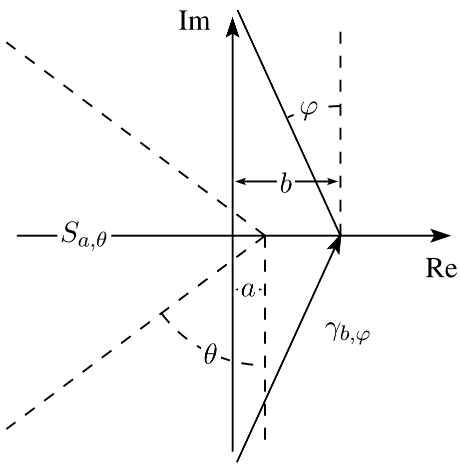
Theorem 4.18 *A closed densely defined operator L on a Banach space $\mathcal{B}$ is the generator of an analytic semigroup if and only if there exists $\theta \in (0, \frac{\pi}{2})$ and $a \geq 0$ such that the spectrum of L is contained in the sector*

$$S_{\theta,a} = \{\lambda \in \mathbf{C} \,:\, \arg(a - \lambda) \in [-\tfrac{\pi}{2} + \theta, \tfrac{\pi}{2} - \theta]\}\;,$$

and there exists $M > 0$ such that the resolvent R_λ satisfies the bound $\|R_\lambda\| \leq M d(\lambda, S_{\theta,a})^{-1}$ for every $\lambda \notin S_{\theta,a}$.

Proof. The fact that generators of analytic semigroups are of the prescribed form is a consequence of Proposition 4.17 and the Hille-Yosida theorem.

To show the converse statement, let L be such an operator, let $\varphi \in (0, \theta)$, let $b > a$, and let $\gamma_{\varphi,b}$ be the curve in the complex plane obtained by going in a counterclockwise way around the boundary of $S_{\varphi,b}$ (see the figure on the right). For t with $|\arg t| < \varphi$, define $S(t)$ by



$$\begin{aligned} S(t) &= \frac{1}{2\pi i}\int_{\gamma_{\varphi,b}} e^{tz} R_z\,dz \qquad (4.5)\\ &= \frac{1}{2\pi i}\int_{\gamma_{\varphi,b}} e^{tz}(z - L)^{-1}\,dz\;. \end{aligned}$$

It follows from the resolvent bound that $\|R_z\|$ is uniformly bounded for $z \in \gamma_{\varphi,b}$. Furthermore, since $|\arg t| < \varphi$, it follows that e^{tz} decays exponentially as $|z| \to \infty$ along $\gamma_{\varphi,b}$, so that this expression is well-defined, does not depend on the choice of b and φ, and (by choosing φ arbitrarily close to θ) determines an analytic function $t \mapsto S(t)$ on the sector $\{t \,:\, |\arg t| < \theta\}$. As in the proof of the Hille-Yosida theorem, the function $(x,t) \mapsto S(t)x$ is jointly continuous because the convergence of the integral defining S is uniform over bounded subsets of $\{t \,:\, |\arg t| < \varphi\}$ for any $|\varphi| < \theta$.

It therefore remains to show that S satisfies the semigroup property on the sector $\{t \,:\, |\arg t| < \theta\}$ and that its generator is indeed given by L. Choosing s and t such that $|\arg s| < \theta$ and $|\arg t| < \theta$ and using the resolvent identity $R_z - R_{z'} = (z' - z)R_z R_{z'}$, we have

$$\begin{aligned} S(s)S(t) &= -\frac{1}{4\pi^2}\int_{\gamma_{\varphi,b'}}\int_{\gamma_{\varphi,b}} e^{tz+sz'} R_z R_{z'}\,dz\,dz' = -\frac{1}{4\pi^2}\int_{\gamma_{\varphi,b'}}\int_{\gamma_{\varphi,b}} e^{tz+sz'}\frac{R_z - R_{z'}}{z' - z}\,dz\,dz'\\ &= -\frac{1}{4\pi^2}\int_{\gamma_{\varphi,b}} e^{tz} R_z \int_{\gamma_{\varphi,b'}} \frac{e^{sz'}}{z' - z}\,dz'\,dz - \frac{1}{4\pi^2}\int_{\gamma_{\varphi,b'}} e^{sz} R_z \int_{\gamma_{\varphi,b}} \frac{e^{tz'}}{z' - z}\,dz'\,dz\;. \end{aligned}$$

Here, the choice of b and b' is arbitrary, as long as $b \neq b'$ so that the inner integrals are well-defined, say $b' > b$ for definiteness. In this case, since the contour $\gamma_{\varphi,b}$ can be 'closed up' to the left but not to the right, the integral $\int_{\gamma_{\varphi,b'}} \frac{e^{sz'}}{z'-z}\,dz'$ is equal to $-2i\pi e^{sz}$ for every $z \in \gamma_{\varphi,b}$, whereas the integral with b and b' inverted vanishes, so that

$$S(s)S(t) = \frac{1}{2i\pi}\int_{\gamma_{\varphi,b}} e^{(t+s)z}R_z = S(s+t)\;,$$

as required. Therefore S is a strongly continuous semigroup; let us call its generator $\hat{L}$ and $\hat{R}_\lambda$ the corresponding resolvent.

To show that $L = \hat{L}$, it suffices to show that $\hat{R}_\lambda = R_\lambda$, so we make use again of Proposition 4.10. Choosing $\mathrm{Re}\lambda > b$ so that $\mathrm{Re}(z-\lambda) < 0$ for every $z \in \gamma_{\varphi,b}$, we have

$$\begin{aligned}\hat{R}_\lambda &= \int_0^\infty e^{-\lambda t}S(t)\,dt = \frac{1}{2\pi i}\int_0^\infty\int_{\gamma_{\varphi,b}} e^{t(z-\lambda)}R_z\,dz\,dt \\ &= \frac{1}{2\pi i}\int_{\gamma_{\varphi,b}}\int_0^\infty e^{t(z-\lambda)}\,dt\,R_z\,dz = \frac{1}{2\pi i}\int_{\gamma_{\varphi,b}}\frac{R_z}{z-\lambda}\,dz = R_\lambda\;.\end{aligned}$$

The last inequality was obtained by using the fact that $\|R_z\|$ decays like $1/|z|$ for large enough z with $|\arg z| \le \frac{\pi}{2} + \varphi$, so that the contour can be 'closed' to enclose the pole at $z = \lambda$. □

Remark 4.19 The generators of analytic semigroups are also called m-sectorial by analogy with the m-dissipative operators generating $\mathcal{C}_0$-semigroups. On a Hilbert space, they can also be characterised as those operators with numerical range contained in a sector $S_{\theta,a}$ and having no extension with the same property.

As a consequence of this characterisation theorem, we can study perturbations of generators of analytic semigroups. The idea is to give a constructive criterion which allows to make sure that an operator of the type $L = L_0 + B$ is the generator of an analytic semigroup, provided that L_0 is such a generator and B satisfies a type of 'relative total boundedness' condition. The precise statement of this result is:

Theorem 4.20 *Let L_0 be the generator of an analytic semigroup and let $B\colon \mathcal{D}(B) \to \mathcal{B}$ be an operator such that*

- *The domain $\mathcal{D}(B)$ contains $\mathcal{D}(L_0)$.*
- *For every $\varepsilon > 0$ there exists $C > 0$ such that $\|Bx\| \le \varepsilon\|L_0x\| + C\|x\|$ for every $x \in \mathcal{D}(L_0)$.*

Then the operator $L = L_0 + B$ (with domain $\mathcal{D}(L) = \mathcal{D}(L_0)$) is also the generator of an analytic semigroup.

Proof. In view of Theorem 4.18 it suffices to show that there exists a sector $S_{\theta,a}$ containing the spectrum of L and such that the resolvent bound $R_\lambda \le Md(\lambda, S_{\theta,a})^{-1}$ holds away from it.

Denote by R_λ^0 the resolvent for L_0 and consider the resolvent equation for L:

$$(\lambda - L_0 - B)x = y\;,\quad x \in \mathcal{D}(L_0)\;.$$

Since (at least for λ outside of some sector) x belongs to the range of R_λ^0, we can set $x = R_\lambda^0 z$ so that this equation is equivalent to

$$z - BR_\lambda^0 z = y\;.$$

The claim therefore follows if we can show that there exists a sector $S_{\theta,a}$ and a constant $c < 1$ such that $\|BR^0_\lambda\| \le c$ for $\lambda \notin S_{\theta,a}$. This is because one then has the bound

$$\|R_\lambda y\| = \|R^0_\lambda z\| \le \frac{\|R^0_\lambda\|}{1-c}\|y\| \;.$$

Using our assumption on B, we have the bound

$$\|BR^0_\lambda z\| \le \varepsilon\|L_0 R^0_\lambda z\| + C\|R^0_\lambda z\| \;. \tag{4.6}$$

Furthermore, one has the identity $L_0 R^0_\lambda = \lambda R^0_\lambda - 1$ and, since L_0 is the generator of an analytic semigroup by assumption, the resolvent bound $\|R^0_\lambda\| \le M d(\lambda, S_{\alpha,b})^{-1}$ for some parameters α, b. Inserting this into (4.6), we obtain the bound

$$\|BR^0_\lambda\| \le \frac{(\varepsilon|\lambda| + C)M}{d(\lambda, S_{\alpha,b})} + \varepsilon \;.$$

Note now that by choosing $\theta \in (0,\alpha)$, we can find some $\delta > 0$ such that $d(\lambda, S_{\alpha,b}) > \delta|\lambda|$ for all $\lambda \notin S_{\theta,a}$ and all $a > 1 \vee (b+1)$. We fix such a θ and we make ε sufficiently small such that one has both $\varepsilon < 1/4$ and $\varepsilon\delta^{-1} < 1/4$.

We can then make a large enough so that $d(\lambda, S_{\alpha,b}) \ge 4CM$ for $\lambda \notin S_{\theta,a}$, so that $\|BR^0_\lambda\| \le 3/4$. for these values of λ, as requested. □

Remark 4.21 As one can see from the proof, one actually needs the bound $\|Bx\| \le \varepsilon\|L_0 x\| + C\|x\|$ only for some particular value of ε that depends on the characteristics of L_0.

As a consequence, we have:

Proposition 4.22 *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 $\mathcal{D}(L) = H^2$ is the generator of an analytic semigroup.

Proof. It is well-known that the operator $(L_0 g)(x) = g''(x)$ with domain $\mathcal{D}(L) = H^2$ is self-adjoint and negative definite, so that it is the generator of an analytic semigroup with angle $\theta = \pi/2$.

Setting $Bg = fg'$, we have for $g \in H^2$ the bound

$$\|Bg\|^2 = \int_{\mathbf{R}} f^2(x)(g'(x))^2\,dx \le \|f\|^2_{L^\infty}\langle g', g'\rangle = -\|f\|^2_{L^\infty}\langle g, g''\rangle \le \|f\|_{L^\infty}\|g\|\|L_0 g\| \;.$$

It now suffices to use the fact that $2|xy| \le \varepsilon x^2 + \varepsilon^{-1}y^2$ to conclude that the assumptions of Theorem 4.20 are satisfied. □

Similarly, one can show:

Exercise 4.23 Show that the generator of an elliptic diffusion with smooth coefficients on a compact Riemannian manifold $\mathcal{M}$ generates an analytic semigroup on $L^2(\mathcal{M}, \varrho)$, where ϱ is the volume measure given by the Riemannian structure.

4.4 Interpolation spaces

The remainder of this section will be devoted to the study of the domains of fractional powers of the generator L of an analytic semigroup $S(t)$. For simplicity, we will assume that there exist $M > 0$ and $w > 0$ such that $\|S(t)\| \le Me^{-wt}$, thus making sure that the resolvent set of L contains all the right half of the complex plane. The general case can be recovered easily by 'shifting the generator to the left'. For $\alpha > 0$, we *define* negative fractional powers of L by

$$(-L)^{-\alpha} \stackrel{\text{def}}{=} \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} S(t)\, dt\;, \tag{4.7}$$

which is a bounded operator by the decay assumption on $\|S(t)\|$. Since $\Gamma(1) = 1$, note that if $\alpha = 1$ one does indeed recover the resolvent of L evaluated at 0. Furthermore, it is straightforward to check that one has the identity $(-L)^{-\alpha}(-L)^{-\beta} = (-L)^{-\alpha-\beta}$, which together justify the definition (4.7).

Note that it follows from this identity that $(-L)^{-\alpha}$ is injective for every $\alpha > 0$. Indeed, given some $\alpha > 0$, one can find an integer $n > 0$ such that $(-L)^{-n} = (-L)^{-n+\alpha}(-L)^{-\alpha}$. A failure for $(-L)^{-\alpha}$ to be injective would therefore result in a failure for $(-L)^n$ and therefore $(-L)^{-1}$ to be injective. This is ruled out by the fact that 0 belongs to the resolvent set of L. We can therefore define $(-L)^\alpha$ as the unbounded operator with domain $\mathcal{D}((-L)^\alpha) = \text{range}(-L)^{-\alpha}$ given by the inverse of $(-L)^{-\alpha}$. This definition is again consistent with the usual definition of $(-L)^\alpha$ for integer values of α. This allows us to set:

Definition 4.24 For $\alpha > 0$ and given an analytic semigroup S on a Banach space $\mathcal{B}$, we define the *interpolation space* $\mathcal{B}_\alpha$ as the domain of $(-L)^\alpha$ endowed with the graph norm. We also define $\mathcal{B}_{-\alpha}$ as the closure of $\mathcal{B}$ under the norm $\|x\|_{-\alpha} = \|(-L)^{-\alpha}x\|$.

Remark 4.25 If the norm of $S(t)$ grows instead of decaying with t, then we use $\lambda - L$ instead of $-L$ for some λ sufficiently large. The choice of different values of λ leads to equivalent norms on $\mathcal{B}_\alpha$.

Exercise 4.26 Show that $\mathcal{B}_\alpha \subset \mathcal{B}_\beta$ if $\alpha \ge \beta$.

Exercise 4.27 Show that for $\alpha \in (0,1)$ and $x \in \mathcal{D}(L)$, one has the identity

$$(-L)^\alpha x = \frac{\sin \alpha\pi}{\pi} \int_0^\infty t^{\alpha-1}(t-L)^{-1}(-L)x\, dt\;.$$

Combine this with the identity $(t-L)^{-1}(-L) = 1 - t(t-L)^{-1}$ to deduce that, for every $\alpha \in (0,1)$, there exists a constant C such that the bound $\|(-L)^\alpha x\| \le C\|Lx\|^\alpha \|x\|^{1-\alpha}$ holds for every $x \in \mathcal{D}(L)$.

Exercise 4.28 Let L and B be as in Exercise 4.35 and denote by S_B the analytic semigroup with generator $L + B$. Use the relation $R_\lambda - R_\lambda^0 = R_\lambda^0 B R_\lambda$ to show that one has the identity

$$S_B(t)x = S(t)x + \int_0^t S(t-s)BS_B(s)x\, ds\;.$$

Hint: Start from the right hand side of the equation and use an argument similar to that of the proof of Theorem 4.18.

Exercise 4.29 Show that $(-L)^\alpha$ commutes with $S(t)$ for every $t > 0$ and every $\alpha \in \mathbf{R}$. Deduce that $S(t)$ leaves $\mathcal{B}_\alpha$ invariant for every $\alpha > 0$.

Exercise 4.30 It follows from Theorem 4.18 that the adjoint L^* of the generator of an analytic semigroup on $\mathcal{B}$ is the generator of an analytic semigroup on the dual space $\mathcal{B}^*$. Denote by $\tilde{\mathcal{B}}_\alpha$ the corresponding interpolation spaces. Show that one has the identity $\tilde{\mathcal{B}}_\alpha = \mathcal{B}_\alpha^*$ for every $\alpha \in \mathbf{R}$.

We now show that an analytic semigroup $S(t)$ always maps $\mathcal{B}$ into $\mathcal{B}_\alpha$ for $t > 0$, so that it has a 'smoothing effect'. Furthermore, the norm in the domains of integer powers of L can be bounded by:

Proposition 4.31 *For every $t > 0$ and every integer $k > 0$, $S(t)$ maps $\mathcal{B}$ into $\mathcal{D}(L^k)$ and there exists a constant C_k such that*

$$\|L^k S(t)x\| \le \frac{C_k}{t^k}$$

for every $t \in (0,1]$.

Proof. In order to show that S maps $\mathcal{B}$ into the domain of every power of L, we use (4.5), together with the identity $LR_\lambda = \lambda R_\lambda - 1$ which is an immediate consequence of the definition of the resolvent R_λ of L. Since $\int_{\gamma_{\varphi,b}} e^{tz}dz = 0$ for every t such that $|\arg t| < \varphi$ and since the domain of L^k is complete under the graph norm, this shows that $S(t)x \in \mathcal{D}(L^k)$ and

$$L^k S(t) = \frac{1}{2\pi i}\int_{\gamma_{\varphi,b}} z^k e^{tz} R_z\, dz\ .$$

It follows that there exist positive constants c_i such that

$$\|L^k S(t)\| \le \frac{1}{2\pi}\int_{\gamma_{\varphi,b}} |z|^k |e^{tz}| \|R_z\|\, d|z| \le c_1 \int_0^\infty (1+x)^k e^{-c_2 t(x-c_3)}(1+x)^{-1} dx\ .$$

Integrating by parts $k-1$ times, we obtain

$$\|L^k S(t)\| \le \frac{c_4}{t^{k-1}}\int_0^\infty e^{-c_2 t(x-c_4)}\, dx = \frac{c_5 e^{c_6 t}}{t^k}\ ,$$

which implies the announced bound. □

It turns out that a similar bound also holds for interpolation spaces with non-integer indices:

Proposition 4.32 *For every $t > 0$ and every $\alpha > 0$, $S(t)$ maps $\mathcal{B}$ into $\mathcal{B}_\alpha$ and there exists a constant C_α such that*

$$\|(-L)^\alpha S(t)x\| \le \frac{C_\alpha}{t^\alpha} \tag{4.8}$$

for every $t \in (0,1]$.

Proof. The fact that $S(t)$ maps $\mathcal{B}$ into $\mathcal{B}_\alpha$ follows from Proposition 4.31 since there exists n such that $\mathcal{D}(L^n) \subset \mathcal{B}_\alpha$. We assume again that the norm of $S(t)$ decays exponentially for large t. The claim for integer values of α is known to hold by Proposition 4.31, so we fix some $\alpha > 0$ which is *not* an integer. Note first that $(-L)^\alpha = (-L)^{\alpha-[\alpha]-1}(-L)^{[\alpha]+1}$, were we denote by $[\alpha]$ the integer part of α. We thus obtain from (4.7) the identity

$$(-L)^\alpha S(t) = \frac{(-1)^{[\alpha]+1}}{\Gamma([\alpha]-\alpha+1)}\int_0^\infty s^{[\alpha]-\alpha} L^{[\alpha]+1} S(t+s)\, ds\ .$$

Using the previous bound for $k = [\alpha]$, we thus get for some $C > 0$ the bound

$$\|(-L)^\alpha S(t)\| \le C \int_0^\infty s^{[\alpha]-\alpha} \frac{e^{-w(t+s)}}{(t+s)^{[\alpha]+1}}\, ds \le Ct^{-\alpha} \int_0^\infty \frac{s^{[\alpha]-\alpha}}{(1+s)^{[\alpha]+1}}\, ds\;,$$

where we used the substitution $s \mapsto ts$. Since the last function is integrable for every $\alpha > 0$, the claim follows at once. □

Exercise 4.33 Using the fact that $S(t)$ commutes with any power of its generator, show that $S(t)$ maps $\mathcal{B}_\alpha$ into $\mathcal{B}_\beta$ for every $\alpha, \beta \in \mathbf{R}$ and that, for $\beta > \alpha$, there exists a constant $C_{\alpha,\beta}$ such that $\|S(t)x\|_{\mathcal{B}_\beta} \le C_{\alpha,\beta}\|x\|_{\mathcal{B}_\alpha} t^{\alpha-\beta}$ for all $t \in (0, 1]$.

Exercise 4.34 Using the bound from the previous exercise and the definition of the resolvent, show that for every $\alpha \in \mathbf{R}$ and every $\beta \in [\alpha, \alpha+1)$ there exists a constant C such that the bound $\|(t-L)^{-1}x\|_{\mathcal{B}_\beta} \le C(1+t)^{\beta-\alpha-1}\|x\|_{\mathcal{B}_\alpha}$ holds for all $t \ge 0$.

Exercise 4.35 Let L be the generator of an analytic semigroup on $\mathcal{B}$ and denote by $\mathcal{B}_\alpha$ the corresponding interpolation spaces. Let B be a (possibly unbounded) operator on $\mathcal{B}$. Using the results from the previous exercise, show that if there exists $\alpha \in [0, 1)$ such that $\mathcal{B}_\alpha \subset \mathcal{D}(B)$ so that B is a bounded operator from $\mathcal{B}_\alpha$ to $\mathcal{B}$, then one has the bound

$$\|Bx\| \le C(\varepsilon\|Lx\| + \varepsilon^{-\alpha/(1-\alpha)}\|x\|)\;,$$

for some constant $C > 0$ and for all $\varepsilon \le 1$. In particular, $L + B$ is also the generator of an analytic semigroup on $\mathcal{B}$.

Exercise 4.36 Consider an analytic semigroup $S(t)$ on $\mathcal{B}$ and denote by $\mathcal{B}_\alpha$ the corresponding interpolation spaces. Fix some $\gamma \in \mathbf{R}$ and denote by $\hat S(t)$ the semigroup S viewed as a semigroup on $\mathcal{B}_\gamma$. Denoting by $\hat{\mathcal{B}}_\alpha$ the interpolation spaces corresponding to $\hat S(t)$, show that one has the identity $\hat{\mathcal{B}}_\alpha = \mathcal{B}_{\gamma+\alpha}$ for every $\alpha \in \mathbf{R}$.

Another question that can be answered in a satisfactory way with the help of interpolation spaces is the speed of convergence of $S(t)x$ to x as $t \to 0$. We know that if $x \in \mathcal{D}(L)$, then $t \mapsto S(t)x$ is differentiable at $t = 0$, so that $\|S(t)x - x\| = t\|Lx\| + o(t)$. Furthermore, one can in general find elements $x \in \mathcal{B}$ so that the convergence $S(t)x \to x$ is arbitrarily slow. This suggests that if $x \in \mathcal{D}((-L)^\alpha)$ for $\alpha \in (0, 1)$, one has $\|S(t)x - x\| = \mathcal{O}(t^\alpha)$. This is indeed the case:

Proposition 4.37 *Let S be an analytic semigroup with generator L on a Banach space $\mathcal{B}$. Then, for every $\alpha \in (0, 1)$, there exists a constant C_α, so that the bound*

$$\|S(t)x - x\| \le C_\alpha t^\alpha \|x\|_{\mathcal{B}_\alpha} \tag{4.9}$$

holds for every $x \in \mathcal{B}_\alpha$ and every $t \in (0, 1]$.

Proof. By density, it is sufficient to show that (4.9) holds for every $x \in \mathcal{D}(L)$. For such an x, one has indeed the chain of inequalities

$$\begin{aligned}\|S(t)x - x\| &= \Big\|\int_0^t S(s)Lx\, dx\Big\| = \Big\|\int_0^t (-L)^{1-\alpha}S(s)(-L)^\alpha x\, dx\Big\| \\ &\le C\|x\|_{\mathcal{B}_\alpha}\int_0^t \|(-L)^{1-\alpha}S(s)\|\, dx \le C\|x\|_{\mathcal{B}_\alpha}\int_0^t s^{\alpha-1}\, ds = C\|x\|_{\mathcal{B}_\alpha} t^\alpha\;.\end{aligned}$$

Here, the constant C depends only on α and changes from one expression to the next. □

We conclude this section with a discussion on the interpolation spaces arising from a perturbed analytic semigroup. As a consequence of Exercises 4.27 and 4.34, we have the following result:

Proposition 4.38 *Let L_0 be the generator of an analytic semigroup on $\mathcal{B}$ and denote by $\mathcal{B}_\gamma^0$ the corresponding interpolation spaces. Let B be a bounded operator from $\mathcal{B}_\alpha^0$ to $\mathcal{B}$ for some $\alpha \in [0,1)$. Let furthermore $\mathcal{B}_\gamma$ be the interpolation spaces associated to $L = L_0 + B$. Then, one has $\mathcal{B}_\gamma = \mathcal{B}_\gamma^0$ for every $\gamma \in [0,1]$.*

Proof. The statement is clear for $\gamma = 0$ and $\gamma = 1$. For intermediate values of γ, we will show that there exists a constant C such that $C^{-1}\|(-L_0)^\gamma x\| \le \|(-L)^\gamma x\| \le C\|(-L_0)^\gamma x\|$ for every $x \in \mathcal{D}(L_0)$.

Since the domain of L is equal to the domain of L_0, we know that the operator BR_t is bounded for every $t > 0$, where R_t is the resolvent of L. Making use of the identity

$$R_t = R_t^0 + R_t^0 B R_t\;, \tag{4.10}$$

(where we similarly denoted by R_t^0 the resolvent of L_0) it then follows from Exercise 4.34 and the assumption on B that one has for every $x \in \mathcal{B}_\gamma^0$ the bound

$$\begin{aligned}\|BR_t x\| &\le \|BR_t^0 x\| + \|BR_t^0 BR_t x\| \le C(\|R_t^0 x\|_{\mathcal{B}_\alpha^0} + \|R_t^0 BR_t x\|_{\mathcal{B}_\alpha^0})\\ &\le C(1+t)^{\alpha-\gamma-1}\|x\|_{\mathcal{B}_\gamma^0} + C(1+t)^{\alpha-1}\|x\|_{\mathcal{B}_\gamma^0}\|BR_t x\|\;.\end{aligned}$$

It follows that, for t sufficiently large, one has the bound

$$\|BR_t x\| \le C(1+t)^{\alpha-\gamma-1}\|x\|_{\mathcal{B}_\gamma^0}\;. \tag{4.11}$$

Since on the other hand one has the resolvent identity $R_s = R_t + (t-s)R_sR_t$, this bound can be extended to all $t > 0$ by possibly changing the constant C.

We now show that $\|(-L)^\gamma x\|$ can be bounded by $\|(-L_0)^\gamma x\|$. We make use of Exercise 4.27 to get, for $x \in \mathcal{D}(L_0)$, the bound

$$\begin{aligned}\|x\|_{\mathcal{B}_\gamma} &= C\Big\|\int_0^\infty t^{\gamma-1} LR_t x\,dt\Big\|\\ &\le C\Big\|\int_0^\infty t^{\gamma-1} L_0 R_t^0 x\,dt\Big\| + C\int_0^\infty t^{\gamma-1}\|(L_0R_t^0 + 1)BR_t x\|\,dt\\ &\le \|x\|_{\mathcal{B}_\gamma^0} + C\int_0^\infty t^{\gamma-1}\|BR_t x\|\,dt\\ &\le \|x\|_{\mathcal{B}_\gamma^0} + C\int_0^\infty t^{\gamma-1}(1+t)^{\alpha-\gamma-1}\,dt\|x\|_{\mathcal{B}_\gamma^0}\;.\end{aligned}$$

Here, we used again the identity (4.10) to obtain the first inequality and we used (4.11) in the last step. Since this integral converges, we have obtained the required bound.

In order to obtain the converse bound, we have similarly to before

$$\|x\|_{\mathcal{B}_\gamma^0} \le \|x\|_{\mathcal{B}_\gamma} + C\int_0^\infty t^{\gamma-1}\|BR_t x\|\,dt\;.$$

Making use of the resolvent identity, this yields for arbitrary $K > 0$ the bound

$$\|x\|_{\mathcal{B}_\gamma^0} \le \|x\|_{\mathcal{B}_\gamma} + C\int_0^\infty t^{\gamma-1}\|BR_{t+K}x\|\,dt + CK\int_0^\infty t^{\gamma-1}\|BR_K R_t x\|\,dt$$

$$\le \|x\|_{\mathcal{B}_\gamma} + C\int_0^\infty t^{\gamma-1}(t+K)^{\alpha-\gamma-1}\,dt\|x\|_{\mathcal{B}^0_\gamma} + CK\int_0^\infty t^{\gamma-1}(1+t)^{-1}\,dt\|x\|$$
$$\le \|x\|_{\mathcal{B}_\gamma} + CK^{\alpha-1}\|x\|_{\mathcal{B}^0_\gamma} + CK\|x\|\;.$$

By making K sufficiently large, the prefactor of the second term can be made smaller than 1, so that the required bound follows. □

Exercise 4.39 Assume that $\mathcal{B} = \mathcal{H}$ is a Hilbert space and that the antisymmetric part of L is 'small' in the sense that $\mathcal{D}(L^*) = \mathcal{D}(L)$ and, for every $\varepsilon > 0$ there exists a constant C such that $\|(L - L^*)x\| \le \varepsilon\|Lx\| + C\|x\|$ for every $x \in \mathcal{D}(L)$. Show that in this case the space $\mathcal{H}_{-\alpha}$ can be identified with the dual of $\mathcal{H}_\alpha$ (under the pairing given by the scalar product of $\mathcal{H}$) for $\alpha \in [0, 1]$.

It is interesting to note that the range $[0, 1]$ appearing in the statement of Proposition 4.38 is not just a restriction of the technique of proof employed here. There are examples of perturbations of generators of analytic semigroups which induce changes in the corresponding interpolation spaces $\mathcal{B}_\alpha$ for $\alpha \notin [0, 1]$.

Consider for example the case $\mathcal{B} = L^2([0, 1])$ and $L_0 = \Delta$, the Laplacian with periodic boundary conditions. Let now $\delta \in (0, 1)$ be arbitrary and let $g \in \mathcal{B}$ be such that $g \notin \mathcal{B}^0_\delta$. Such an element g exists since Δ is an unbounded operator. Define B as the operator with domain $\mathcal{C}([0, 1]) \subset \mathcal{B}$ given by

$$(Bf)(x) = f'(1/2)g(x)\;. \tag{4.12}$$

It turns out that $\mathcal{B}^0_\alpha \subset \mathcal{C}([0, 1])$ for $\alpha > 3/4$ (see Lemma 6.14), so that the assumptions of Proposition 4.38 are indeed satisfied. Consider now the interpolation spaces of index $1 + \delta$. Since we know that $\mathcal{B}_\delta = \mathcal{B}^0_\delta$, we have the characterisations

$$\mathcal{B}_{1+\delta} = \{f \in \mathcal{D}(\Delta) \,:\, \Delta f + f'(1/2)g \in \mathcal{B}^0_\delta\}\;,$$
$$\mathcal{B}^0_{1+\delta} = \{f \in \mathcal{D}(\Delta) \,:\, \Delta f \in \mathcal{B}^0_\delta\}\;.$$

Since on the other hand $g \notin \mathcal{B}^0_\delta$ by assumption, it follows that $\mathcal{B}_{1+\delta} \cap \mathcal{B}^0_{1+\delta}$ consists of functions with vanishing derivative at $1/2$, so that in particular $\mathcal{B}_{1+\delta} \neq \mathcal{B}^0_{1+\delta}$.

Exercise 4.40 Show that in the example above, one has $\mathcal{B}_{-1/4} \neq \mathcal{B}^0_{-1/4}$. **Hint:** Consider the adjoint of L as the generator of an analytic semigroup on $H^{-2} = (\mathcal{B}^0_1)^*$ and make use of Exercise 4.30, using the fact that $\delta' \notin \mathcal{B}^0_{-3/4}$.

Exercise 4.41 Show, again in the same setting as above, that if $g \in \mathcal{B}^0_\delta$ for some $\delta > 0$, then one has $\mathcal{B}_\alpha = \mathcal{B}^0_\alpha$ for every $\alpha \in [0, 1 + \delta)$.

Remark 4.42 The operator B defined in (4.12) is not a closed operator on $\mathcal{B}$. In fact, it is not even closable! This is however of no consequence for Proposition 4.38 since the operator $L = L_0 + B$ is closed and this is all that matters.

5 Linear SPDEs / Stochastic Convolutions

We now apply the knowledge gathered in the previous sections to discuss the solution to linear stochastic PDEs. Most of the material from this section can be found in the monographs [DPZ92b, DPZ96]. The aim of this section is to define what we mean by the solution to a linear stochastic PDE of the form

$$dx = Lx\,dt + Q\,dW(t)\;, \quad x(0) = x_0\;, \tag{5.1}$$

where we want x to take values in a separable Banach space $\mathcal{B}$, L is the generator of a $\mathcal{C}_0$ semigroup on $\mathcal{B}$, W is a cylindrical Wiener process on some Hilbert space $\mathcal{K}$, and $Q\colon \mathcal{K} \to \mathcal{B}$ is a bounded linear operator.

We do not in general expect x to take values in $\mathcal{D}(L)$ and we do not even in general expect $QW(t)$ to be a $\mathcal{B}$-valued Wiener process, so that the usual way of defining solutions to (5.1) by simply integrating both sides of the equality does not work. However, if we apply some $\ell \in \mathcal{D}(L^*)$ to both sides of (5.1), then the usual definition makes sense. This motivates the following definition:

Definition 5.1 A $\mathcal{B}$-valued process $x(t)$ is said to be a *weak solution* to (5.1) if, for every $t > 0$, $\int_0^t \|x(s)\|\,ds < \infty$ almost surely and the identity

$$\langle \ell, x(t)\rangle = \langle \ell, x_0\rangle + \int_0^t \langle L^*\ell, x(s)\rangle\,ds + \int_0^t \langle Q^*\ell, dW(s)\rangle\;, \tag{5.2}$$

holds almost surely for every $\ell \in \mathcal{D}(L^*)$.

Remark 5.2 The stochastic integral in (5.2) can be interpreted in the sense of Section 3.4 since the map $Q^*\ell\colon \mathcal{K} \to \mathbf{R}$ is Hilbert-Schmidt for every $\ell \in \mathcal{B}^*$.

Remark 5.3 The term 'weak' refers to the PDE notion of a weak solution and *not* to the probabilistic notion of a weak solution to a stochastic differential equation.

Remark 5.4 Although separability of $\mathcal{B}$ was not required in the previous section on semigroup theory, it is again needed in this section, since many of the results from the section on Gaussian measure theory would not hold otherwise.

On the other hand, suppose that $f\colon \mathbf{R}_+ \to \mathcal{D}(L)$ is a continuous function and consider the function $x\colon \mathbf{R}_+ \to \mathcal{D}(L)$ given by $x(t) = S(t)x_0 + \int_0^t S(t-s)f(s)\,ds$, where S is the $\mathcal{C}_0$-semigroup generated by L. If $x_0 \in \mathcal{D}(L)$ as well, then this function is differentiable and it is easy to check, using Proposition 4.6, that it satisfies the differential equation $\partial_t x = Lx + f$. Formally replacing $f(s)\,ds$ by $Q\,dW$, this suggests the following alternative definition of a solution to (5.1):

Definition 5.5 A $\mathcal{B}$-valued process $x(t)$ is said to be a *mild solution* to (5.1) if the identity

$$x(t) = S(t)x_0 + \int_0^t S(t-s)Q\,dW(s)\;, \tag{5.3}$$

holds almost surely for every $t > 0$. The right hand side of (5.3) is also sometimes called a *stochastic convolution*.

Remark 5.6 By the results from Section 3.4, the right hand side of (5.3) makes sense in any Hilbert space $\mathcal{H}$ containing $\mathcal{B}$ and such that $\int_0^t \operatorname{tr} \iota S(t-s)QQ^*S(t-s)^*\iota^*\,ds < \infty$, where $\iota\colon \mathcal{B} \to \mathcal{H}$ is the inclusion map. The statement should then be interpreted as saying that the right hand side belongs to $\mathcal{B} \subset \mathcal{H}$ almost surely. In the case where $\mathcal{B}$ is itself a Hilbert space, (5.3) makes sense if and only if $\int_0^t \operatorname{tr} S(t-s)QQ^*S(t-s)^*\,ds < \infty$.

It turns out that these two notions of solutions are actually equivalent. In order to prepare the proof of this result, we first show that:

Lemma 5.7 *If $S(t)$ is a $\mathcal{C}_0$-semigroup on $\mathcal{B}$, then $S^*(t)$ is a $\mathcal{C}_0$-semigroup on the closure of $\mathcal{D}(L^*)$ in $\mathcal{B}^*$ and its generator is given by L^*.*

Proof. Since the resolvent of the adjoint is the adjoint of the resolvent, it is immediate that L^* satisfies the conditions of the Hille-Yosida theorem on the closure of $\mathcal{D}(L^*)$ in $\mathcal{B}^*$. It is therefore the generator of a strongly continuous semigroup $T(t)$. The fact that $T(t) = S(t)^*$ follows from the corresponding relation between the semigroups generated by the Yosida approximations of L and L^*. □

Proposition 5.8 *If the mild solution is almost surely integrable, then it is also a weak solution. Conversely, every weak solution is a mild solution.*

Proof. Note first that, by considering the process $x(t) - S(t)x_0$ and using Proposition 4.6, we can assume without loss of generality that $x_0 = 0$.

We now assume that the process $x(t)$ defined by (5.3) takes values in $\mathcal{B}$ almost surely and we show that this implies that it satisfies (5.2). Fixing an arbitrary $\ell \in \mathcal{D}(L^*)$, applying $L^*\ell$ to both sides of (5.3), and integrating the result between 0 and t, we obtain:

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

Using Proposition 4.6 and the fact that, by Lemma 5.7, S^* is a strongly continuous semigroup on $\bar{\mathcal{B}}^*$, the closure of $\mathcal{D}(L^*)$ in $\mathcal{B}^*$, we obtain

$$\begin{aligned}\int_0^t \langle L^*\ell, x(s)\rangle\, ds &= \int_0^t \langle Q^*S^*(t-r)\ell, dW(r)\rangle - \int_0^t \langle Q^*\ell, dW(r)\rangle \\ &= \Big\langle \ell, \int_0^t S(t-r)Q\, dW(r)\Big\rangle - \int_0^t \langle Q^*\ell, dW(r)\rangle \\ &= \langle \ell, x(t)\rangle - \int_0^t \langle Q^*\ell, dW(r)\rangle\ ,\end{aligned}$$

thus showing that x is indeed a weak solution to (5.1).

To show the converse, let now $x(t)$ be any weak solution to (5.1) (again with $x_0 = 0$). Fix an arbitrary $\ell \in \mathcal{D}(L^*)$, some final time $t > 0$, and consider the function $f(s) = S^*(t-s)\ell$. Since $\ell \in \mathcal{D}(L^*)$, it follows from Proposition 4.6 that this function belongs to $\mathcal{E} \stackrel{\text{def}}{=} \mathcal{C}([0,t], \mathcal{D}(L^*)) \cap \mathcal{C}^1([0,t], \bar{\mathcal{B}}^*)$. We are going to show that one has for such functions the almost sure identity

$$\langle f(t), x(t)\rangle = \int_0^t \langle \dot f(s) + L^*f(s), x(s)\rangle + \int_0^t \langle f(s), Q\, dW(s)\rangle\ . \tag{5.4}$$

Since in our case $\dot f(s) + L^*f(s) = 0$, this implies that the identity

$$\langle \ell, x(t)\rangle = \int_0^t \langle \ell, S(t-s)Q\, dW(s)\rangle\ ,$$

holds almost surely for all $\ell \in \mathcal{D}(L^*)$. Since, by the closed graph theorem, $\mathcal{D}(L^*)$ is large enough to separate points in $\mathcal{B}$[1] and since $\mathcal{B}$ is separable (so that countably many elements of $\mathcal{D}(L^*)$ are sufficient), this implies that x is indeed a mild solution.

It remains to show that (5.4) holds for all $f \in \mathcal{E}$. Since linear combinations of functions of the type $\varphi_\ell(s) = \ell\varphi(s)$ for $\varphi \in \mathcal{C}^1([0,t], \mathbf{R})$ and $\ell \in \mathcal{D}(L^*)$ are dense in $\mathcal{E}$ (see Exercise 5.10

[1] Assume that, for some $x, y \in \mathcal{B}$, we have $\langle \ell, x\rangle = \langle \ell, y\rangle$ for every $\ell \in \mathcal{D}(L^*)$. We can also assume without loss of generality that the range of L is $\mathcal{B}$, so that $x = Lx'$ and $y = Ly'$, thus yielding $\langle L^*\ell, x'\rangle = \langle L^*\ell, y'\rangle$. Since L is injective and has dense domain, the closed graph theorem states that the range of L^* is all of $\mathcal{B}^*$, so that $x' = y'$ and thus also $x = y$.

below) and since x is almost surely integrable, it suffices to show that (5.4) holds for $f = \varphi\ell$. Since $\langle \ell, QW(s)\rangle$ is a standard one-dimensional Brownian motion, we can apply Itô's formula to $\varphi(s)\langle \ell, x(s)\rangle$, yielding

$$\varphi(t)\langle \ell, x(t)\rangle = \int_0^t \varphi(s)\langle L^*\ell, x(s)\rangle + \int_0^t \dot\varphi(s)\langle \ell, x(s)\rangle + \int_0^t \varphi(s)\langle \ell, Q\,dW(s)\rangle\;,$$

which coincides with (5.4) as required. □

Remark 5.9 It is actually possible to show that if the right hand side of (5.3) makes sense for some t, then it makes sense for all t and the resulting process belongs almost surely to $L^p([0,T],\mathcal{B})$ for every p. Therefore, the concepts of mild and weak solution actually always coincide. This follows from the fact that the covariance of $x(t)$ increases with t, see for example [DJT95].

Exercise 5.10 Let $f \in \mathcal{C}([0,1],\mathcal{D}(L^*)) \cap \mathcal{C}^1([0,1],\bar{\mathcal{B}}^*)$ and, for $n > 0$, define f_n on the interval $s \in [k/n,(k+1)/n]$ by cubic spline interpolation:

$$\begin{aligned}f_n(s) &= f(k/n)(k+1-ns)^2(1+2ns-2k) + f((k+1)/n)(ns-k)^2(3-2ns+2k)\\ &\quad + (ns-k)(k+1-ns)^2 n(f((k+\tfrac12)/n) - f((k-\tfrac12)/n))\\ &\quad + (ns-k)^2(ns-k-1)n(f((k+\tfrac32)/n) - f((k+\tfrac12)/n))\;.\end{aligned}$$

Show that f_n is a finite linear combinations of functions of the form $\ell\varphi(s)$ with $\varphi \in \mathcal{C}^1([0,1],\mathbf{R})$ and that $f_n \to f$ in $\mathcal{C}([0,1],\mathcal{D}(L^*)) \cap \mathcal{C}^1([0,1],\bar{\mathcal{B}}^*)$.

5.1 Time and space regularity

In this subsection, we are going to study the space and time regularity of solutions to linear stochastic PDEs. For example, we are going to see how one can easily derive the fact that the solutions to the stochastic heat equation are 'almost' $\frac14$-Hölder continuous in time and 'almost' $\frac12$-Hölder continuous in space. Since we are often going to use the Hilbert-Schmidt norm of a linear operator, we introduce the notation

$$\|A\|_{\mathrm{HS}}^2 = \operatorname{tr} AA^*\;.$$

For most of this section, we are going to make use of the theory of analytic semigroups. However, we start with a very weak regularity result for the solutions to stochastic PDEs whose linear operator L generates an arbitrary $\mathcal{C}_0$-semigroup:

Theorem 5.11 *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 $Q\colon \mathcal{K} \to \mathcal{H}$ 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 $\alpha \in (0,\frac12)$ such that $\int_0^1 t^{-2\alpha}\|S(t)Q\|_{\mathrm{HS}}^2\,dt < \infty$. Then the solution x to (5.1) has almost surely continuous sample paths in $\mathcal{H}$.*

Proof. Note first that $\|S(t+s)Q\|_{\mathrm{HS}} \le \|S(s)\|\,\|S(t)Q\|_{\mathrm{HS}}$, so that the assumptions of the theorem imply that $\int_0^T t^{-2\alpha}\|S(t)Q\|_{\mathrm{HS}}^2\,dt < \infty$ for every $T > 0$. Let us fix an arbitrary terminal time T from now on. Defining the process y by

$$y(t) = \int_0^t (t-s)^{-\alpha} S(t-s)Q\,dW(s)\;,$$

we obtain the existence of a constant C such that

$$\mathbf{E}\|y(t)\|^2 = \int_0^t (t-s)^{-2\alpha}\|S(t-s)Q\|_{\mathrm{HS}}^2\,ds = \int_0^t s^{-2\alpha}\|S(s)Q\|_{\mathrm{HS}}^2\,ds \le C\;,$$

uniformly for $t \in [0,T]$. It therefore follows from Fernique's theorem that for every $p > 0$ there exist a constant C_p such that

$$\mathbf{E}\int_0^T \|y(t)\|^p\,dt < C_p\;. \tag{5.5}$$

Note now that there exists a constant c_α (actually $c_\alpha = (\sin 2\pi\alpha)/\pi$) such that the identity

$$\int_s^t (t-r)^{\alpha-1}(r-s)^{-\alpha}\,dr = \frac{1}{c_\alpha}\;,$$

holds for every $t > s$. It follows that one has the identity

$$\begin{aligned} x(t) &= S(t)x_0 + c_\alpha \int_0^t \int_s^t (t-r)^{\alpha-1}(r-s)^{-\alpha} S(t-s)\,dr\,Q\,dW(s) \\ &= S(t)x_0 + c_\alpha \int_0^t \int_0^r (t-r)^{\alpha-1}(r-s)^{-\alpha} S(t-s)Q\,dW(s)\,dr \\ &= S(t)x_0 + c_\alpha \int_0^t S(t-r) \int_0^r (r-s)^{-\alpha} S(r-s)Q\,dW(s)\,(t-r)^{\alpha-1}\,dr \\ &= S(t)x_0 + c_\alpha \int_0^t S(t-r)y(r)\,(t-r)^{\alpha-1}\,dr\;. \end{aligned} \tag{5.6}$$

The claim thus follows from (5.5) if we can show that for every $\alpha \in (0,\frac{1}{2})$ there exists $p > 0$ such that the map

$$y \mapsto F_y\;,\quad F_y(t) = \int_0^t (t-r)^{\alpha-1} S(t-r)y(r)\,dr$$

maps $L^p([0,T],\mathcal{H})$ into $\mathcal{C}([0,T],\mathcal{H})$. Since the semigroup $t \mapsto S(t)$ is uniformly bounded (in the usual operator norm) on any bounded time interval and since $t \mapsto (t-r)^{\alpha-1}$ belongs to L^q for $q \in [1, 1/(1-\alpha))$, we deduce from Hölder's inequality that there exists a constant C_T such that one does indeed have the bound $\sup_{t\in[0,T]} \|F_y(t)\|^p \le C_T \int_0^T \|y(t)\|^p\,dt$, provided that $p > \frac{1}{\alpha}$. Since continuous functions are dense in L^p, the proof is complete if we can show that F_y is continuous for every continuous function y with $y(0) = 0$.

Fixing such a y, we first show that F_y is right-continuous and then that it is left continuous. Fixing $t > 0$, we have for $h > 0$ the bound

$$\begin{aligned} \|F_y(t+h) - F_y(t)\| \le &\int_0^t \|((t+h-r)^{\alpha-1}S(h) - (t-r)^{\alpha-1})S(t-r)y(r)\|\,dr \\ &+ \int_t^{t+h} (t+h-r)^{\alpha-1}\|S(t+h-r)y(r)\|\,dr \end{aligned}$$

The second term is bounded by $\mathcal{O}(h^\delta)$ for some $\delta > 0$ by Hölder's inequality. It follows from the strong continuity of S that the integrand of the first term converges to 0 pointwise as $h \to 0$. Since on the other hand the integrand is bounded by $C(t-r)^{\alpha-1}\|y(r)\|$ for some constant C, this term also converges to 0 by the dominated convergence theorem. This shows that F_y is right continuous.

To show that F_y is also left continuous, we write

$$\begin{aligned} \|F_y(t) - F_y(t-h)\| \le &\int_0^{t-h} \|((t-r)^{\alpha-1}S(h) - (t-h-r)^{\alpha-1})S(t-h-r)y(r)\|\,dr \\ &+ \int_{t-h}^t (t-r)^{\alpha-1}\|S(t-r)y(r)\|\,dr\;. \end{aligned}$$

We bound the second term by Hölder's inequality as before. The second term can be rewritten as

$$\int_0^t \|((t+h-r)^{\alpha-1}S(h)-(t-r)^{\alpha-1})S(t-r)y(r-h)\|\,dr\;,$$

with the understanding that $y(r)=0$ for $r<0$. Since we assumed that y is continuous, we can again use the dominated convergence theorem to show that this term tends to 0 as $h\to 0$. □

Remark 5.12 The trick employed in (5.6) is sometimes called the "factorisation method" and was introduced in the context of stochastic convolutions by Da Prato, Kwapień, and Zabczyk [DPKZ87, DPZ92a].

This theorem is quite sharp in the sense that, without any further assumption on Q and L, it is not possible in general to deduce that $t\mapsto x(t)$ has more regularity than just continuity, even if we start with a very regular initial condition, say $x_0=0$. We illustrate this fact with the following exercise:

Exercise 5.13 Consider the case $\mathcal{H}=L^2(\mathbf{R})$, $\mathcal{K}=\mathbf{R}$, $L=\partial_x$ and $Q=g$ for some $g\in L^2(\mathbf{R})$ such that $g\geq 0$ and $g(x)=|x|^{-\beta}$ for some $\beta\in(0,\frac{1}{2})$ and all $|x|<1$. This satisfies the conditions of Theorem 5.11 for any $\alpha<1$.

Since L generates the translation group, the solution to

$$du(x,t)=\partial_x u(x,t)\,dt+g(x)\,dW(t)\;,\quad u(x,0)=0\;,$$

is given by

$$u(x,t)=\int_0^t g(x+t-s)\,dW(s)\;.$$

Convince yourself that for fixed t, the map $x\mapsto u(x,t)$ is in general γ-Hölder continuous for $\gamma<\frac{1}{2}-\beta$, but no better. Deduce from this that the map $t\mapsto u(\cdot,t)$ is in general also γ-Hölder continuous for $\gamma<\frac{1}{2}-\beta$ (if we consider it either as an $\mathcal{H}$-valued map or as a $\mathcal{C}_b(\mathbf{R})$-valued map), but cannot be expected to have more regularity than that. Since β can be chosen arbitrarily close to $\frac{1}{2}$, it follows that the exponent α appearing in Theorem 5.11 is in general independent of the Hölder regularity of the solution.

One of the main insights of regularity theory for parabolic PDEs (both deterministic and stochastic) is that space regularity is intimately linked to time regularity in several ways. Very often, the knowledge that a solution has a certain spatial regularity for fixed time implies that it also has a certain temporal regularity at a given spatial location.

From a slightly different point of view, if we consider time-regularity of the solution to a PDE viewed as an evolution equation in some infinite-dimensional space of functions, then the amount of regularity that one obtains depends on the functional space under consideration. As a general rule, the smaller the space (and therefore the more spatial regularity it imposes) the lower the regularity of the solution, viewed as a function with values in that space.

We start by considering the case of linear SPDEs with a self-adjoint linear part that take values in a Hilbert space $\mathcal{H}$ and we obtain conditions for the solutions to take values in one of the interpolation spaces:

Theorem 5.14 *Consider (5.1) on a Hilbert space $\mathcal{H}$, assume that L generates an analytic semigroup, and denote by $\mathcal{H}_\alpha$ the corresponding interpolation spaces. If there exists $\alpha\geq 0$ such that $Q\colon\mathcal{K}\to\mathcal{H}_\alpha$ is bounded and $\beta\in(0,\frac{1}{2}+\alpha]$ such that $\|(-L)^{-\beta}\|_{\mathrm{HS}}<\infty$ then the solution x takes values in $\mathcal{H}_\gamma$ for every $\gamma<\gamma_0=\frac{1}{2}+\alpha-\beta$.*

Proof. As usual, we can assume without loss of generality that 0 belongs to the resolvent set of L. It suffices to show that

$$I(T) \stackrel{\text{def}}{=} \int_0^T \|(-L)^\gamma S(t)Q\|_{\text{HS}}^2 \, dt < \infty \;, \qquad \forall T > 0 \;.$$

Since Q is assumed to be bounded from $\mathcal{K}$ to $\mathcal{H}_\alpha$, there exists a constant C such that

$$I(T) \le C \int_0^T \|(-L)^\gamma S(t)(-L)^{-\alpha}\|_{\text{HS}}^2 \, dt = C \int_0^T \|(-L)^{\gamma-\alpha} S(t)\|_{\text{HS}}^2 \, dt \;.$$

Since $(-L)^{-\beta}$ is Hilbert-Schmidt, we have the bound

$$\|(-L)^{\gamma-\alpha} S(t)\|_{\text{HS}} \le \|(-L)^{-\beta}\|_{\text{HS}} \|(-L)^{\beta+\gamma-\alpha} S(t)\| \le C(1 \vee t^{\alpha-\gamma-\beta}) \;.$$

For this expression to be square integrable near $t = 0$, we need $\alpha - \gamma - \beta > -\frac{1}{2}$, which is precisely the stated condition. □

Exercise 5.15 Show that if we are in the setting of Theorem 5.14 and L is selfadjoint, then the solutions to (5.1) actually belong to $\mathcal{H}_\gamma$ for $\gamma = \gamma_0$.

Exercise 5.16 Show that the solution to the stochastic heat equation on $[0, 1]$ with periodic boundary conditions has solutions in the fractional Sobolev space H^s for every $s < 1/2$. Recall that H^s is the Hilbert space with scalar product $\langle f, g\rangle_s = \sum_k \hat f_k \hat g_k (1 + k^2)^s$, where $\hat f_k$ denotes the kth Fourier coefficient of f.

Exercise 5.17 Consider the following modified stochastic heat equation on $[0, 1]^d$ with periodic boundary conditions:

$$dx = \Delta x \, dt + (1 - \Delta)^{-\gamma} \, dW \;,$$

where W is a cylindrical Wiener process on $L^2([0, 1]^d)$. For any given $s \ge 0$, how large does γ need to be for x to take values in H^s ?

Using this knowledge about the spatial regularity of solutions, we can now turn to the time-regularity. We have:

Theorem 5.18 *Consider the same setting as in Theorem 5.14 and fix $\gamma < \gamma_0$. Then, at all times $t > 0$, the process x is almost surely δ-Hölder continuous in $\mathcal{H}_\gamma$ for every $\delta < \frac{1}{2} \wedge (\gamma_0 - \gamma)$.*

Proof. It follows from Kolmogorov's continuity criteria that it suffices to check that the bound

$$\mathbf{E}\|x(t) - x(s)\|_\gamma^2 \le C|t - s|^{1\wedge 2(\tilde\gamma-\gamma)}$$

holds uniformly in $s, t \in [t_0, T]$ for every $t_0, T > 0$ and for every $\tilde\gamma < \gamma_0$. Here and below, C is an unspecified constant that changes from expression to expression. Assume that $t > s$ from now on. It follows from the semigroup property and the independence of the increments of W that

$$x(t) = S(t - s)x(s) + \int_s^t S(t - r)Q \, dW(r) \;. \tag{5.7}$$

Furthermore, $x(s)$ is independent of the increments of W over the interval $[s, t]$, so that Proposition 4.37 allows us to get the bound

$$\begin{aligned}\mathbf{E}\|x(t) - x(s)\|_\gamma^2 &= \mathbf{E}\|S(t - s)x(s) - x(s)\|_\gamma^2 + \int_0^{t-s} \|(-L)^\gamma S(r)Q\|_{\text{HS}}^2 \, dr \\ &\le C|t - s|^{2(\tilde\gamma-\gamma)\wedge 2}\mathbf{E}\|x(s)\|_{\tilde\gamma}^2 + C \int_0^{t-s} (1 \vee r^{\alpha-\gamma-\beta})^2 \, dr \;.\end{aligned}$$

Here, we obtained the bound on the second term in exactly the same way as in the proof of Theorem 5.14. The claim now follows from the fact that $\alpha - \gamma - \beta = (\gamma_0 - \gamma) - \frac{1}{2}$. □

5.2 Long-time behaviour

This section is devoted to the behaviour of the solutions to (5.1) for large times. Let's again start with an example that illustrates some of the possible behaviours.

Example 5.19 Let $x \mapsto V(x)$ be some smooth 'potential' and let $\mathcal{H} = L^2(\mathbf{R}, \exp(-V(x))\,dx)$. Let S denote the translation semigroup (to the right) on $\mathcal{H}$ and denote its generator by $-\partial_x$. Let us first discuss which conditions on V ensure that S is a strongly continuous semigroup on $\mathcal{H}$. It is clear that it is a semigroup and that $S(t)u \to u$ for u any smooth function with compact support. It therefore only remains to show that $\|S(t)\|$ is uniformly bounded for $t \in [0,1]$ say. We have

$$\|S(t)u\|^2 = \int u^2(x-t)e^{-V(x)}\,dx = \int u^2(x)e^{-V(x)}e^{V(x)-V(x+t)}\,dx\;. \tag{5.8}$$

This shows that a necessary and sufficient condition for S to be a strongly continuous semigroup on $\mathcal{H}$ is that, for every $t > 0$, there exists C_t such that $\sup_{x\in\mathbf{R}}(V(x) - V(x+t)) \le C_t$ and such that C_t remains bounded as $t \to 0$. Examples of potentials leading to a $\mathcal{C}_0$-semigroup are x, $\sqrt{1+x^2}$, $\log(1+x^2)$, etc or any increasing function. Note however that the potential $V(x) = x^2$ does *not* lead to a strongly continuous semigroup. One different way of interpreting this is to consider the unitary transformation $K\colon u \mapsto \exp(\frac{1}{2}V)u$ from the 'flat' space L^2 into $\mathcal{H}$. Under this transformation, the generator $-\partial_x$ is turned into

$$-(K^{-1}\partial_x K u)(x) = -\partial_x u(x) - \tfrac{1}{2}V'(x)u(x)\;.$$

Considering the characterisation of generators of $\mathcal{C}_0$-semigroups given by the Hille-Yosida theorem, one would expect this to be the generator of a strongly continuous semigroup if V' is bounded from below, which is indeed a sufficient condition.

Let now V be such that S is a $\mathcal{C}_0$-semigroup and consider the SPDE on $\mathcal{H}$ given by

$$du(x,t) = -\partial_x u(x,t)\,dt + f(x)\,dW(t)\;, \tag{5.9}$$

where W is a one-dimensional Wiener process and f is some function in $\mathcal{H}$. The solution to (5.9) with initial condition $u_0 = 0$ is given as before by

$$u(x,t) = \int_0^t f(x+s-t)\,dW(s)\;. \tag{5.10}$$

If we fix the time t, we can make the change of variable $s \mapsto t-s$, so that $u(x,t)$ is equal in distribution to $\int_0^t f(x-s)\,dW(s)$.

We see that if f happens to be also square integrable (we will assume that this is the case in the sequel and we will also assume that f is not identically zero), then (5.10) has a limit in distribution as $t \to \infty$ given by

$$\tilde u(x) = \int_0^\infty f(x-s)\,dW(s)\;. \tag{5.11}$$

It is however not clear *a priori* that $\tilde u$ does belong to $\mathcal{H}$. On one hand, we have the bound

$$\mathbf{E}\int_{\mathbf{R}} \tilde u(x)^2 e^{-V(x)}\,dx = \int_{\mathbf{R}}\int_0^\infty f^2(x-t)\,dt\,e^{-V(x)}\,dx \le \int_{\mathbf{R}} f^2(t)\,dt\int_{\mathbf{R}} e^{-V(x)}\,dx\;,$$

thus showing that $\tilde u$ definitely belongs to $\mathcal{H}$ if e^{-V} has finite mass. On the other hand, there are examples where $\tilde u \in \mathcal{H}$ even though e^{-V} has infinite mass. For example, if $f(x) = 0$ for $x \le 0$, then it is necessary and sufficient to have $\int_0^\infty e^{-V(x)}\,dx < \infty$. Denote by ν the law of $\tilde u$ for further reference.

Furthermore, if e^{-V} is integrable, there are many measures on $\mathcal{H}$ that are invariant under the action of the semigroup S. For example, given a function $h \in \mathcal{H}$ which is periodic with period τ (that is $S(\tau)h = h$), we can check that the push-forward of the Lebesgue measure on $[0, \tau]$ under the map $t \mapsto S(t)h$ is invariant under the action of S. This is simply a consequence of the invariance of Lebesgue measure under the shift map. Given any invariant probability measure μ_h of this type, let v be an $\mathcal{H}$-valued random variable with law μ_h that is independent of W. We can then consider the solution to (5.9) with initial condition v. Since the law of $S(t)v$ is equal to the law of v by construction, it follows that the law of the solution converges to the distribution of the random variable $\tilde{u} + v$, with the understanding that $\tilde{u}$ and v are independent.

This shows that in the case $\int e^{-V(x)}\,dx < \infty$, it is possible to construct solutions u to (5.9) such that the law of $u(\cdot\,, t)$ converges to $\mu_h \star \nu$ for any periodic function h.

Exercise 5.20 Construct an example of a potential V such that the semigroup S is *not* strongly continuous by choosing it such that $\lim_{t\to 0} \|S(t)\| = +\infty$, even though each of the operators $S(t)$ for $t > 0$ is bounded! **Hint:** Choose V of the form $V(x) = x^3 - \sum_{n>0} nW(\frac{x-c_n}{n})$, where W is an isolated 'spike' and c_n are suitably chosen constants.

This example shows that in general, the long-time behaviour of solutions to (5.1) may depend on the choice of initial condition. It also shows that depending on the behaviour of $\mathcal{H}$, L and Q, the law of the solutions may or may not converge to a limiting distribution in the space in which solutions are considered.

In order to formalise the concept of 'long-time behaviour of solutions' for (5.1), it is convenient to introduce the *Markov semigroup* associated to (5.1). Given a linear SPDE with solutions in $\mathcal{B}$, we can define a family $\mathcal{P}_t$ of bounded linear operators on $\mathscr{B}_b(\mathcal{B})$, the space of Borel measurable bounded functions from $\mathcal{B}$ to $\mathbf{R}$ by

$$(\mathcal{P}_t\varphi)(x) = \mathbf{E}\varphi\Big(S(t)x + \int_0^t S(t-s)Q\,dW(s)\Big)\;. \tag{5.12}$$

It follows from (5.7) and the independence of increments of W that $\mathcal{P}_t$ satisfies the semigroup property $\mathcal{P}_{t+s} = \mathcal{P}_t \circ \mathcal{P}_s$ for any two times $s, t \geq 0$.

Exercise 5.21 Show that $\mathcal{P}_t$ maps the space $\mathcal{C}_b(\mathcal{B})$ of continuous bounded functions from $\mathcal{B}$ to $\mathbf{R}$ into itself. (Recall that we assumed $\mathcal{B}$ to be separable.)

If we denote by $\mathcal{P}_t(x, \cdot\,)$ the law of $S(t)x + \int_0^t S(t-s)Q\,dW(s)$, then $\mathcal{P}_t$ can alternatively be represented as

$$(\mathcal{P}_t\varphi)(x) = \int_{\mathcal{B}} \varphi(y)\,\mathcal{P}_t(x, dy)\;.$$

It follows that its dual $\mathcal{P}_t^*$ acts on measures with finite total variation by

$$(\mathcal{P}_t^*\mu)(A) = \int_{\mathcal{B}} \mathcal{P}_t(x, A)\,\mu(dx)\;.$$

Since it preserves the mass of positive measures, $\mathcal{P}_t^*$ is a continuous map from the space $\mathscr{P}_1(\mathcal{B})$ of Borel probability measures on $\mathcal{B}$ (endowed with the total variation topology) into itself. It follows from (5.12) and the definition of the dual that $\mathcal{P}_t\mu$ is nothing but the law at time t of the solution to (5.1) with its initial condition u_0 distributed according to μ, independently of the increments of W over $[0, t]$. With these notations in place, we define:

Definition 5.22 A Borel probability measure μ on $\mathcal{B}$ is an *invariant measure* for (5.1) if $\mathcal{P}_t^*\mu = \mu$ for every $t > 0$, where $\mathcal{P}_t$ is the Markov semigroup associated to solutions of (5.1) via (5.12).

In the case $\mathcal{B} = \mathcal{H}$ where we consider (5.1) on a Hilbert space $\mathcal{H}$, the situations in which such an invariant measure exists are characterised in the following theorem:

Theorem 5.23 *Consider (5.1) with solutions in a Hilbert space $\mathcal{H}$ and define the operator $Q_t\colon \mathcal{H} \to \mathcal{H}$ by*

$$Q_t = \int_0^t S(t)QQ^* S^*(t)\,dt\;.$$

Then there exists an invariant measure μ for (5.1) if and only if one of the following two equivalent conditions are satisfied:

1. *There exists a positive definite trace class operator $Q_\infty\colon \mathcal{H} \to \mathcal{H}$ such that $2\mathrm{Re}\langle Q_\infty L^* x, x\rangle + \|Q^* x\|^2 = 0$ for every $x \in \mathcal{D}(L^*)$.*
2. *One has $\sup_{t>0} \operatorname{tr} Q_t < \infty$.*

Furthermore, any invariant measure 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. The proof goes as follows. We first show that μ being invariant implies that 2. holds. Then we show that 2. implies 1., and we conclude the first part by showing that 1. implies the existence of an invariant measure.

Let us start by showing that if μ is an invariant measure for (5.1), then 2. is satisfied. By choosing $\varphi(x) = e^{i\langle h,x\rangle}$ for arbitrary $h \in \mathcal{H}$, it follows from (5.12) that the Fourier transform of $\mathcal{P}_t\mu$ satisfies the equation

$$\widehat{\mathcal{P}_t\mu}(x) = \hat\mu(S^*(t)x)e^{-\frac{1}{2}\langle x,Q_t x\rangle}\;. \tag{5.13}$$

Taking logarithms and using the fact that $|\hat\mu(x)| \le 1$ for every $x \in \mathcal{H}$ and every probability measure μ, It follows that if μ is invariant, then

$$\langle x, Q_t x\rangle \le -2\log|\hat\mu(x)|\;,\quad \forall x \in \mathcal{H}\;,\quad \forall t > 0\;. \tag{5.14}$$

Choose now a sufficiently large value of $R > 0$ so that $\mu(\|x\| > R) < 1/8$ (say) and define a symmetric positive definite operator $A_R\colon \mathcal{H} \to \mathcal{H}$ by

$$\langle h, A_R h\rangle = \int_{\|x\|\le R} |\langle x, h\rangle|^2\,\mu(dx)\;.$$

Since, for any orthonormal basis, one has $\|x\|^2 = \sum_n |\langle x, e_n\rangle|^2$, it follows that A_R is trace class and that $\operatorname{tr} A_R \le R^2$. Furthermore, one has the bound

$$|1 - \hat\mu(h)| \le \int_{\mathcal{H}} |1 - e^{i\langle h,x\rangle}|\,\mu(dx) \le \sqrt{\langle h, A_R h\rangle} + \frac{1}{4}\;.$$

Combining this with (5.14), it follows that $\langle x, Q_t x\rangle \le 2\log 4$ for every $x \in \mathcal{H}$ such that $\langle x, A_R x\rangle \le 1/4$ so that, by homogeneity,

$$\langle x, Q_t x\rangle \le (8\log 4)\langle x, A_R x\rangle\;.$$

It follows that $\operatorname{tr} Q_t \le (8\log 4)R^2$, so that 2. is satisfied. To show that 2. implies 1., note that $\sup \operatorname{tr} Q_t < \infty$ implies that

$$Q_\infty = \int_0^\infty S(t)QQ^* S^*(t)\,dt\;,$$

is a well-defined positive definite trace class operator (since $t \mapsto Q_t^{1/2}$ forms a Cauchy sequence in the space of Hilbert-Schmidt operators). Furthermore, one has the identity

$$\langle x, Q_\infty x\rangle = \langle S^*(t)x, Q_\infty S^*(t)x\rangle + \int_0^t \|Q^* S^*(s)x\|^2\, ds\ .$$

for $x \in \mathcal{D}(L^*)$, both terms on the right hand side of this expression are differentiable. Taking the derivative at $t = 0$, we get

$$0 = 2\mathrm{Re}\langle Q_\infty L^* x, x\rangle + \|Q^* x\|^2\ ,$$

which is precisely the identity in 1.

Let now Q_∞ be a given operator as in 1., we want to show that the centred Gaussian measure μ_∞ with covariance Q_∞ is indeed invariant for $\mathcal{P}_t$. For $x \in \mathcal{D}(L^*)$, it follows from Proposition 4.6 that the map $F_x : t \mapsto \langle Q_\infty S^*(t)x, S^*(t)x\rangle$ is differentiable with derivative given by $\partial_t F_x(t) = 2\mathrm{Re}\langle Q_\infty L^* S^*(t)x, S^*(t)x\rangle$. It follows that

$$F_x(t) - F_x(0) = 2\int_0^t \mathrm{Re}\langle Q_\infty L^* S^*(s)x, S^*(s)x\rangle\, ds = -\int_0^t \|Q^* S^*(s)x\|^2\, ds\ ,$$

so that one has the identity

$$Q_\infty = S(t)Q_\infty S^*(t) + \int_0^t S(s)QQ^* S^*(s)\, ds = S(t)Q_\infty S^*(t) + Q_t\ .$$

Inserting this into (5.13), the claim follows. Here, we used the fact that $\mathcal{D}(L^*)$ is dense in $\mathcal{H}$, which is always the case on a Hilbert space, see [Yos95, p. 196].

Since it is obvious from (5.13) that every measure of the type $\nu \star \mu_\infty$ with ν invariant for S is also invariant for $\mathcal{P}_t$, it remains to show that the converse also holds. Let μ be invariant for $\mathcal{P}_t$ and define μ_t as the push-forward of μ under the map $S(t)$. Since $\hat\mu_t(x) = \hat\mu(S^*(t)x)$, it follows from (5.13) and the invariance of μ that there exists a function $\psi : \mathcal{H} \to \mathbf{R}$ such that $\hat\mu_t(x) \to \psi(x)$ uniformly on bounded sets, $\psi \circ S(t)^* = \psi$, and such that $\hat\mu(x) = \psi(x)\exp(-\frac{1}{2}\langle x, Q_\infty x\rangle)$. It therefore only remains to show that there exists a probability measure ν on $\mathcal{H}$ such that $\psi = \hat\nu$.

In order to show this, it suffices to show that the family of measures $\{\mu_t\}$ is tight, that is for every $\varepsilon > 0$ there exists a compact set K such that $\mu_t(K) \geq 1-\varepsilon$ for every t. Prokhorov's theorem [Bil68, p. 37] then ensures the existence of a sequence t_n increasing to ∞ and a measure ν such that $\mu_{t_n} \to \nu$ weakly. In particular, $\hat\mu_{t_n}(x) \to \hat\nu(x)$ for every $x \in \mathcal{H}$, thus concluding the proof.

To show tightness, denote by ν_t the centred Gaussian measure on $\mathcal{H}$ with covariance Q_t and note that one can find a sequence of bounded linear operators $A_n : \mathcal{H} \to \mathcal{H}$ with the following properties:

a. One has $\|A_{n+1}x\| \geq \|A_n x\|$ for every $x \in \mathcal{H}$ and every $n \geq 0$.

b. The set $B_R = \{x : \sup_n \|A_n x\| \leq R\}$ is compact for every $R > 0$.

c. One has $\sup_n \mathrm{tr}\, A_n Q_\infty A_n^* < \infty$.

(By diagonalising Q_∞, the construction of such a family of operators is similar to the construction, given a positive sequence $\{\lambda_n\}$ with $\sum_n \lambda_n < \infty$, of a positive sequence a_n with $\lim_{n\to\infty} a_n = +\infty$ and $\sum_n a_n\lambda_n < \infty$.) Let now $\varepsilon > 0$ be arbitrary. It follows from Prokhorov's theorem that there exists a compact set $\hat K \subset \mathcal{H}$ such that $\mu(\mathcal{H} \setminus \hat K) \leq \frac{\varepsilon}{2}$. Furthermore, it follows from property *c.* above and the fact that $Q_\infty \geq Q_t$ that there exists $R > 0$ such that $\nu_t(\mathcal{H} \setminus B_R) \leq \frac{\varepsilon}{2}$. Define a set $K \subset \mathcal{H}$ by

$$K = \{z - y : z \in \hat K\ , y \in B_R\}\ .$$

It is straightforward to check, using the Heine-Borel theorem, that K is precompact.

If we now take X and Y to be independent $\mathcal{H}$-valued random variables with laws μ_t and ν_t respectively, then it follows from the definition of a mild solution and the invariance of μ that $Z = X + Y$ has law μ. Since one has the obvious implication $\{Z \in \hat{K}\}\&\{Y \in B_R\} \Rightarrow \{X \in K\}$, it follows that

$$\mu_t(\mathcal{H} \setminus K) = \mathbf{P}(X \notin K) \le \mathbf{P}(Z \notin \hat{K}) + \mathbf{P}(Y \notin B_R) \le \varepsilon ,$$

thus showing that the sequence $\{\mu_t\}$ is tight as requested. □

It is clear from Theorem 5.23 that if (5.1) does have a solution in some Hilbert space $\mathcal{H}$ and if $\|S(t)\| \to 0$ as $t \to \infty$ in that same Hilbert space, then it also possesses a unique invariant measure on $\mathcal{H}$. It turns out that as far as the "uniqueness" part of this statement is concerned, it is sufficient to have $\lim_{t\to\infty} \|S(t)x\| = 0$ for every $x \in \mathcal{H}$:

Proposition 5.24 *If $\lim_{t\to\infty} \|S(t)x\| = 0$ for every $x \in \mathcal{H}$, then (5.1) 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. In view of Theorem 5.23, the first claim follows if we show that δ_0 is the only measure that is invariant under the action of the semigroup S. Let ν be an arbitrary probability measure on $\mathcal{H}$ such that $S(t)^*\nu = \nu$ for every $t > 0$ and let $\varphi\colon \mathcal{H} \to \mathbf{R}$ be a bounded continuous function. On then has indeed

$$\int_{\mathcal{H}} \varphi(x)\nu(dx) = \lim_{t\to\infty} \int_{\mathcal{H}} \varphi(S(t)x)\nu(dx) = \varphi(0) , \tag{5.15}$$

where we first used the invariance of ν and then the dominated convergence theorem.

To show that $\mathcal{P}_t^*\nu \to \mu_\infty$ whenever an invariant measure exists we use the fact that in this case, by Theorem 5.23, one has $Q_t \uparrow Q_\infty$ in the trace class topology. Denoting by μ_t the centred Gaussian measure with covariance Q_t, the fact that L^2 convergence implies weak convergence then implies that there exists a measure $\hat{\mu}_\infty$ such that $\mu_t \to \hat{\mu}_\infty$ weakly. Furthermore, the same reasoning as in (5.15) shows that $S(t)^*\nu \to \delta_0$ weakly as $t \to \infty$. The claim then follows from the fact that $\mathcal{P}_t^*\nu = (S(t)^*\nu) \star \mu_t$ and from the fact that convolving two probability measures is a continuous operation in the topology of weak convergence. □

Note that the condition $\lim_{t\to\infty} \|S(t)x\| = 0$ for every x is *not* sufficient in general to guarantee the existence of an invariant measure for (5.1). This can be seen again with the aid of Example 5.19. Take an increasing function V with $\lim_{x\to\infty} V(x) = \infty$, but such that $\int_0^\infty e^{-V(x)}\,dx = \infty$. Then, since $\exp(V(x) - V(x+t)) \le 1$ and $\lim_{t\to\infty} \exp(V(x) - V(x+t)) = 0$ for every $x \in \mathbf{R}$, it follows from (5.8) and the dominated convergence theorem that $\lim_{t\to\infty} \|S(t)u\| = 0$ for every $u \in \mathcal{H}$. However, the fact that $\int_0^\infty e^{-V(x)}\,dx = \infty$ prevents the random process $\tilde{u}$ defined in (5.11) from belonging to $\mathcal{H}$, so that (5.9) has no invariant measure in this particular situation.

Exercise 5.25 Show that if (5.1) has an invariant measure μ_∞ but there exists $x \in \mathcal{H}$ such that $\limsup_{t\to\infty} \|S(t)x\| > 0$, then one cannot have $\mathcal{P}_t^*\delta_x \to \mu_\infty$ weakly. In this sense, the statement of Proposition 5.24 is sharp.

5.3 Convergence in other topologies

Proposition 5.24 shows that if (5.1) has an invariant measure μ_∞, one can in many cases expect to have $\mathcal{P}_t^*\nu \to \mu_\infty$ weakly for every initial measure ν. It is however not clear *a priori* whether such a convergence also holds in some stronger topologies on the space of probability measures. If we consider the finite-dimensional case (that is $\mathcal{H} = \mathbf{R}^n$ for some $n > 0$), the situation is clear: the condition $\lim_{t\to\infty} \|S(t)x\| = 0$ for every $x \in \mathcal{H}$ then implies that $\lim_{t\to\infty} \|S(t)\| = 0$, so that L has to be a matrix whose eigenvalues all have strictly negative real parts. One then has:

Proposition 5.26 *In the finite-dimensional case, assume that all eigenvalues of L strictly negative real parts and that Q_∞ has full rank. Then, there exists $T > 0$ such that $\mathcal{P}_t^* \delta_x$ has a smooth density $p_{t,x}$ with respect to Lebesgue measure for every $t > T$. Furthermore, μ_∞ has a smooth density p_∞ with respect to Lebesgue measure and there exists $c > 0$ such that, for every $\lambda > 0$, one has*

$$\lim_{t\to\infty} e^{ct} \sup_{y\in\mathbf{R}^n} e^{\lambda|y|}|p_\infty(y) - p_{t,x}(y)| = 0 .$$

In other words, $p_{t,x}$ converges to p_∞ exponentially fast in any weighted norm with exponentially increasing weight.

The proof of Proposition 5.26 is left as an exercise. It follows in a straightforward way from the explicit expression for the density of a Gaussian measure.

In the infinite-dimensional case, the situation is much less straightforward. The reason is that there exists no natural reference measure (the equivalent of the Lebesgue measure) with respect to which one could form densities.

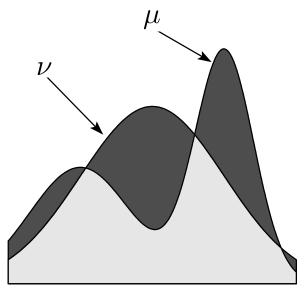
In particular, even though one always has $\|\mu_t - \mu_\infty\|_\infty \to 0$ in the finite-dimensional case (provided that μ_∞ exists and that all eigenvalues of L have strictly negative real part), one cannot expect this to be true in general. Consider for example the SPDE

$$dx = -x\,dt + Q\,dW(t) , \quad x(t) \in \mathcal{H} ,$$

where W is a cylindrical process on $\mathcal{H}$ and $Q\colon \mathcal{H} \to \mathcal{H}$ is a Hilbert-Schmidt operator. One then has

$$Q_t = \frac{1-e^{-2t}}{2} QQ^* , \qquad Q_\infty = \frac{1}{2} QQ^* .$$

Combining this with Proposition 3.30 (dilates of an infinite-dimensional Gaussian measure are mutually singular) shows that if QQ^* has infinitely many non-zero eigenvalues, then μ_t and μ_∞ are mutually singular in this case.



One question that one may ask oneself is under which conditions the convergence $\mathcal{P}_t^\nu \to \mu_\infty$ takes place in the total variation norm. The total variation distance between two probability measures is determined by their 'overlap' as depicted in the figure on the right: the total variation distance between μ and ν is given by the dark gray area, which represents the parts that do not overlap. If μ and ν have densities $\mathcal{D}_\mu$ and $\mathcal{D}_\nu$ with respect to some common reference measure π (one can always take $\pi = \frac{1}{2}(\mu+\nu)$), then one has $\|\mu - \nu\|_{\mathrm{TV}} = \int |\mathcal{D}_\mu(x) - \mathcal{D}_\nu(x)|\,\pi(dx)$.

Exercise 5.27 Show that this definition of the total variation distance does not depend on the particular choice of a reference measure.

The total variation distance between two probability measures μ and ν on a separable Banach space $\mathcal{B}$ can alternatively be characterised as

$$\|\mu - \nu\|_{\mathrm{TV}} = 2 \inf_{\pi\in\mathscr{C}(\mu,\nu)} \pi(\{x \neq y\}) , \tag{5.16}$$

where the infimum runs over the set $\mathscr{C}(\mu,\nu)$ of all probability measures π on $\mathcal{B}\times\mathcal{B}$ with marginals μ and ν. In other words, if the total variation distance between μ and ν is smaller than 2ε, then it is possible to construct $\mathcal{B}$-valued random variables X and Y with respective laws μ and ν such that $X = Y$ with probability larger than $1 - \varepsilon$.

This yields a straightforward interpretation to the total variation convergence $\mathcal{P}_t^\nu \to \mu_\infty$: for large times, a sample drawn from the invariant distribution is with high probability indistinguishable from a sample drawn from the Markov process at time t. Compare this with the notion of weak convergence which relies on the topology of the underlying space and only asserts that the two samples are close with high probability in the sense determined by the topology in question. For example, $\|\delta_x - \delta_y\|$ is always equal to 2 if $x \neq y$, whereas $\delta_x \to \delta_y$ weakly if $x \to y$.

Exercise 5.28 Show that the two definitions of the total variation distance given above are indeed equivalent by constructing a coupling that realises the infimum in (5.16). It is useful for this to consider the measure $\mu \wedge \nu$ which, in μ and ν have densities $\mathcal{D}_\mu$ and $\mathcal{D}_\nu$ with respect to some common reference measure π, is given by $(\mathcal{D}_\mu(x) \wedge \mathcal{D}_\nu(x))\pi(dx)$.

An alternative characterisation of the total variation norm is as the dual norm to the supremum norm on the space $\mathscr{B}_b(\mathcal{B})$ of bounded Borel measurable functions on $\mathcal{B}$:

$$\|\mu - \nu\|_{\mathrm{TV}} = \sup\Big\{\int \varphi(x)\mu(dx) - \int \varphi(x)\nu(dx) \;:\; \sup_{x\in\mathcal{B}} |\varphi(x)| \le 1\Big\}\;.$$

It turns out that, instead of showing directly that $\mathcal{P}_t^*\nu \to \mu_\infty$ in the total variation norm, it is somewhat easier to show that one has $\mathcal{P}_t^*\nu \to \mu_\infty$ in a type of 'weighted total variation norm', which is slightly stronger than the usual total variation norm. Given a weight function $V\colon \mathcal{B} \to \mathbf{R}_+$, we define a weighted supremum norm on measurable functions by

$$\|\varphi\|_V = \sup_{x\in\mathcal{B}} \frac{|\varphi(x)|}{1+V(x)}\;,$$

as well as the dual norm on measures by

$$\|\mu - \nu\|_{\mathrm{TV},V} = \sup\Big\{\int \varphi(x)\mu(dx) - \int \varphi(x)\nu(dx) \;:\; \|\varphi\|_V \le 1\Big\}\;. \tag{5.17}$$

Since we assumed that $V > 0$, it is obvious that one has the relation $\|\mu - \nu\|_{\mathrm{TV}} \le \|\mu - \nu\|_{\mathrm{TV},V}$, so that convergence in the weighted norm immediately implies convergence in the usual total variation norm. By considering the Jordan decomposition of $\mu - \nu = \varrho_+ - \varrho_-$, it is clear that the supremum in (5.17) is attained at functions φ such that $\varphi(x) = 1 + V(x)$ for ϱ_+-almost every x and $\varphi(x) = -1 - V(x)$ for ϱ_--almost every x. In other words, an alternative expression for the weighted total variation norm is given by

$$\|\mu - \nu\|_{\mathrm{TV},V} = \int_{\mathcal{X}} (1+V(x))\,|\mu-\nu|(dx)\;, \tag{5.18}$$

just like the total variation norm is given by $\|\mu - \nu\|_{\mathrm{TV}} = |\mu - \nu|(\mathcal{X})$.

The reason why it turns out to be easier to work in a weighted norm is the following: For a suitable choice of V, we are going to see that in a large class of examples, one can construct a weight function V and find constants $c < 1$ and $T > 0$ such that

$$\|\mathcal{P}_T^*\mu - \mathcal{P}_T^*\nu\|_{\mathrm{TV},V} \le c\|\mu - \nu\|_{\mathrm{TV},V}\;, \tag{5.19}$$

for any two probability measures μ and ν. This implies that the map $\mathcal{P}_T$ is a contraction on the space of probability measures, which must therefore have exactly one fixed point, yielding both the existence of an invariant measure μ_∞ and the exponential convergence of $\mathcal{P}_t^*\nu$ to μ_∞ for every initial probability measure ν which integrates V.

This argument is based on the following abstract result that works for arbitrary Markov semigroups on Polish (that is separable, complete, metric) spaces:

Theorem 5.29 (Harris) *Let $\mathcal{P}_t$ be a Markov semigroup on a Polish space $\mathcal{X}$ such that there exists a time $T_0 > 0$ and a function $V\colon \mathcal{X} \to \mathbf{R}_+$ such that:*

- *The exist constants $\gamma < 1$ and $K > 0$ such that $\mathcal{P}_{T_0} V(x) \le \gamma V(x) + K$ for every $x \in \mathcal{X}$.*
- *For every $K' > 0$, there exists $\delta > 0$ such that $\|\mathcal{P}^*_{T_0}\delta_x - \mathcal{P}^*_{T_0}\delta_y\|_{\mathrm{TV}} \le 2 - \delta$ for every pair x, y such that $V(x) + V(y) \le K'$.*

Then, there exists $T > 0$ such that (5.19) holds for some $c < 1$.

In a nutshell, the argument for the proof of Theorem 5.29 is the following. There are two mechanisms that allow to decrease the weighted total variation distance between two probability measures:

2. The mass of the two measures moves into regions where the weight $V(x)$ becomes smaller.
1. The two measures 'spread out' in such a way that there is an increase in the overlap between them.

The two conditions of Theorem 5.29 are tailored such as to combine these two effects in order to obtain an exponential convergence of $\mathcal{P}^*_t\mu$ to the unique invariant measure for $\mathcal{P}_t$ as $t \to \infty$.

Remark 5.30 Traditional proofs of Theorem 5.29 as given for example in [MT93] tend to make use of coupling arguments and estimates of return times of the Markov process described by $\mathcal{P}_t$ to level sets of V. Such proofs are quite involved at a technical level and are by consequent not easy to follow. Furthermore, they require more background in advanced probability theory than what is assumed for the scope of these notes. The elementary proof given here is taken from [HM08] and is based on the arguments first exposed in [HM06]. It has the disadvantage of being less intuitively appealing than proofs based on coupling arguments, but this is more than offset by the advantage of fitting into less than two pages without having to appeal to advanced mathematical concepts. It also has the advantage of being generalisable to situations where (5.19) does not hold, see [].

Before we turn to the proof of Theorem 5.29, we define for every $\beta > 0$ the distance function

$$d_\beta(x,y) = \begin{cases} 0 & \text{if } x = y \\ 2 + \beta V(x) + \beta V(y) & \text{if } x \neq y. \end{cases}$$

One can check that the positivity of V implies that this is indeed a distance function, albeit a rather strange one. We define the corresponding 'Lipschitz' seminorm on functions $\varphi\colon \mathcal{X} \to \mathbf{R}$ by

$$\|\varphi\|_{\mathrm{Lip}_\beta} = \sup_{x\neq y} \frac{|\varphi(x) - \varphi(y)|}{d_\beta(x,y)} .$$

We are going to make use of the following lemma:

Lemma 5.31 *With the above notations, one has $\|\varphi\|_{\mathrm{Lip}_\beta} = \inf_{c\in\mathbf{R}} \|\varphi + c\|_{\beta V}$.*

Proof. It is obvious that $\|\varphi\|_{\mathrm{Lip}_\beta} \le \|\varphi + c\|_{\beta V}$ for every $c \in \mathbf{R}$. On the other hand, if x_0 is any fixed point in $\mathcal{X}$, one has

$$|\varphi(x)| \le |\varphi(x_0)| + \|\varphi\|_{\mathrm{Lip}_\beta}(2 + \beta V(x) + \beta V(x_0)) , \tag{5.20}$$

for all $x \in \mathcal{X}$. Set now

$$c = -\sup_{x\in\mathcal{X}} (\varphi(x) - \|\varphi\|_{\mathrm{Lip}_\beta}(1 + \beta V(x))) .$$

It follows from (5.20) that c is finite. Furthermore, one has

$$\varphi(y) + c \le \varphi(y) - (\varphi(y) - \|\varphi\|_{\mathrm{Lip}_\beta}(1 + \beta V(y))) = \|\varphi\|_{\mathrm{Lip}_\beta}(1 + \beta V(y)) ,$$

and

$$\begin{aligned}\varphi(y) + c &= \inf_{x\in\mathcal{X}} (\varphi(y) - \varphi(x) + \|\varphi\|_{\mathrm{Lip}_\beta}(1 + \beta V(x))) \\ &\ge \inf_{x\in\mathcal{X}} \|\varphi\|_{\mathrm{Lip}_\beta}(1 + \beta V(x) - d_\beta(x,y)) = -\|\varphi\|_{\mathrm{Lip}_\beta}(1 + \beta V(y)) ,\end{aligned}$$

which implies that $\|\varphi + c\|_{\beta V} \le \|\varphi\|_{\mathrm{Lip}_\beta}$. □

Proof of Theorem 5.29. During this proof, we use the notation $\mathcal{P} \stackrel{\text{def}}{=} \mathcal{P}_{T_0}$ for simplicity. We are going to show that there exists a choice of $\beta \in (0,1)$ such that there is $\alpha < 1$ satisfying the bound

$$|\mathcal{P}\varphi(x) - \mathcal{P}\varphi(y)| \le \alpha d_\beta(x,y)\|\varphi\|_{\mathrm{Lip}_\beta} , \tag{5.21}$$

uniformly over all measurable functions $\varphi\colon \mathcal{X} \to \mathbf{R}$ and all pairs $x, y \in \mathcal{X}$. Note that this is equivalent to the bound $\|\mathcal{P}\varphi\|_{\mathrm{Lip}_\beta} \le \alpha\|\varphi\|_{\mathrm{Lip}_\beta}$. Combining this with Lemma 5.31 and (5.18), we obtain that, for $T = nT_0$, one has the bound

$$\begin{aligned}\|\mathcal{P}_T^*\mu - \mathcal{P}_T^*\nu\|_{\mathrm{TV},V} &= \inf_{\|\varphi\|_V\le 1} \int_{\mathcal{X}} (\mathcal{P}_T\varphi)(x)\,(\mu-\nu)(dx) \\ &= \inf_{\|\varphi\|_V\le 1} \inf_{c\in\mathbf{R}} \int_{\mathcal{X}} ((\mathcal{P}_T\varphi)(x) + c)\,(\mu-\nu)(dx) \\ &\le \inf_{\|\varphi\|_V\le 1} \inf_{c\in\mathbf{R}} \|\mathcal{P}_T\varphi + c\|_V \int_{\mathcal{X}} (1 + V(x))\,|\mu-\nu|(dx) \\ &= \inf_{\|\varphi\|_V\le 1} \beta^{-1} \inf_{c\in\mathbf{R}} \|\mathcal{P}_T\varphi + c\|_{\beta V}\|\mu-\nu\|_{\mathrm{TV},V} \\ &= \beta^{-1} \inf_{\|\varphi\|_V\le 1} \|\mathcal{P}_T\varphi\|_{\mathrm{Lip}_\beta}\|\mu-\nu\|_{\mathrm{TV},V} \\ &\le \frac{\alpha^n}{\beta} \inf_{\|\varphi\|_V\le 1} \|\varphi\|_{\mathrm{Lip}_\beta}\|\mu-\nu\|_{\mathrm{TV},V} \le \frac{\alpha^n}{\beta^2}\|\mu-\nu\|_{\mathrm{TV},V} .\end{aligned}$$

Since $\alpha < 1$, the result (5.19) then follows at once by choosing n sufficiently large.

Let us turn now to (5.21). If $x = y$, there is nothing to prove, so we assume that $x \neq y$. Fix an arbitrary non-constant function φ and assume without loss of generality that $\|\varphi\|_{\mathrm{Lip}_\beta} = 1$. It follows from Lemma 5.31 that, by adding a constant to it if necessary, we can assume that $|\varphi(x) + c| \le (1 + \beta V(x))$.

This immediately implies the bound

$$\begin{aligned}|\mathcal{P}\varphi(x) - \mathcal{P}\varphi(y)| &\le (2 + \beta\mathcal{P}V(x) + \beta\mathcal{P}V(y)) \\ &\le 2 + 2\beta K + \beta\gamma V(x) + \beta\gamma V(y) .\end{aligned}$$

Suppose now that x and y are such that $V(x) + V(y) \ge \frac{2K+2}{1-\gamma}$. A straightforward calculation shows that in this case, for every $\beta > 0$ there exists $\alpha_1 < 1$ such that (5.21) holds. One can choose for example

$$\alpha_1 = 1 - \frac{1}{2}\frac{\beta}{1 - \gamma + \beta K + \beta} .$$

Take now a pair x, y such that $V(x) + V(y) \le \frac{2K+2}{1-\gamma}$. Note that we can write $\varphi = \varphi_1 + \varphi_2$ with $|\varphi_1(x)| \le 1$ and $|\varphi_2(x)| \le \beta V(x)$. (Set $\varphi_1(x) = (\varphi(x) \wedge 1) \vee (-1)$.) It follows from the assumptions on V that there exists some $\delta > 0$ such that

$$\begin{aligned}|\mathcal{P}\varphi(x) - \mathcal{P}\varphi(y)| &\le |\mathcal{P}\varphi_1(x) - \mathcal{P}\varphi_1(y)| + |\mathcal{P}\varphi_2(x) - \mathcal{P}\varphi_2(y)| \\ &\le \|\mathcal{P}^*\delta_x - \mathcal{P}^*\delta_y\|_{\mathrm{TV}} + \beta(\mathcal{P}V)(x) + \beta(\mathcal{P}V)(y) \\ &\le 2 - \delta + \beta(2K + \gamma V(x) + \gamma V(y)) \le 2 - \delta + 2\beta K \frac{1+\gamma}{1-\gamma} .\end{aligned}$$

If we now choose $\beta < \frac{\delta}{4K}\frac{1-\gamma}{1+\gamma}$, (5.21) holds with $\alpha = 1 - \frac{1}{2}\delta < 1$. Combining this estimate with the one obtained previously shows that one can indeed find α and β such that (5.21) holds for all x and y in $\mathcal{X}$, thus concluding the proof of Theorem 5.29. □

One could argue that this theorem does not guarantee the existence of an invariant measure since the fact that $\mathcal{P}_T^*\mu = \mu$ does not guarantee that $\mathcal{P}_t\mu = \mu$ for every t. However, one has:

Lemma 5.32 *If there exists a probability measure such that $\mathcal{P}_T^*\mu = \mu$ for some fixed time $T > 0$, then there also exists a probability measure μ_∞ such that $\mathcal{P}_t^*\mu_\infty = \mu_\infty$ for all $t > 0$.*

Proof. Define the measure μ_∞ by

$$\mu_\infty(A) = \frac{1}{T}\int_0^T \mathcal{P}_t^*\mu(A)\,dt .$$

It is then a straightforward exercise to check that it does have the requested property. □

We are now able to use this theorem to obtain the following result on the convergence of the solutions to (5.1) to an invariant measure in the total variation topology:

Theorem 5.33 *Assume that (5.1) has a solution in some Hilbert space $\mathcal{H}$ and that there exists a time T such that $\|S(T)\| < 1$ and such that $S(T)$ maps $\mathcal{H}$ into the image of $Q_T^{1/2}$. Then (5.1) admits a unique invariant measure μ_∞ and there exists $\gamma > 0$ such that*

$$\|\mathcal{P}_t^*\nu - \mu_\infty\|_{\mathrm{TV}} \le C(\nu)e^{-\gamma t} ,$$

for every probability measure ν on $\mathcal{H}$ with finite second moment.

Proof. Let $V(x) = \|x\|$ and denote by μ_t the centred Gaussian measure with covariance Q_t. We then have

$$\mathcal{P}_t V(x) \le \|S(t)x\| + \int_{\mathcal{H}} \|x\|\,\mu_t(dx) ,$$

which shows that the first assumption of Theorem 5.29 is satisfied. A simple variation of Exercise 3.24 (use the decomposition $\mathcal{H} = \tilde{\mathcal{H}} \oplus \ker K$) shows that the Cameron-Martin space of μ_T is given by $\mathrm{Im}\, Q_T^{1/2}$ endowed with the norm

$$\|h\|_T = \inf\{\|x\| \,:\, h = Q_T^{1/2}x\} .$$

Since we assumed that $S(T)$ maps $\mathcal{H}$ into the image of $Q_T^{1/2}$, it follows from the closed graph theorem that there exists a constant C such that $\|S(T)x\|_T \le C\|x\|$ for every $x \in \mathcal{H}$. It follows from the Cameron-Martin formula that the total variation distance between $\mathcal{P}_T^*\delta_x$ and $\mathcal{P}_T^*\delta_y$ is equal to the total variation distance between $\mathcal{N}(0,1)$ and $\mathcal{N}(\|S(T)x - S(T)y\|_T, 1)$, so that the second assumption of Theorem 5.29 is also satisfied.

Both the existence of μ_∞ and the exponential convergence of $\mathcal{P}_t^*\nu$ towards it is then a consequence of Banach's fixed point theorem in the dual to the space of measurable functions with $\|\varphi\|_V < \infty$. □

Remark 5.34 The proof of Theorem 5.33 shows that if its assumptions are satisfied, then the map $x \mapsto \mathcal{P}_t^*\delta_x$ is continuous in the total variation distance for $t \geq T$.

Remark 5.35 Since $\operatorname{Im} S(t)$ decreases with t and $\operatorname{Im} Q_t^{1/2}$ increases with t, it follows that if $\operatorname{Im} S(t) \subset \operatorname{Im} Q_t^{1/2}$ for some t, then this also holds for any subsequent time. This is consistent with the fact that Markov operators are contraction operators in the supremum norm, so that if $x \mapsto \mathcal{P}_t^*\delta_x$ is continuous in the total variation distance for some t, the same must be true for all subsequent times.

While Theorem 5.33 is very general, it is sometimes not straightforward at all to verify its assumptions for arbitrary linear SPDEs. In particular, it might be complicated *a priori* to determine the image of $Q_t^{1/2}$. The task of identifying this subspace can be made easier by the following result:

Proposition 5.36 *The image of $Q_t^{1/2}$ is equal to the image of the map A_t given by*

$$A_t\colon L^2([0,t],\mathcal{K}) \to \mathcal{H}\;, \qquad A_t\colon h \mapsto \int_0^t S(s)Qh(s)\,ds\;.$$

Proof. Since $Q_t = A_tA_t^*$, we can use polar decomposition [RS80, Thm VI.10] to find an isometry J_t of $(\ker A_t)^\perp \subset \mathcal{H}$ (which extends to $\mathcal{H}$ by setting $J_tx = 0$ for $x \in \ker A_t$) such that $Q_t^{1/2} = A_tJ_t$.

Alternatively, one can show that, in the situation of Theorem 3.34, the Cameron-Martin space of $\tilde\mu = A^*\mu$ is given by the image under A of the Cameron-Martin space of μ. This follows from Proposition 3.21 since, as a consequence of the definition of the push-forward of a measure, the composition with A yields an isometry between $L^2(\mathcal{B},\mu)$ and $L^2(\mathcal{B},\tilde\mu)$. □

One case where it is straightforward to check whether $S(t)$ maps $\mathcal{H}$ into the image of $Q_t^{1/2}$ is the following:

Example 5.37 Consider the case where $\mathcal{K} = \mathcal{H}$, L is selfadjoint, and there exists a function $f\colon \mathbf{R} \to \mathbf{R}_+$ such that $Q = f(L)$. (This identity should be interpreted in the sense of the functional calculus already mentioned in Theorem 4.14.)

If we assume furthermore that $f(\lambda) > 0$ for every $\lambda \in \mathbf{R}$, then the existence of an invariant measure is equivalent to the existence of a constant $c > 0$ such that $\langle x, Lx\rangle \leq -c\|x\|^2$ for every $x \in \mathcal{H}$. Using functional calculus, we see that the operator Q_T is then given by

$$Q_T = \frac{f^2(L)}{2L}(1 - e^{-2LT})\;,$$

and, for every $T > 0$, the Cameron-Martin norm for μ_T is equivalent to the norm

$$\|x\|_f = \Big\|\frac{\sqrt{L}}{f(L)}x\Big\|\;.$$

In order to obtain convergence $\mathcal{P}_t^*\nu \to \mu_\infty$ in the total variation topology, it is therefore sufficient that there exist constants $c, C > 0$ such that $f(\lambda) \geq Ce^{-c\lambda}$ for $\lambda \geq 0$.

This shows that one cannot expect convergence in the total variation topology to take place under similarly weak conditions as in Proposition 5.24. In particular, convergence in the total variation topology requires some non-degeneracy of the driving noise which was not the case for weak convergence.

Exercise 5.38 Consider again the case $\mathcal{K} = \mathcal{H}$ and L selfadjoint with $\langle x, Lx\rangle \le -c\|x\|^2$ for some $c > 0$. Assume furthermore that Q is selfadjoint and that Q and L commute, so that there exists a space $L^2(\mathcal{M}, \mu)$ isometric to $\mathcal{H}$ and such that both Q and L can be realised as multiplication operators (say f and g respectively) on that space. Show that:

- In order for there to exist solutions in $\mathcal{H}$, the set $A_Q \stackrel{\text{def}}{=} \{\lambda \in \mathcal{M} \;:\; f(\lambda) \neq 0\}$ must be 'essentially countable' in the sense that it can be written as the union of a countable set and a set of μ-measure 0.
- If there exists $T > 0$ such that $\operatorname{Im} S(T) \subset \operatorname{Im} Q_T^{1/2}$, then μ is purely atomic and there exists some possibly different time $t > 0$ such that $S(t)$ is trace class.

Exercise 5.37 suggests that there are many cases where, if $S(t)$ maps $\mathcal{H}$ to $\operatorname{Im} Q_t^{1/2}$ for some $t > 0$, then it does so for all $t > 0$. It also shows that, in the case where L and Q are selfadjoint and commute, Q must have an orthonormal basis of eigenvectors with all eigenvalues non-zero. Both statements are certainly not true in general. We see from the following example that there can be infinite-dimensional situations where $S(t)$ maps $\mathcal{H}$ to $\operatorname{Im} Q_t^{1/2}$ even though Q is of rank one!

Example 5.39 Consider the space $\mathcal{H} = \mathbf{R} \oplus L^2([0,1], \mathbf{R})$ and denote elements of $\mathcal{H}$ by (a, u) with $a \in \mathbf{R}$. Consider the semigroup S on $\mathcal{H}$ given by

$$S(T)(a,u) = (a, \tilde u)\;, \quad \tilde u(x) = \left\{ \begin{array}{rl} a & \text{for } x \le t \\ u(x-t) & \text{for } x > t. \end{array}\right.$$

It is easy to check that S is indeed a strongly continuous semigroup on $\mathcal{H}$ and we denote its generator by $(0, \partial_x)$. We drive this equation by adding noise only on the first component of $\mathcal{H}$. In other words, we set $\mathcal{K} = \mathbf{R}$ and $Q1 = (1, 0)$ so that, formally, we are considering the equation

$$da = dW(t)\;, \qquad du = \partial_x u\, dt\;.$$

Even though, at a formal level, the equations for a and u look decoupled, they are actually coupled via the domain of the generator of S. In order to check whether $S(t)$ maps $\mathcal{H}$ into $\mathcal{H}_t \stackrel{\text{def}}{=} \operatorname{Im} Q_t^{1/2}$, we make use of Proposition 5.36. This shows that $\mathcal{H}_t$ consists of elements of the form

$$\int_0^t h(s)\chi_s\, ds\;,$$

where $h \in L^2([0,t])$ and χ_s is the image of $(1, 0)$ under $S(s)$, which is given by $(1, \mathbf{1}_{[0,s\wedge 1]})$. On the other hand, the image of $S(t)$ consists of all elements (a, u) such that $u(x) = a$ for $x \le t$. Since one has $\chi_s(x) = 0$ for $x > s$, it is obvious that $\operatorname{Im} S(t) \not\subset \mathcal{H}_t$ for $t < 1$.

On the other hand, for $t > 1$, given any $a > 0$, we can find a function $h \in L^2([0,t])$ such that $h(x) = 0$ for $x \le 1$ and $\int_0^t h(x)\, dx = a$. Since, for $s \ge 1$, one has $\chi_s(x) = 1$ for every $x \in [0,1]$, it follows that one does have $\operatorname{Im} S(t) \subset \mathcal{H}_t$ for $t < 1$.

6 Semilinear SPDEs

Now that we have a good working knowledge of the behaviour of solutions to linear stochastic PDEs, we are prepared to turn to nonlinear SPDEs. In these notes, we will restrict ourselves to the study of *semilinear* SPDEs with *additive* noise.

In this context, a *semilinear* SPDE is one such that the nonlinearity can be treated as a perturbation of the linear part of the equation. The word *additive* for the noise refers to the fact that, as in (5.1), we will only consider noises described by a fixed operator $Q\colon \mathcal{K} \to \mathcal{B}$, rather than by an operator-valued function of the solution. We will therefore consider equations of the type

$$dx = Lx\,dt + F(x)\,dt + Q\,dW(t)\;, \quad x(0) = x_0 \in \mathcal{B}\;, \tag{6.1}$$

where L is the generator of a strongly continuous semigroup S on a separable Banach space $\mathcal{B}$, W is a cylindrical Wiener process on some separable Hilbert space $\mathcal{K}$, and $Q\colon \mathcal{K} \to \mathcal{B}$ is bounded. Furthermore, F is a measurable function from some linear subspace $\mathcal{D}(F) \subset \mathcal{B}$ into $\mathcal{B}$. We will say that a process $t \mapsto x(t) \in \mathcal{D}(F)$ is a *mild solution* to (6.1) if the identity

$$x(t) = S(t)x_0 + \int_0^t S(t-s)F(x(s))\,ds + \int_0^t S(t-s)Q\,dW(s)\;. \tag{6.2}$$

holds almost surely for every $t > 0$.

6.1 Local solutions

Throughout this section, we will make the standing assumption that the linearisation to (6.1) (that is the corresponding equation with $F = 0$) does have a continuous solution with values in $\mathcal{B}$. In order to simplify notations, we are going to write

$$W_L(t) \stackrel{\text{def}}{=} \int_0^t S(t-s)Q\,dW(s)\;,$$

In the nonlinear case, there are situations where solutions explode after a finite (but possibly random) time interval. In order to be able to account for such a situation, we introduce the notion of a local solution. Recall first that, given a cylindrical Wiener process W defined on some probability space $(\Omega, \mathbf{P})$, we can associate to it the natural filtration $\{\mathscr{F}_t\}_{t\ge 0}$ generated by the increments of W.

In this context, a *stopping time* is a positive random variable τ such that the event $\{\tau \le T\}$ is $\mathcal{F}_T$-measurable for every $T \ge 0$. With this definition at hand, we have:

Definition 6.1 A *local mild solution* to (6.1) is a $\mathcal{D}(F)$-valued stochastic process x together with a stopping time τ such that $\tau > 0$ almost surely and such that the identity

$$x(t) = S(t)x_0 + \int_0^t S(t-s)F(x(s))\,ds + W_L(t)\;, \tag{6.3}$$

holds almost surely for every stopping time t such that $t < \tau$ almost surely.

Remark 6.2 In some situations, it might be of advantage to allow F to be a map from $\mathcal{D}(F)$ to $\mathcal{B}'$ for some superspace $\mathcal{B}'$ such that $\mathcal{B} \subset \mathcal{B}'$ densely and such that $S(t)$ extends to a continuous linear map from $\mathcal{B}'$ to $\mathcal{B}$. The prime example of such a space $\mathcal{B}'$ is an interpolation space with negative index in the case where the semigroup S is analytic. The definition of a mild solution carries over to this situation without any change.

A local mild solution (x, τ) is called *maximal* if, for every mild solution $(\tilde{x}, \tilde{\tau})$, one has $\tilde{\tau} \leq \tau$ almost surely.

Exercise 6.3 Show that mild solutions to (6.1) coincide with mild solutions to (6.1) with L replaced by $\tilde{L} = L - c$ and F replaced by $\tilde{F} = F + c$ for any constant $c \in \mathbf{R}$.

Our first result on the existence and uniqueness of mild solutions to nonlinear SPDEs makes the strong assumption that the nonlinearity F is defined on the whole space $\mathcal{B}$ and that it is locally Lipschitz there:

Proposition 6.4 *Consider (6.1) on a Banach space $\mathcal{B}$ and assume that W_L is a continuous $\mathcal{B}$-valued process. Assume furthermore that $F\colon \mathcal{B} \to \mathcal{B}$ is such that it restriction to every bounded set is Lipschitz continuous. Then, there exists a unique maximal mild solution (x, τ) to (6.1). 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. Given any realisation $W_L \in \mathcal{C}(\mathbf{R}_+, \mathcal{B})$ of the stochastic convolution, we are going to show that there exists a time $\tau > 0$ depending only on W_L up to time τ and a unique continuous function $x\colon [0, \tau) \to \mathcal{B}$ such that (6.3) holds for every $t < \tau$. Furthermore, the construction will be such that either $\tau = \infty$, or one has $\lim_{t\uparrow\tau} \|x(t)\| = \infty$, thus showing that (x, τ) is maximal.

The proof relies on the Banach fixed point theorem. Given a terminal time $T > 0$ and a continuous function $g\colon \mathbf{R}_+ \to \mathcal{B}$, we define the map $M_{g,T}\colon \mathcal{C}([0, T], \mathcal{B}) \to \mathcal{C}([0, T], \mathcal{B})$ by

$$(M_{g,T}u)(t) = \int_0^t S(t - s)F(u(s))\, ds + g(t) \;.$$

The proof then works in almost exactly the same way as the usual proof of uniqueness of a maximal solution for ordinary differential equations with Lipschitz coefficients. Note that we can assume without loss of generality that the semigroup S is bounded, since we can always subtract a constant to L and add it back to F. Using the fact that $\|S(t)\| \leq M$ for some constant M, one has for any $T > 0$ the bound

$$\sup_{t\in[0,T]} \|M_{g,T}u(t) - M_{g,T}v(t)\| \leq MT \sup_{t\in[0,T]} \|F(u(t)) - F(v(t))\| \;. \tag{6.4}$$

Furthermore, one has

$$\sup_{t\in[0,T]} \|M_{g,T}u(t) - g(t)\| \leq MT \sup_{t\in[0,T]} \|F(u(t))\| \;. \tag{6.5}$$

Fix now an arbitrary constant $R > 0$. Since F is locally Lipschitz, it follows from (6.4) and (6.5) that there exists a maximal $T > 0$ such that $M_{g,T}$ maps the ball of radius R around g in $\mathcal{C}([0, T], \mathcal{B})$ into itself and is a contraction with contraction constant $\frac{1}{2}$ there. This shows that $M_{g,T}$ has a unique fixed point for T small enough and the choice of T was obviously performed by using knowledge of g only up to time T. Setting $g(t) = S(t)x_0 + W_L(t)$, the pair (x, T), where T is as just constructed and x is the unique fixed point of $M_{g,T}$ thus yields a local mild solution to (6.1).

In order to construct the maximal solution, we iterate this construction in the same way as in the finite-dimensional case. Uniqueness and continuity in time also follows as in the finite-dimensional case. In the case where F is globally Lipschitz continuous, denote its Lipschitz constant by K. We then see from (6.4) that $M_{g,T}$ is a contraction on the whole space for $T < 1/(KM)$, so that the choice of T can be made independently of the initial condition, thus showing that the solution exists for all times. □

While this setting is very straightforward and did not make use of any PDE theory, it nevertheless allows to construct solutions for an important class of examples, since every composition operator of the form $(N(u))(\xi) = (f \circ u)(\xi)$ is locally Lipschitz on $\mathcal{C}(K, \mathbf{R}^d)$ (for K a compact subset of $\mathbf{R}^n$, say), provided that $f\colon \mathbf{R}^d \to \mathbf{R}^d$ is locally Lipschitz continuous.

A much larger class of examples can be found if we restrict the regularity properties of F, but assume that L generates an analytic semigroup:

Proposition 6.5 *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 $\alpha \geq 0$. Assume furthermore that there exists $\gamma \geq 0$ 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 grows at most polynomially.*

Then, (6.1) has a unique maximal mild solution (x,τ) with x taking values in $\mathcal{B}_\beta$ for every $\beta < \beta_\star \stackrel{\text{def}}{=} \alpha \wedge (\gamma + 1 - \delta)$.

Proof. In order to show that (6.1) has a unique mild solution, we proceed in a way similar to the proof of Proposition 6.4 and we make use of Exercise 4.33 to bound $\|S(t-s)F(u(s))\|$ in terms of $\|F(u(s))\|_{-\delta}$. This yields instead of (6.4) the bound

$$\sup_{t\in[0,T]} \|M_{g,T}u(t) - M_{g,T}v(t)\| \leq MT^{1-\delta} \sup_{t\in[0,T]} \|F(u(t)) - F(v(t))\| , \tag{6.6}$$

and similarly for (6.5), thus showing that (6.1) has a unique $\mathcal{B}$-valued maximal mild solution (x,τ). In order to show that $x(t)$ actually belongs to $\mathcal{B}_\beta$ for $t < \tau$ and $\beta \leq \alpha \wedge \gamma$, we make use of a 'bootstrapping' argument, which is essentially an induction on β.

For notational convenience, we denote by $W_L(s,t) = \int_s^t S(t-r)Q\,dW(r)$ the stochastic convolution between times s and t. We are actually going to show the following stronger statement. Fix an arbitrary time $T > 0$. Then, for every $\beta \in [0,\beta_\star)$ there exist exponents $p_\beta \geq 1$ and $q_\beta \geq 0$ and a constant C such that the bound

$$\|x_t\|_\beta \leq Ct^{-q_\beta}(1 + \sup_{s\in[\frac{t}{2},t]} \|x_s\| + \sup_{\frac{t}{2}\leq s<r\leq t} \|W_L(s,r)\|_\beta)^{p_\beta} , \tag{6.7}$$

holds almost surely for all $t \in (0,T]$.

The bound (6.7) is obviously true for $\beta = 0$ with $p_\beta = 1$ and $q_\beta = 0$. Assume now that, for some $a = a_0 \in [1/2, 1)$ and for some $\beta = \beta_0 \in [0,\gamma]$, we have the bound

$$\|x_t\|_\beta \leq Ct^{-q_\beta}(1 + \sup_{s\in[at,t]} \|x_s\| + \sup_{at\leq s<r\leq t} \|W_L(s,r)\|_\beta)^{p_\beta} , \tag{6.8}$$

for all $t \in (0,T]$.

We will then argue that, for any arbitrary $\varepsilon \in (0, 1-\delta)$, the statement (6.8) also holds for $\beta = \beta_0 + \varepsilon$ (and therefore also for all intermediate values) and $a = a_0^2$. Since it is possible to go from $\beta = 0$ to any value of $\beta < \gamma + 1 - \delta$ in a finite number of such steps and since we are allowed to choose a as close to 1 as we wish, the claim then follows at once.

From the definition of a mild solution, we have the identity

$$x_t = S((1-a)t)x_{at} + \int_{at}^t S(t-s)F(x(s))\,ds + W_L(at,t) .$$

Since $\beta \leq \gamma$, it follows from our polynomial growth assumption on F that there exists $n > 0$ such that, for $t \in (0,T]$,

$$\|x_t\|_{\beta+\varepsilon} \leq Ct^{-\varepsilon}\|x_{at}\|_\beta + \|W_L(at,t)\|_{\beta+\varepsilon} + C\int_{at}^t (t-s)^{-(\varepsilon+\delta)}(1 + \|x_s\|_\beta)^n\,ds$$

$$\leq C(t^{-\varepsilon} + t^{1-\varepsilon-\delta}) \sup_{at\leq s\leq t} (1 + \|x_s\|_\beta^n) + \|W_L(at,t)\|_{\beta+\varepsilon}$$
$$\leq Ct^{-\varepsilon} \sup_{at\leq s\leq t} (1 + \|x_s\|_\beta^n) + \|W_L(at,t)\|_{\beta+\varepsilon} \;.$$

Here, the constant C depends on everything but t and x_0. Using the induction hypothesis, this yields the bound

$$\|x_t\|_{\beta+\varepsilon} \leq Ct^{-\varepsilon-nq_\beta}(1 + \sup_{s\in[a^2t,t]} \|x_s\| + \sup_{a^2\leq s<r\leq t} \|W_L(s,r)\|_\beta)^{np_\beta} + \|W_L(at,t)\|_{\beta+\varepsilon} \;,$$

thus showing that (6.8) does indeed hold for $\beta = \beta_0 + \varepsilon$ and $a = a_0^2$ with $p_{\beta+\varepsilon} = np_\beta$ and $q_{\beta+\varepsilon} = \varepsilon + nq_\beta$. This concludes the proof of Proposition 6.5. □

Remark 6.6 We slightly cheated in this proof since it appears to rely on the fact that the quantity $\sup_{at\leq s<r\leq t} \|W_L(s,r)\|_\alpha$ is almost surely finite, which does not immediately follow from the fact that $\sup_{0\leq s\leq t} \|W_L(s)\|_\alpha$ is almost surely finite. However, a closer inspection of the proof shows that, in order to obtain (6.7) for some fixed t, one actually only needs to have $\sup_n \sup_{t_n\leq s\leq t} \|W_L(t_n,s)\|_\alpha < \infty$ for a finite number of values t_n, which is always the case.

6.2 Interpolation inequalities and Sobolev embeddings

The kind of bootstrapping arguments used in the proof of Proposition 6.5 above are extremely useful to obtain regularity properties of the solutions to semilinear parabolic stochastic PDEs. However, they rely on obtaining bounds on the regularity of F from one interpolation space into another. In many important situations, the interpolation spaces turn out to be given by fractional Sobolev spaces. For the purpose of these notes, we are going to restrict ourselves to the analytically easier situation where the space variable of the stochastic PDE under consideration takes values in the d-dimensional torus $\mathbf{T}^d$. For the corresponding embeddings on more general manifolds or unbounded domains, see for example the comprehensive monographs [Tri83, Tri92, Tri06].

Recall that, given a distribution $u \in L^2(\mathbf{T}^d)$, we can decompose it as a Fourier series:

$$u(x) = \sum_{k\in\mathbf{Z}^d} u_k e^{i\langle k,x\rangle} \;,$$

where the identity holds for (Lebesgue) almost every $x \in \mathbf{T}^d$. Furthermore, the L^2 norm of u is given by Parseval's identity $\|u\|^2 = \sum |u_k|^2$. We have

Definition 6.7 The fractional Sobolev space $H^s(\mathbf{T}^d)$ for $s \geq 0$ is given by the subspace of functions $u \in L^2(\mathbf{T}^d)$ such that

$$\|u\|_{H^s}^2 \stackrel{\text{def}}{=} \sum_{k\in\mathbf{Z}^d} (1 + |k|^2)^s |u_k|^2 < \infty \;. \tag{6.9}$$

Note that this is a separable Hilbert space and that $H^0 = L^2$. For $s < 0$, we define H^s as the closure of L^2 under the norm (6.9).

Remark 6.8 By the spectral decomposition theorem, H^s for $s > 0$ is the domain of $(1 - \Delta)^{s/2}$ and we have $\|u\|_{H^s} = \|(1 - \Delta)^{s/2} u\|_{L^2}$.

A very important situation is the case where L is a differential operator with constant coefficients (formally $L = P(\partial_x)$ for some polynomial $P\colon \mathbf{R}^d \to \mathbf{R}$) and $\mathcal{H}$ is either an L^2 space or some Sobolev space. In this case, one has

Lemma 6.9 *Assume that $P\colon \mathbf{R}^d \to \mathbf{R}$ is a polynomial of degree $2m$ such that there exist positive constants c, C such that the bound*

$$(-1)^{m+1} c|k|^{2m} \leq P(k) \leq (-1)^{m+1} C|k|^{2m} ,$$

holds for all k outside of some compact set. Then, the operator $P(\partial_x)$ generates an analytic semigroup on $\mathcal{H} = H^s$ for every $s \in \mathbf{R}$ and the corresponding interpolation spaces are given by $\mathcal{H}_\alpha = H^{s+2m\alpha}$.

Proof. By inspection, noting that $P(\partial_x)$ is conjugate to the multiplication operator by $P(ik)$ via the Fourier decomposition. □

Note first that for any two positive real numbers a and b and any pair of positive conjugate exponents p and q, one has Young's inequality

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} , \qquad \frac{1}{p} + \frac{1}{q} = 1 . \tag{6.10}$$

As a corollary of this elementary bound, we obtain Hölder's inequality, which can be viewed as a generalisation of the Cauchy-Schwartz inequality:

Proposition 6.10 (Hölder's inequality) *Let $(\mathcal{M}, \mu)$ be a measure space and let p and q be a pair of positive conjugate exponents. Then, for any pair of measurable functions $u, v\colon \mathcal{M} \to \mathbf{R}$, one has*

$$\int_{\mathcal{M}} |u(x)v(x)|\, \mu(dx) \leq \|u\|_p \, \|v\|_q ,$$

for any pair (p, q) of conjugate exponents.

Proof. It follows from (6.10) that, for every $\varepsilon > 0$, one has the bound

$$\int_{\mathcal{M}} |u(x)v(x)|\, \mu(dx) \leq \frac{\varepsilon^p \|u\|_p^p}{p} + \frac{\|v\|_q^q}{q\varepsilon^q} ,$$

Setting $\varepsilon = \|v\|_q^{\frac{1}{p}} \|u\|_p^{\frac{1}{p}-1}$ concludes the proof. □

One interesting consequence of Hölder's inequality is the following interpolation inequality for powers of selfadjoint operators:

Proposition 6.11 *Let A be a positive definite selfadjoint operator on a separable Hilbert space $\mathcal{H}$ and let $\alpha \in [0, 1]$. Then, the bound $\|A^\alpha u\| \leq \|Au\|^\alpha \|u\|^{1-\alpha}$ holds for every $u \in \mathcal{D}(A^\alpha) \subset \mathcal{H}$.*

Proof. The extreme cases $\alpha \in \{0, 1\}$ are obvious, so we assume $\alpha \in (0, 1)$. By the spectral theorem, we can assume that $\mathcal{H} = L^2(\mathcal{M}, \mu)$ and that A is the multiplication operator by some positive function f. Applying Hölder's inequality with $p = 1/\alpha$ and $q = 1/(1 - \alpha)$, one then has

$$\begin{aligned} \|A^\alpha u\|^2 &= \int f^{2\alpha}(x) u^2(x)\, \mu(dx) = \int |fu|^{2\alpha}(x)\, |u|^{2-2\alpha}(x)\, \mu(dx) \\ &\leq \Bigl(\int f^2(x) u^2(x)\, \mu(dx)\Bigr)^\alpha \Bigl(\int u^2(x)\, \mu(dx)\Bigr)^{1-\alpha} , \end{aligned}$$

which is exactly the bound we wanted to show. □

An immediate corollary is:

Corollary 6.12 *For any $t > s$ and any $r \in [s,t]$, the bound*

$$\|u\|_{H^r}^{t-s} \le \|u\|_{H^t}^{r-s} \|u\|_{H^s}^{t-r} \tag{6.11}$$

is valid for every $u \in H^t$.

Proof. Apply Proposition 6.11 with $\mathcal{H} = H^s$, $A = (1-\Delta)^{\frac{t-s}{2}}$, and $\alpha = (r-s)/(t-s)$. □

Exercise 6.13 As a consequence of Hölder's inequality, show that for any collection of n measurable functions and any exponents $p_i > 1$ such that $\sum_{i=1}^n p_i^{-1} = 1$, one has the bound

$$\int_{\mathcal{M}} |u_1(x)\cdots u_n(x)|\,\mu(dx) \le \|u_1\|_{p_1}\cdots\|u_n\|_{p_n}\;.$$

Following our earlier discussion regarding fractional Sobolev spaces, it would be convenient to be able to bound the L^p norm of a function in terms of one of the fractional Sobolev norms. It turns out that bounding the L^∞ norm is rather straightforward:

Lemma 6.14 *For every $s > \frac{d}{2}$, the space $H^s(\mathbf{T}^d)$ is contained in the space of continuous functions and there exists a constant C such that $\|u\|_{L^\infty} \le C\|u\|_{H^s}$.*

Proof. It follows from Cauchy-Schwarz that

$$\sum_{k\in\mathbf{Z}^d} |u_k| \le \Big(\sum_{k\in\mathbf{Z}^d} (1+|k|^2)^s |u_k|^2\Big)^{1/2}\Big(\sum_{k\in\mathbf{Z}^d}(1+|k|^2)^{-s}\Big)^{1/2}\;.$$

Since the sum in the second factor converges if and only if $s > \frac{d}{2}$, the claim follows. □

Exercise 6.15 In dimension $d = 2$, find an example of an unbounded function u such that $\|u\|_{H^1} < \infty$.

Exercise 6.16 Show that for $s > \frac{d}{2}$, H^s is contained in the space $\mathcal{C}^\alpha(\mathbf{T}^d)$ for every $\alpha < s - \frac{d}{2}$.

As a consequence of Lemma 6.14, we are able to obtain a more general Sobolev embedding for all L^p spaces:

Theorem 6.17 (Sobolev embeddings) *Let $p \in [2,\infty]$. Then, for every $s > \frac{d}{2} - \frac{d}{p}$, the space $H^s(\mathbf{T}^d)$ is contained in the space $L^p(\mathbf{T}^d)$ and there exists a constant C such that $\|u\|_{L^p} \le C\|u\|_{H^s}$.*

Proof. The case $p = 2$ is obvious and the case $p = \infty$ has already been shown, so it remains to show the claim for $p \in (2,\infty)$. The idea is to divide Fourier space into 'blocks' corresponding to different length scales and to estimate separately the L^p norm of every block. More precisely, we define a sequence of functions $u^{(n)}$ by

$$u^{-1}(x) = u_0\;,\quad u^{(n)}(x) = \sum_{2^n \le |k| < 2^{n+1}} u_k e^{i\langle k,x\rangle}\;,$$

so that one has $u = \sum_{n\ge -1} u^{(n)}$. For $n \ge 0$, one has

$$\|u^{(n)}\|_{L^p}^p \le \|u^{(n)}\|_{L^2}^2 \|u^{(n)}\|_{L^\infty}^{2-p}\;. \tag{6.12}$$

Choose now $s' = \frac{d}{2} + \varepsilon$ and note that the construction of $u^{(n)}$, together with Lemma 6.14, guarantees that one has the bounds

$$\|u^{(n)}\|_{L^2} \le 2^{-ns}\|u^{(n)}\|_{H^s}\;, \quad \|u^{(n)}\|_{L^\infty} \le C\|u^{(n)}\|_{H^{s'}} \le C2^{n(s'-s)}\|u^{(n)}\|_{H^s}\;.$$

Inserting this into (6.12), we obtain

$$\|u^{(n)}\|_{L^p} \le C\|u^{(n)}\|_{H^s}2^{n\left((s'-s)\frac{2-p}{p} - \frac{2s}{p}\right)} = C\|u^{(n)}\|_{H^s}2^{n\left(\frac{d}{p}-\frac{d}{2}-s\right)} \le C\|u\|_{H^s}2^{n\left(\frac{d}{p}-\frac{d}{2}-s\right)}\;.$$

It follows that $\|u\|_{L^p} \le |u_0| + \sum_{n\ge 0}\|u^{(n)}\|_{L^p} \le C\|u\|_{H^s}$, provided that the exponent appearing in this expression is negative, which is precisely the case whenever $s > \frac{d}{2} - \frac{d}{p}$. □

Remark 6.18 For $p \neq \infty$, one actually has $H^s(\mathbf{T}^d) \subset L^p(\mathbf{T}^d)$ with $s = \frac{d}{2} - \frac{d}{p}$, but this borderline case is more difficult to obtain.

Combining the Sobolev embedding theorem and Hölder's inequality, it is eventually possible to estimate in a similar way the fractional Sobolev norm of a product of two functions:

Theorem 6.19 *Let r, s and t be positive exponents such that $s \wedge r > t$ and $s + r > t + \frac{d}{2}$. Then, if $u \in H^r$ and $v \in H^s$, the product $w = uv$ belongs to H^t.*

Proof. Define $u^{(n)}$ and $v^{(m)}$ as in the proof of the Sobolev embedding theorem and set $w^{(m,n)} = u^{(m)}v^{(n)}$. Note that one has $w_k^{(m,n)} = 0$ if $|k| > 2^{3+(m\vee n)}$. It then follows from Hölder's inequality that if $p, q \ge 2$ are such that $p^{-1} + q^{-1} = \frac{1}{2}$, one has the bound

$$\|w^{(m,n)}\|_{H^t} \le C2^{t(m\vee n)}\|w^{(m,n)}\|_{L^2} \le C2^{t(m\vee n)}\|u^{(m)}\|_{L^p}\|v^{(n)}\|_{L^q}\;.$$

Assume now that $m > n$. The conditions on r, s and t are such that there exists a pair (p,q) as above with

$$r > t + \frac{d}{2} - \frac{d}{p}\;, \qquad s > \frac{d}{2} - \frac{d}{q}\;.$$

In particular, we can find some $\varepsilon > 0$ such that

$$\|u^{(m)}\|_{L^p} \le C\|u^{(m)}\|_{H^{r-t-\varepsilon}} \le C2^{-m(t+\varepsilon)}\|u\|_{H^r}\;, \quad \|v^{(n)}\|_{L^p} \le C\|v^{(n)}\|_{H^{s-\varepsilon}} \le C2^{-m\varepsilon}\|v\|_{H^s}\;.$$

Inserting this into the previous expression, we find that

$$\|w^{(m,n)}\|_{H^t} \le C2^{-m\varepsilon - n\varepsilon}\|u\|_{H^r}\|u\|_{H^s}\;.$$

Since our assumptions are symmetric in u and v, we obtain a similar bound for the case $m \le n$, so that

$$\|w\|_{H^t} \le \sum_{m,n>0}\|w^{(m,n)}\|_{H^t} \le C\|u\|_{H^r}\|u\|_{H^s}\sum_{m,n>0}2^{-m\varepsilon-n\varepsilon} \le C\|u\|_{H^r}\|u\|_{H^s}\;,$$

as requested. □

Exercise 6.20 Show that the conclusion of Theorem 6.19 still holds if $s = t = r$ is a positive integer, provided that $s > \frac{d}{2}$.

Exercise 6.21 Similarly to Exercise 6.13, show that one can iterate this bound so that if $s_i > s \ge 0$ are exponents such that $\sum_i s_i > s + \frac{(n-1)d}{2}$, then one has the bound

$$\|u_1\cdots u_n\|_s \le C\|u_1\|_{s_1}\cdots\|u_n\|_{s_n}\;.$$

Hint: The case $s \ge \frac{d}{2}$ is simple, so it suffices to consider the case $s < \frac{d}{2}$.

The functional inequalities from the previous section allow to check that the assumptions of Propositions 6.4 and 6.5 are verified by a number of interesting equations.

6.3 Reaction-Diffusion equations

This is a class of partial differential equations that model the evolution of reactants in a gel, described by a spatial domain D. They are of the type

$$du = \Delta u\, dt + f \circ u\, dt + Q\, dW(t)\;, \tag{6.13}$$

where $u(x,t) \in \mathbf{R}^d$, $x \in D$, describes the density of the various components of the reaction at time t and location x. The nonlinearity $f\colon \mathbf{R}^d \to \mathbf{R}^d$ describes the reaction itself and the term Δu describes the diffusion of the reactants in the gel. The noise term $Q\, dW$ should be interpreted as a crude attempt to describe the fluctuations in the quantities of reactant due both to the discrete nature of the underlying particle system and the interaction with the environment[2].

Equations of the type (6.13) also appear in the theory of amplitude equations, where they appear as a kind of 'normal form' near a change of linear instability. In this particular case, one often has $d = 2$ and $f(u) = \kappa u - u|u|^2$ for some $\kappa \in \mathbf{R}$, see [BHP05]. A natural choice for the Banach space $\mathcal{B}$ in which to consider solutions to (6.13) is the space of bounded continuous functions $\mathcal{B} = \mathcal{C}(D, \mathbf{R}^d)$ since the composition operator $u \mapsto f \circ u$ (also sometimes called Nemitskii operator) then maps $\mathcal{B}$ into itself and inherits the regularity properties of f. If the domain D is sufficiently regular then the semigroup generated by the Laplacian Δ is the Markov semigroup for a Brownian motion in D. The precise description of the domain of Δ is related to the behaviour of the corresponding Brownian motion when it hits the boundary of D. In order to avoid technicalities, let us assume from now on that D consists of the torus $\mathbf{T}^n$, so that there is no boundary to consider.

Exercise 6.22 Show that in this case, Δ generates an analytic semigroup on $\mathcal{B} = \mathcal{C}(\mathbf{T}^n, \mathbf{R}^d)$ and that for $\alpha \in \mathbf{N}$, the interpolation space $\mathcal{B}_\alpha$ is given by $\mathcal{B}_\alpha = \mathcal{C}^{2\alpha}(\mathbf{T}^n, \mathbf{R}^d)$.

If Q is such that the stochastic convolution has continuous sample paths in $\mathcal{B}$ almost surely and f is locally Lipschitz continuous, we can directly apply Proposition 6.4 to obtain the existence of a unique local solution to (6.13) in $\mathcal{C}(\mathbf{T}^n, \mathbf{R}^d)$. We would like to obtain conditions on f that ensure that this local solution is also a global solution, that is the blow-up time τ is equal to infinity almost surely.

If f happens to be a globally Lipschitz continuous function, then the existence and uniqueness of global solutions follows from Proposition 6.4. Obtaining global solutions when f is not globally Lipschitz continuous is slightly more tricky. The idea is to obtain some *a priori* estimate on some functional of the solution which dominates the supremum norm and ensures that it cannot blow up in finite time.

Let us first consider the deterministic part of the equation alone. The structure we are going to exploit is the fact that the Laplacian generates a Markov semigroup. We have the following general result:

Lemma 6.23 *Let $\mathcal{P}_t$ be a Feller[3] Markov semigroup over a Polish space $\mathcal{X}$. Extend it to $\mathcal{C}_b(\mathcal{X}, \mathbf{R}^d)$ by applying it to each component independently. Let $V\colon \mathbf{R}^d \to \mathbf{R}_+$ be convex (that is $V(\alpha x + (1-\alpha)y) \le \alpha V(x) + (1-\alpha)V(y)$ for all $x, y \in \mathbf{R}^d$ and $\alpha \in [0,1]$) and define $\tilde V\colon \mathcal{C}_b(\mathcal{X}, \mathbf{R}^d) \to \mathbf{R}_+$ by $\tilde V(u) = \sup_{x\in\mathcal{X}} V(u(x))$. Then $\tilde V(\mathcal{P}_t u) \le \tilde V(u)$ for every $t \ge 0$ and every $u \in \mathcal{C}_b(\mathcal{X}, \mathbf{R}^d)$.*

[2] A more realistic description of these fluctuations would result in a covariance Q that depends on the solution u. Since we have not developed the tools necessary to treat this type of equations, we restrict ourselves to the simple case of a constant covariance operator Q.

[3] A Markov semigroup is Feller if it maps continuous functions into continuous functions.

Proof. Note first that if V is convex, then it is continuous and, for every probability measure μ on $\mathbf{R}^d$, one has the inequality

$$V\Big(\int_{\mathbf{R}^d} x\,\mu(dx)\Big) \le \int_{\mathbf{R}^d} V(x)\,\mu(dx)\;. \tag{6.14}$$

One can indeed check by induction that (6.14) holds if μ is a 'simple' measure consisting of a convex combination of finitely many Dirac measures. The general case then follows from the continuity of V and the fact that every probability measure on $\mathbf{R}^d$ can be approximated (in the topology of weak convergence) by a sequence of simple measures.

Denote now by $\mathcal{P}_t(x,\cdot\,)$ the transition probabilities for $\mathcal{P}_t$, so that $(\mathcal{P}_t u)(x) = \int_{\mathcal{X}} u(y)\,\mathcal{P}_t(x,dy)$. One then has

$$\begin{aligned}\tilde V(\mathcal{P}_t u) &= \sup_{x\in\mathcal{X}} V\Big(\int_{\mathcal{X}} u(y)\,\mathcal{P}_t(x,dy)\Big) = \sup_{x\in\mathcal{X}} V\Big(\int_{\mathbf{R}^d} v\,(u^*\mathcal{P}_t(x,\cdot\,))(dv)\Big)\\ &\le \sup_{x\in\mathcal{X}} \int_{\mathbf{R}^d} V(v)\,(u^*\mathcal{P}_t(x,\cdot\,))(dv) = \sup_{x\in\mathcal{X}} \int_{\mathcal{X}} V(u(y))\,\mathcal{P}_t(x,dy)\\ &\le \sup_{y\in\mathcal{X}} V(u(y)) = \tilde V(u)\;,\end{aligned}$$

as required. □

In particular, this result can be applied to the semigroup $S(t)$ generated by the Laplacian in (6.13), so that $\tilde V(S(t)u) \le \tilde V(u)$ for every convex V and every $u \in \mathcal{C}(\mathbf{T}^n,\mathbf{R}^d)$. This is the main ingredient allowing us to obtain a priori estimates on the solution to (6.13):

Proposition 6.24 *Consider the setting for equation (6.13) described above. Assume that Q is such that W_Δ has continuous sample paths in $\mathcal{B} = \mathcal{C}(\mathbf{T}^n,\mathbf{R}^d)$ and that there exists a convex twice differentiable function $V\colon \mathbf{R}^d \to \mathbf{R}_+$ such that $\lim_{|x|\to\infty} V(x) = \infty$ and such that, for every $R > 0$, there exists a constant C such that $\langle \nabla V(x), f(x+y)\rangle \le CV(x)$ for every $x \in \mathbf{R}^d$ and every y with $|y| \le R$. Then (6.13) has a global solution in $\mathcal{B}$.*

Proof. We denote by $u(t)$ the local mild solution to (6.13). Our aim is to obtain *a priori* bounds on $\tilde V(u(t))$ that are sufficiently good to show that one cannot have $\lim_{t\to\tau} \|u(t)\| = \infty$ for any finite (stopping) time τ.

Setting $v(t) = u(t) - W_\Delta(t)$, the definition of a mild solution shows that v satisfies the equation

$$v(t) = e^{\Delta t}v(0) + \int_0^t e^{\Delta(t-s)}(f\circ(v(s)+W_\Delta(s)))\,ds \stackrel{\text{def}}{=} e^{\Delta t}v(0) + \int_0^t e^{\Delta(t-s)}F(s)\,ds\;.$$

Since $t \mapsto v(t)$ is continuous by Proposition 6.4 and the same holds for W_Δ by assumption, the function $t \mapsto F(t)$ is continuous in $\mathcal{B}$. Therefore, one has

$$\lim_{h\to0} \frac{1}{h}\Big(\int_0^h e^{\Delta(h-s)}F(s)\,ds - he^{\Delta h}F(0)\Big) = 0\;.$$

We therefore obtain for $\tilde V(v)$ the bound

$$\limsup_{h\to0} h^{-1}(\tilde V(v(t+h)) - \tilde V(v(t))) = \limsup_{h\to0} h^{-1}(\tilde V(v(t)+hF(t)) - \tilde V(v(t)))\;.$$

Since V belongs to $\mathcal{C}^2$ by assumption, we have

$$\tilde V(v(t)+hF(t)) = \sup_{x\in\mathbf{T}^n} (V(v(x,t)) + h\langle \nabla V(v(x,t)), F(x,t)\rangle) + \mathcal{O}(h^2)\;.$$

Using the definition of F and the assumptions on V, it follows that for every $R > 0$ there exists a constant C such that, provided that $\|W_\Delta(t)\| \le R$, one has

$$\limsup_{h\to 0} h^{-1}(\tilde V(v(t+h)) - \tilde V(v(t))) \le C\tilde V(v(t)) \;.$$

A standard comparison argument for ODEs then shows that $\tilde V(v(t))$ cannot blow up as long as $\|W_\Delta(t)\|$ does not blow up, thus concluding the proof. □

Exercise 6.25 In the case $d = 1$, show that the assumptions of Proposition 6.24 are satisfied for $V(u) = u^2$ if f is any polynomial of odd degree with negative leading coefficient.

Exercise 6.26 Show that in the case $d = 3$, (6.13) has a unique global solution when we take for f the right-hand side of the Lorentz attractor:

$$f(u) = \begin{pmatrix} \sigma(u_2 - u_1) \\ u_1(\varrho - u_3) - u_2 \\ u_1u_2 - \beta u_3 \end{pmatrix} \;,$$

where ϱ, σ and β are three arbitrary positive constants.

6.4 The stochastic Navier-Stokes equations

The incompressible stochastic Navier-Stokes equations on the torus $\mathbf{R}^2$ are given by

$$du = \nu\Delta u\,dt - (u\cdot\nabla)u\,dt - \nabla p\,dt + Q\,dW(t)\;, \qquad \operatorname{div} u = 0\;, \tag{6.15}$$

where the pressure p is determined by the incompressibility condition $\operatorname{div} u = 0$ and $\nu > 0$ denotes the kinematic viscosity of the fluid. In order to put these equations into the more familiar form (6.1), we denote by Π the orthogonal projection onto the space of divergence-free vector fields. In Fourier components, Π is given by

$$(\Pi u)_k = u_k - \frac{k\langle k, u_k\rangle}{|k|^2}\;. \tag{6.16}$$

(Note here that the Fourier coefficients of a vector field are themselves vectors.) With this notation, one has

$$du = \nu\Delta u\,dt + \Pi(u\cdot\nabla)u\,dt + Q\,dW(t) \stackrel{\text{def}}{=} \Delta u\,dt + F(u)\,dt + Q\,dW(t)\;.$$

It is clear from (6.16) that Π is a contraction in any fractional Sobolev space. For $t \ge 0$, it therefore follows that

$$\|F(u)\|_{H^t} \le \|u\|_{H^s}\|\nabla u\|_{H^r} \le C\|u\|_{H^s}^2\;, \tag{6.17}$$

provided that $s > t \vee (\frac{t}{2} + \frac{1}{2} + \frac{d}{4}) = t \vee (\frac{t}{2} + 1)$. In particular, this bound holds for $s = t+1$, provided that $t > 0$.

Furthermore, in this setting, since L is just the Laplacian, if we choose $\mathcal{H} = H^s$, then the interpolation spaces $\mathcal{H}_\alpha$ are given by $\mathcal{H}_\alpha = H^{s+2\alpha}$. This allows us to apply Proposition 6.5 to show that the stochastic Navier-Stokes equations admit local solutions for any initial condition in H^s, provided that $s > 1$, and that the stochastic convolution takes values in that space. Furthermore, these solutions will immediately lie in any higher order Sobolev space, all the way up to the space in which the stochastic convolution lies.

This line of reasoning does however not yield any *a priori* bounds on the solution, so that it may blow up in finite time. The Navier-Stokes nonlinearity satisfies $\langle u, F(u)\rangle = 0$ (the scalar product is the L^2 scalar product), so one may expect bounds in L^2, but we do not know at this stage whether initial conditions in L^2 do indeed lead to local solutions. We would therefore like to obtain bounds on $F(u)$ in negative Sobolev spaces. In order to do this, we exploit the fact that H^{-s} can naturally be identified with the dual of H^s, so that

$$\|F(u)\|_{H^{-s}} = \sup\Big\{\int F(u)(x)\, v(x)\, dx\ , \quad v \in \mathcal{C}^\infty\ , \quad \|v\|_{H^s} \le 1\Big\}\ .$$

Making use of the fact that we are working with divergence-free vector fields, one has (using Einstein's convention of summation over repeated indices):

$$\int F(u)\, v\, dx = -\int v_j u_i \partial_i u_j\, dx \le \|v\|_{L^p} \|\nabla u\|_{L^2} \|u\|_{L^q}\ ,$$

provided that $p, q > 2$ and $\frac{1}{p} + \frac{1}{q} = \frac{1}{2}$. We now make use of the fact that $\|u\|_{L^q} \le C_q \|\nabla u\|_2$ for every $q \in [2, \infty)$ (but $q = \infty$ is excluded) to conclude that for every $s > 0$ there exists a constant C such that

$$\|F(u)\|_{-s} \le C \|\nabla u\|_{L^2}^2\ . \tag{6.18}$$

In order to get *a priori* bounds for the solution to the 2D stochastic Navier-Stokes equations, one can then make use of the following trick: introduce the vorticity $w = \nabla \wedge u = \partial_1 u_2 - \partial_2 u_1$. Then, provided that $\int u\, dx = 0$ (which, provided that the range of Q consists of vector fields with mean 0, is a condition that is preserved under the solutions to (6.15)), the vorticity is sufficient to describe u completely by making use of the incompressibility assumption div $u = 0$. Actually, the map $w \mapsto u$ can be described explicitly by

$$u_k = (Kw)_k = \frac{k^\perp w_k}{|k|^2}\ , \qquad (k_1, k_2)^\perp = (-k_2, k_1)\ .$$

This shows in particular that K is actually a bounded operator from H^s into H^{s+1} for every s. It follows that one can rewrite (6.15) as

$$dw = \nu \Delta w\, dt + (Kw \cdot \nabla) w\, dt + \tilde{Q}\, dW(t) \stackrel{\text{def}}{=} \Delta w\, dt + \tilde{F}(w)\, dt + \tilde{Q}\, dW(t)\ . \tag{6.19}$$

Since $\tilde{F}(w) = \nabla \wedge F(Kw)$, it follows from (6.18) that one has the bounds

$$\|\tilde{F}(w)\|_{-1-s} \le C \|w\|_{L^2}^2\ ,$$

so that $\tilde{F}$ is a locally Lipschitz continuous map from L^2 into H^s for every $s < -1$. This shows that (6.19) has unique local solutions for every initial condition in L^2 and that these solutions immediately become as regular as the corresponding stochastic convolution.

Denote now by $\tilde{W}_L$ the stochastic convolution

$$\tilde{W}_L(t) = \int_0^t e^{\Delta(t-s)} \tilde{Q}\, dW(s)\ ,$$

and define the process $v(t) = w(t) - W_L(t)$. With this notation, v is the unique solution to the random PDE

$$\partial_t v = \nu \Delta v + \tilde{F}(v + \tilde{W}_L)\ .$$

It follows from (6.17) that $\|\tilde{F}(w)\|_{H^{-s}} \le C \|w\|_{H^s}^2$, provided that $s > 1/3$. For the sake of simplicity, assume from now on that $\tilde{W}_L$ takes values in $H^{1/2}$ almost surely. Using the fact that $\langle v, \tilde{F}(v) \rangle = 0$, we then obtain for the L^2-norm of v the following *a priori* bound:

$$\partial_t \|v\|^2 = -2\nu \|\nabla v\|^2 - 2\langle \tilde{W}_L, \tilde{F}(v + \tilde{W}_L) \rangle$$

$$
\begin{aligned}
&\leq -2\nu\|\nabla v\|^2 + 2\|\tilde W_L\|_{H^{1/2}}\|v + \tilde W_L\|_{H^{1/2}}^2 \\
&\leq -2\nu\|\nabla v\|^2 + 4\|\tilde W_L\|_{H^{1/2}}(\|v\|_{H^{1/2}}^2 + \|\tilde W_L\|_{H^{1/2}}^2) \\
&\leq -2\nu\|\nabla v\|^2 + 4\|\tilde W_L\|_{H^{1/2}}(\|v\|\|\nabla v\| + \|\tilde W_L\|_{H^{1/2}}^2) \\
&\leq \frac{8}{\nu}\|\tilde W_L\|_{H^{1/2}}^2\|v\|^2 + 2\|\tilde W_L\|_{H^{1/2}}^3 \;,
\end{aligned}
$$

so that global existence of solutions then follows from Gronwall's inequality.

This calculation is only format, since it is not known in general whether the norm of v is differentiable in time. The bound that one obtains can however be made rigorous in exactly the same way as for the example of the stochastic reaction-diffusion equation.

References

[BHP05] D. BLÖMKER, M. HAIRER, and G. A. PAVLIOTIS. Modulation equations: stochastic bifurcation in large domains. *Comm. Math. Phys.* **258**, no. 2, (2005), 479–512.

[Bil68] P. BILLINGSLEY. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.

[Bog98] V. I. BOGACHEV. *Gaussian measures*, vol. 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1998.

[Dav80] E. B. DAVIES. *One-parameter semigroups*, vol. 15 of *London Mathematical Society Monographs*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1980.

[DJT95] J. DIESTEL, H. JARCHOW, and A. TONGE. *Absolutely summing operators*, vol. 43 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.

[DPKZ87] G. DA PRATO, S. KWAPIEŃ, and J. ZABCZYK. Regularity of solutions of linear stochastic equations in Hilbert spaces. *Stochastics* **23**, no. 1, (1987), 1–23.

[DPZ92a] G. DA PRATO and J. ZABCZYK. A note on stochastic convolution. *Stochastic Anal. Appl.* **10**, no. 2, (1992), 143–153.

[DPZ92b] G. DA PRATO and J. ZABCZYK. *Stochastic equations in infinite dimensions*, vol. 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1992.

[DPZ96] G. DA PRATO and J. ZABCZYK. *Ergodicity for infinite-dimensional systems*, vol. 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.

[HM06] M. HAIRER and J. C. MATTINGLY. Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations, 2006. Preprint.

[HM08] M. HAIRER and J. C. MATTINGLY. Yet another look at harris ergodic theorem for markov chains, 2008. Preprint.

[Kat82] M. P. KATS. Continuity of universally measurable linear mappings. *Sibirsk. Mat. Zh.* **23**, no. 3, (1982), 83–90, 221.

[Law04] G. F. LAWLER. An introduction to the stochastic Loewner evolution. In *Random walks and geometry*, 261–293. Walter de Gruyter GmbH & Co. KG, Berlin, 2004.

[Lun95] A. LUNARDI. *Analytic semigroups and optimal regularity in parabolic problems*. Progress in Nonlinear Differential Equations and their Applications, 16. Birkhäuser Verlag, Basel, 1995.

[MT93] S. P. MEYN and R. L. TWEEDIE. *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag London Ltd., London, 1993.

[PR07] C. PRÉVÔT and M. RÖCKNER. *A concise course on stochastic partial differential equations*, vol. 1905 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.

[PZ07] S. PESZAT and J. ZABCZYK. *Stochastic partial differential equations with Lévy noise*, vol. 113 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2007. An evolution equation approach.

[RS80] M. REED and B. SIMON. *Methods of Modern Mathematical Physics I–IV*. Academic Press, San Diego, California, 1980.

[RW94] L. C. G. ROGERS and D. WILLIAMS. *Diffusions, Markov processes, and martingales. Vol. 1*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Ltd., Chichester, second ed., 1994. Foundations.

[RY94] D. REVUZ and M. YOR. *Continuous martingales and Brownian motion*, vol. 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second ed., 1994.

[Sch66] L. SCHWARTZ. Sur le théorème du graphe fermé. *C. R. Acad. Sci. Paris Sér. A-B* **263**, (1966), A602–A605.

[She07] S. SHEFFIELD. Gaussian free fields for mathematicians. *Probab. Theory Related Fields* **139**, no. 3-4, (2007), 521–541.

[SS05] M. SANZ-SOLÉ. *Malliavin calculus*. Fundamental Sciences. EPFL Press, Lausanne, 2005. With applications to stochastic partial differential equations.

[SS06] O. SCHRAMM and S. SHEFFIELD. Contour lines of the two-dimensional discrete gaussian free field, 2006. Preprint.

[Tri83] H. TRIEBEL. *Theory of function spaces*, vol. 38 of *Mathematik und ihre Anwendungen in Physik und Technik [Mathematics and its Applications in Physics and Technology]*. Akademische Verlagsgesellschaft Geest & Portig K.-G., Leipzig, 1983.

[Tri92] H. TRIEBEL. *Theory of function spaces. II*, vol. 84 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1992.

[Tri06] H. TRIEBEL. *Theory of function spaces. III*, vol. 100 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 2006.

[Yos95] K. YOSIDA. *Functional analysis*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.