Analysis of non-autonomous Ornstein-Uhlenbeck Processes with càd-làg paths

Dissertation
zur Erlangung des Doktorgrades
an der Fakultät für Mathematik
der Universität Bielefeld

vorgelegt von

Florian Knäble

im Februar 2012

Gedruckt auf alterungsbeständigem Papier nach DIN ISO-9706.

Contents

Chapter 1

Introduction

The first part of this thesis is centered around the analysis of a particular Ornstein-Uhlenbeck stochastic partial differential equation.

$$\begin{cases} dX_t &= (A(t)X_t + f(t))dt + B(t)dL_t \\ X_s &= x \end{cases} \tag{1.1}$$

The most important features being, that the coefficients are time-dependent, that the driving noise L is Lévy and that the state space is an infinite dimensional Hilbert space H.

Stochastic differential equations like (1.1) arise when the evolution of some object (X_t) is not only determined by a deterministic differential equation, but also subject to random influences (in the form of the noise L_t) . These might come in via noisy coefficients of an otherwise deterministic equation, as an error induced by a (knowingly) inadequate approximation, or as a model for a random environment, as for the behavior of a particle under random collisions.

When X_t does not describe a finite random vector like the position of a single molecule or the values of a finite number of asset prices, but a continuous quantity like the temperature distribution in some area or a whole curve of interest rates (modeled as vectors in an infinite dimensional state space) then we speak about a stochastic partial differential equation (SPDE). See e.g. [41] for a recent introduction on SPDE.

Another important feature is the nature of the driving noise in terms of a particular stochastic process L_t. The best-known case is the one of L_t being a Brownian motion, a process with independent and time-homogeneous increments and continuous sample path. Relaxing the latter restriction to paths which allow for jump discontinuities we open up the vast field of Lévy stochastic processes with prominent examples such as the stable and compound Poisson processes. A nice introduction to Lévy processes is given in the book [4].

A linear equation with additive noise like equation (1.1) is also called an Ornstein-Uhlenbeck equation referring to the fundamental work [36] of the physicists L. Ornstein and G. Uhlenbeck. The first treatment of Ornstein-

Uhlenbeck equations with Lévy noise (but time–independent linear drift A) in an infinite-dimensional setting appeared in [11].

We will analyze this equation in the following way: First we will elaborate on the notion of solution for (1.1). For each t fixed $A(t)$ is an (unbounded) linear operator, thus for a *strong* solution we would have to assure that the solution is in the right domain of definition for Lebesgue almost all times. This we cannot do, but we can prove the existence of a solution in the *mild* and *weak* sense. For a mild solution, one needs the linear drift to generate a contraction semigroup, in order to exploit its smoothing properties. In our non-autonomous setting we thus have to require that the linear operators $A(t)$ generate a two-parameter semigroup, say $U(t,s)$ solving the associated Cauchy problem, that is

$$\frac{d}{dt}U(t,s) = A(t)U(t,s).$$

Together with some necessary smoothing properties, these conditions are wrapped up in the definition of an *exponentially stable evolution semigroup* which we present in Section 3.1. Thus we are able to define the unique mild solution of (1.1) as:

$$X(t,s,x) = U(t,s)x + \int_s^t U(t,r)f(r)dr + \int_s^t U(t,r)B(r)dL_r \qquad (1.2)$$

For an introduction to the mild approach of solving SPDE we refer to the classical reference [19] for the Gaussian case and to [40] for SPDE with Lévy noise. The latter one also includes some motivation on the choice of discontinuous noise.

The integral $\int_s^t U(t,r)B(r)dL_r$ in the mild solution is called a *stochastic convolution* and its precise definition is given in Section 2.2 after a reminder on integration with respect to Lévy processes. Then we prove in Section 3.2 that formula (1.2) provides a solution in the (analytically) weak sense as well, that is we have for $s \le t$

$$\langle X_t, y\rangle = \langle x, y\rangle + \int_s^t \langle X_r, A^*(r)y\rangle dr + \int_s^t \langle f(r), y\rangle dr + \int_s^t B^*(r)y dL_r \qquad (1.3)$$

where y is chosen arbitrarily from the common domain of the operators $A(t)$, so that, as in the mild case, the actual solution X_t is not required to be in any domain of definition. It is worth mentioning, that , unlike in the autonomous case, the adjoint semigroup $U^*(t,s)$ does not necessarily enjoy the same smoothing properties as the original semigroup $U(t,s)$. This makes the proof of equation (1.3) quite technical. Moreover, only under stronger conditions (on the smoothing properties of the adjoint semigroup), we can prove that a weak solution is also a mild one, which yields uniqueness for the weak solution.

Then in Section 4 we turn our attention to the (inhomogeneous) Markov semigroup generated by our solution. Let us denote by $X(t,s,x)$ the (unique) solution at time t which started at time s in x and then define the two-parameter family of operators

$$P(s,t)f(x) := \mathbb{E}[f(X(t,s,x))] \qquad (t \geq s).$$

Looking at equation (1.2) one can see that $P(s,t)$ can be written as

$$\mathbb{E}[f(X(t,s,x)] = \int_H f\left(U(t,s)x + \int_s^t U(t,r)f(r)dr + y\right)\mu_{s,t}(dy),$$

where $\mu_{s,t}$ is the distribution of the stochastic convolution $\int_s^t U(t,r)B(r)dL_r$. Semigroups of this type are known as generalized Mehler semigroups in the autonomous case and their representations as generalized Mehler formulae. They were first discussed in [7] and [25]. In the Gaussian case the distribution of such a stochastic convolution is well-known, it is again Gaussian and the covariance operator can be calculated in terms of the underlying Wiener process. Much of the elegance of our methods relies on the fact that we have an analogous result in the Lévy case which is proved in Lemma 4.1.1. For example, the proof that $P(s,t)$ fulfills indeed the semigroup property as stated in Lemma 4.1.6 is an easy consequence of this result. This is also the reason why we refrain from including a non-linear drift in equation (1.1). Though questions of existence and uniqueness can still be settled using contraction methods (provided the nonlinear drift is Lipschitz), there is no chance to obtain explicit (Mehler) formulae for the solution in the general semilinear case.

We want to establish and study a generator associated to this semigroup. Remember, that in the autonomous case, one usually studies the generator on an L^p space with respect to an invariant measure. If ν is an invariant measure for a (one-parameter) semigroup $(P_\tau)_{\tau\geq 0}$ it should fulfill for bounded functions f:

$$\int_H P_\tau f(x)\nu(dx) = \int_H f(x)\nu(dx) \tag{1.4}$$

For a two-parameter semigroup we cannot hope for (1.4) to hold with $P(s,t)$ in place of P_τ. In fact, the additional second (time) parameter in the semigroup calls for an additional time parameter in the measure. Thus, what we can establish (in Theorem 4.2.4), is a collection of measures $(\nu_t)_{t\in\mathbb{R}}$ satisfying for $t \geq s$ and f as above:

$$\int_H P(s,t)f(x)\nu_s(dx) = \int_H f(x)\nu_t(dx) \tag{1.5}$$

In a dual perspective it is best to view such an *evolution system of measures* as a flow of measures generated by the dual semigroup: $\nu_t = P^*(s,t)\nu_s$.

Nevertheless, we want to get back to an equation like (1.4). This can be done by taking the additional time parameter into the state space and thus getting back a one-parameter semigroup of operators, which still holds all the information of the two-parameter semigroup. This procedure is also known as space–time homogenization. In detail, if f is a bounded function on the extended state space $\mathbb{R} \times H$ then this one-parameter semigroup P_τ acts as

$$P_\tau f(t,x) = P(t,t+\tau)f(t+\tau,\cdot)(x)$$

just shifting the time variable forward and acting with the two-parameter semigroup on the space variable, where the second time parameter is supplied by the time variable. Note that if we choose f independent of t, P_τ just acts as the two-parameter semigroup $P(s,t)$ in which we are actually interested in. Thus, this is how we will obtain useful information back in the end.

The technical advantage of this formulation is, that we can associate an invariant measure ν with the semigroup P_τ. Thus we are back in the autonomous framework and can use its well-developed theory. The measure ν (now to be defined on the extended state space $\mathbb{R} \times H$) is found by averaging the evolution system of measures ν_t over time. That is for a measure ξ on the time line $\mathbb{R}$, we form the measure $\nu_t \otimes \xi$ being uniquely defined by:

$$(\nu_t \otimes \xi)(A \times [s,t]) := \int_A \nu_r(A)\xi(dr)$$

for arbitrary Borel subsets $A \in H$ and $s \leq t$. I turns out that for ν to be a truly invariant measure we would need the measure ξ to be invariant with respect to translations. Alas, we simply cannot normalize Lebesgue measure on the whole of $\mathbb{R}$, to obtain an invariant *probability* measure. If the coefficients of equation (1.1) are periodic with some finite period $T > 0$, then all information about the semigroup can be captured on the state space $[0,T] \times H$ and ξ can be taken to be normalized Lebesgue measure on $[0,T]$. This construction is explained in Chapter 6. In the non-periodic case we can choose ξ to be Lebesgue measure normalized with an exponential weight: $\xi(dt) := e^{-|t|}dt$. For the details we refer to Definition 4.3.3. Then we will not obtain equation (1.4) but its weaker analogue

$$\int_H P_\tau f(x)\nu(dx) \leq C(\tau) \int_H f(x)\nu(dx) \quad , f \geq 0 \tag{1.6}$$

for some function $C : \mathbb{R} \to \mathbb{R}_+$. A measure satisfying equation (1.6) is called a *subinvariant* measure and in many important respects these measures are as good as a fully invariant measure. In particular, the space $L^2(\nu)$ is suitable for the analysis of the generator of the semigroup P_τ as the latter one is proven to be strongly continuous on this space. In Section 4.4 we calculate the generator on a dense space of exponential test functions finding the simple form

$$Gu(t,x) = u_t(t,x) + G_t u(t,x). \tag{1.7}$$

where G_t is just the generator corresponding to the equation with coefficients frozen at t. For an explicit form see Remark 4.4.2.

This helps us in Section 5 to characterize an important function of the generator. If G is the generator of the semigroup P_τ then its *square field operator* Γ is defined as

$$\Gamma(f,g) := G(fg) - fGg - gGf$$

for functions $f, g \in D(G)$. An explicit formula for this expression, (which in the autonomous case had already been obtained in [30]), then allows for vital estimates leading to a Poincaré and a Harnack inequality for our semigroup.

Let us now elaborate on related literature, in particular on results we based our work upon. For $H = \mathbb{R}^d$ and L_t a d-dimensional Brownian motion, equation (1.1) was studied intensively in [14]. Inspired by their paper, part of our work is a generalization of their results to the case where H is infinite-dimensional and L_t a Lévy process. A number of our arguments are adapted from [14], although the Lévy setting forces us to work more heavily with Fourier transforms and the infinite dimensional setting requires extra care. In parallel to our work, evolution systems of measures have been studied in a finite dimensional Lévy setting in [47] obtaining explicit Lebesgue-densities in the particular case of α-stable Lévy noise. In [26] evolution systems of measures (which are in general not unique) have been characterized in the Gaussian case.

Concerning the generalized Mehler semigroups mentioned above, these are already well understood in the autonomous case. Invariant measures measures are established in [11] in the Gaussian case and in [25] in the non-Gaussian case. Generators are examined in [29] and the square field operator is identified in [30]. The recent paper [38] considers generalized Mehler semigroups with a noise more general than Lévy. The authors drop the assumption of stationary increments, introducing explicit time-dependence in the noise term itself. For a recent treatment of general semilinear and multiplicative non-autonomous equations with Lévy noise, see [45]. In this generality, however, nothing is known about the existence of evolution systems of measures.

Finally, let us remark that most of our above-mentioned results have already been published in [28].

In the second part of this thesis, contained in Chapter 7, we consider the Fokker-Planck equation corresponding to (1.1) with $f \equiv 0, B \equiv Id$ and L_t a Wiener process

$$\partial_t \rho_t = -2G_t^* \rho_t. \tag{1.8}$$

Here, G_t^* is the adjoint of the generator G_t from (1.7).
Following the ideas in the seminal paper [27], our aim is to interpret (1.8) in a variational sense. In [27] the authors discovered that to a (autonomous and Gaussian) Fokker-Planck equation one can associate an energy functional involving logarithmic entropy. The Fokker-Planck equation can then be solved by implementing a steepest descent method with respect to this energy functional on a space of probability measures.

In the case of an Ornstein-Uhlenbeck equation with a unique invariant measure ν this energy functional is given, for probability measures ρ, by the relative entropy with respect to the measure ν

$$\mathrm{Ent}_\nu(\rho) := \int_H \log\left(\frac{d\rho}{d\nu}(x)\right) \rho(dx).$$

See Definition 7.1.3 for details.
The steepest descent scheme to minimize this functional is then constructed iteratively as follows:

For a given value ρ_k at time $k\tau$, define the approximation ρ_{k+1} at time $(k+1)\tau$ as a minimizer of the functional

$$\rho \mapsto \frac{1}{2\tau} d^2(\rho, \rho_k) + \mathrm{Ent}_\nu(\rho). \tag{1.9}$$

Here, the distance d is the Wasserstein distance on the space of probability measures introduced in definition 7.2.1. To see that equation (1.9) is a natural way to model gradient flows in metric spaces let us compare it to the Euclidian case. In order to solve, for a smooth and convex potential $F : \mathbb{R}^n \to \mathbb{R}$, the equation

$$\frac{d}{dt} y(t) = -\nabla F(y(t)),$$

we can employ an implicit Euler scheme, setting (with $t_k := k\tau$)

$$\frac{1}{\tau} y(t_{k+1}) - y(t_k) = -\nabla F(y(t_{k+1})). \tag{1.10}$$

But the solution $y(t_{k+1})$ of (1.10) is nothing but a minimizer of the functional

$$y \mapsto \frac{1}{2\tau} |y - y(t_k)|^2 + F(y). \tag{1.11}$$

Note the similarities between equations (1.9) and (1.11). In both cases the first term measures the distance between old and new states – penalizing big deviations – while the second term enforces a reduction of the respective energy.

Following the work [27], there was thus a growing interest in gradient flow type dynamics on the Wasserstein space, culminating in the book [3]. While gradient flows on Hilbert spaces were well-understood since the 1970s (see e.g.[8]), the authors in [3] gave the first unified treatment, developing rigorously the underlying metric structures with a particular emphasis on the Wasserstein case.

As another vital contribution, the authors in [3] refined the notion of convexity for functionals of probability measures, adapting the seminal definition of displacement convexity from the paper [9] to the particular case of the Wasserstein space. In this way they were able to prove evolution variational inequalities for the solutions of the steepest descent schemes and thus gave a rigorous meaning to the notion of a gradient flow on the Wasserstein space. We give a short account on convexity in the Wasserstein space in Section 7.2.2.

Moreover, while the setting in [27] was finite dimensional, the framework in [3] was general enough to cover the case where the Wasserstein space is formed by probability measures defined over an infinite-dimensional Hilbert space.

Building on this general theory, the paper [22] extended the variational interpretation of the Fokker-Planck equation to abstract Wiener spaces and in [32] Fokker-Planck equations associated to Ornstein-Uhlenbeck equations on Banach spaces were treated.

Let us now come to our contribution and how it is structured. Let us first stress that, while we treated general Lévy noise in the first part of this work, we were not able to go beyond the Gaussian framework in the second part. We still hope that the results can be extended also to this case, but we are sure that this necessitates a substantial change in the ingredients of the minimizing scheme. It raises the question: What is the role both of relative entropy and of the 2-Wasserstein distance in this setting? The recent survey paper [1] gives a first answer, suggesting that the role of the entropy can be explained by large deviation principles. Maybe this link will present a way to cover more general noise.

Thus, our point will be to stick to the Gaussian framework but to generalize to a non-autonomous setting. Our use of generalized invariant measures in the first part of this thesis might suggest that one could hope to find a solution to the Fokker-Planck equation by considering an autonomous gradient flow on an extended state space. Alas, this is not the case. On the one hand, the entropy functional generated by such an invariant measure lacks important convexity properties, on the other hand the space-time structure simply does not seem to capture the non-autonomous dynamics. Instead we will work with a time-dependent energy functional, replacing the functional Ent_ν by Ent_{ν_t} with changing reference measures ν_t. Indeed, in Section 7.1 we start out by setting up our framework and we define ν_t to be the invariant measure of the underlying equation with coefficients frozen at t. As laid out in Section 7.2 the steepest descent scheme is then defined analogously to the autonomous case. In this section we also take a closer look at the Wasserstein distance employed in the scheme. In the infinite dimensional setting it is necessary to fit this distance to the Gaussian driving noise. This is done as in [32]. Then we recall some of the above mentioned theory from [3] on convexity of entropy functionals and of the Wasserstein distance itself, in order to show that the scheme is well-defined. In the autonomous case these convexity properties also imply convergence of the approximations given by (1.9) towards a continuous flow of probability measures. In our time-dependent case this is not straightforward. This is in part due to the fact, that we do not possess a single energy functional anymore which can act as a Lyapunov functional. Nevertheless, in Section 7.3 we obtain compactness of the approximating measure flows and in Section 7.4 we prove convergence towards a continuous curve of measures which is then shown to solve the Fokker-Planck equation. Our methods require in particular that the reference measures ν_t are equivalent for varying t, a rather demanding assumption in infinite dimensions. Section 7.5 thus presents an example of coefficients which lead to such well-behaved reference measures.

Acknowledgement

First of all, I thank my advisor Prof. Michael Röckner for his constant support, guidance and understanding. Furthermore I am grateful to my immediate colleagues from our graduate school "Internationales Graduiertenkolleg Stochastics and Real World Models" who helped me each in their own way. I have appreciated the courses given by Prof. Philippe Clement within the scope of our college, as well as his personal advice. I would like to thank Prof. Jinghai Shao for important feedback on part of this thesis. Finally I am indebted to the German Science Foundation for their financial support through said graduate college.

Chapter 2

Prerequisites on Lévy Processes and on integration with respect to Lévy martingale measures

In this chapter we give a short reminder on stochastic integration with respect to Lévy processes in order to prepare for the solution of our equation in the next chapter.

In the following let H be a real separable Hilbert space with scalar product $\langle \cdot, \cdot \rangle := \langle \cdot, \cdot \rangle_H$ and norm $\| \cdot \| := \| \cdot \|_H$. An H-valued stochastic process L adapted to a filtration $(\mathcal{F}_t)_{t \geq 0}$ is called a *Lévy process* if it has independent and stationary increments, is stochastically continuous, and we have $P(L_0 = 0) = 1$. We say that $A \in \mathcal{B}(H)$ is bounded below if $0 \notin \bar{A}$.
We denote by $N(t, A)$ the (random) number of "jumps of size A" up to time t, that is $N(t, A) := \text{card}\{0 < s \leq t | \Delta L_s \in A\}$, where $\Delta L_s := L_s - \lim_{r \nearrow s} L_r$.
If A is bounded below, then $N(t, A)$ is a Poisson process, with intensity $\mu(A)$, where $\mu(A) := \mathbb{E}[N(1, A)]$.
$N(t, A)$ (also denoted $N_t(A)$) is called the Poisson random measure associated to L, $\tilde{N}(t, A) := N(t, A) - t\mu(A)$ is called the compensated Poisson random measure.

Definition 2.0.1 *A Borel σ-finite measure μ on $H \setminus \{0\}$ with :*

$$\int_{H \setminus \{0\}} \min(1, \|x\|^2) \mu(dx) < \infty$$

is called a Lévy measure.
It is convenient to extend this measure to all of H by setting $\mu(\{0\}) = 0$.

The following theorem is an important step in understanding the general structure of Lévy processes. It shows that every Lévy process can be split into a drift part, a Gaussian part and a jump part. This decomposition opens up possibilities for stochastic integration, as the jump part can be split again: into a process with bounded jumps which can be compensated to yield a martingale and into a process with unbounded jumps which produces paths of locally bounded variation.

Theorem 2.0.2 (Lévy-Ito Decomposition) *If L is an H-valued Lévy process, there are a drift vector $b \in H$, a Wiener process W_R on H with covariance operator R of trace class, such that W_R is independent of $N_t(A)$ for any A that is bounded below and we have:*

$$L_t = bt + W_R(t) + \int_{\|x\|<1} x\tilde{N}_t(dx) + \int_{\|x\|\geq 1} xN_t(dx)$$

PROOF See e.g. [2] Theorem 4.1 . ■

The next theorem is the counterpart of Theorem 2.0.2 in terms of Fourier transforms. Interestingly, the whole jump structure of a Lévy process can be captured by a single measure, which automatically fulfills the assumptions of a Lévy measure. It is precisely the measure giving the intensity of the Poisson process $N_t(A)$ counting the jumps in a Borel set A of the state space.

Theorem 2.0.3 (Lévy-Khinchine Representation) *If L is an H-valued Lévy process with Lévy-Ito decomposition as in Theorem 2.0.2, then its characteristic function takes the form: $\mathbb{E}[e^{i\langle L_t, u\rangle}] = e^{t\lambda(u)}$ and*

$$\lambda(u) = i\langle b, u\rangle - \frac{1}{2}\langle u, Ru\rangle + \int_{H\setminus\{0\}} \left[e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle\chi_{\{\|x\|\leq 1\}}\right]\mu(dx) \quad (2.1)$$

PROOF See [31] Theorem 5.7.3 . ■

Since a finite Borel measure is characterized by its Fourier transform we will say that a measure is associated to a triple $[b, R, \mu]$ if its characteristic exponent has the form (2.1).

Remark 2.0.4 *Actually the Lévy-Khinchine representation holds not only for Lévy processes but for any infinitely divisible random variable. (See [43] for an account of infinite divisibility.) Moreover, Lévy processes and infinitely divisible measures can be brought in a one to one correspondence. In particular the converse of Theorem 2.0.3 is true: any function of the form*

$$\exp\left\{i\langle b, u\rangle - \frac{1}{2}\langle u, Ru\rangle + \int_{H\setminus\{0\}} \left[e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle\chi_{\{\|x\|\leq 1\}}\right]\mu(dx)\right\}$$

is the characteristic function of a probability measure.

2.1 Stochastic Integration with respect to Lévy martingale measures

In this subsection we follow [5], where the proofs of all results can be found.

Definition 2.1.1 *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ be a filtered probability space. Let H be a Hilbert space and let $\mathcal{R}_1 := \{A \in \mathcal{B}(U) \mid 0 \notin \bar{A}\}$, where $U := \{x \in H \mid \|x\| \leq 1\}$. Finally, for $n \in \mathbb{N}$, let $S_n := \{x \in H \mid \frac{1}{n} \leq \|x\| \leq 1\}$.*

A Lévy martingale measure on H is a set function $M : \mathbb{R}_+ \times \mathcal{R}_1 \times \Omega \to H$ satisfying:

- $M(0, A) = 0$ *almost surely for all* $A \in \mathcal{R}_1$
- $M(t, \emptyset) = 0$ *almost surely*
- *almost surely we have:* $M(t, A \cup B) = M(t, A) + M(t, B)$ *for all t and all disjoint A, $B \in \mathcal{R}_1$*
- $M(t, A)_{\{t\geq 0\}}$ *is a square-integrable martingale for each* $A \in \mathcal{R}_1$
- *if $A \cap B = \emptyset$ $M(t, A)_{\{t\geq 0\}}$ and $M(t, B)_{\{t\geq 0\}}$ are orthogonal, that is: $\langle M(t, A), M(t, B)\rangle$ is a real-valued martingale for such $A, B \in \mathcal{R}_1$*
- $\sup\{\mathbb{E}[\|M(t, A)\|^2] \mid A \in \mathcal{B}(S_n)\} < \infty$ *for every* $n \in \mathbb{N}$
- *for every sequence A_j decreasing to the empty set such that $A_j \subset \mathcal{B}(S_n)$ for all j we have:* $\lim_{j\to\infty} \mathbb{E}[\|M(t, A_j)\|^2] = 0$
- *for every $s < t$ and every $A \in \mathcal{R}_1$ we have that $M(t, A) - M(s, A)$ is independent of $\mathcal{F}_s$*

Proposition 2.1.2 *If $\tilde{N}$ is the compensated Poisson random measure of an H-valued Lévy process, then $M(t, A) = \int_A x\tilde{N}_t(dx)$, $A \in \mathcal{R}_1$ defines a Lévy martingale measure on H.*

Similarly as a Wiener process is characterized by its covariance operator, we can describe the covariance structure of a Lévy martingale measure by a family of operators parametrized by our ring $\mathcal{R}_1$.

Proposition 2.1.3

$$\mathbb{E}[|\langle M(t, A), v\rangle|^2] = t\langle v, T_A v\rangle$$

for all $t \geq 0$, $v \in H$ $A \in \mathcal{R}_1$, where the operators T_A are given by $T_A v := \int_A T_x v\mu(dx)$ and $T_x v := \langle x, v\rangle x$.

We will establish a limited theory of integration only, as for our purposes it will be sufficient to integrate deterministic operator valued functions. We do not even need them to depend on the jump size. We follow the general approach, so let us introduce the space of our integrands, the approximating simple functions, and state how the integral is defined for them. For convenience, we set $M([s, t], A) := M(t, A) - M(s, A)$.

Definition 2.1.4 *Let H' be another real separable Hilbert space, $0 \leq T_- < T_+$. Let $\mathcal{H}^2$ be the space of all $R : [T_-, T_+] \times \{x \in H \mid \|x\| \leq 1\} \to \mathcal{L}(H, H')$ such that R is strongly measurable and we have:*

$$\|R\|_{\mathcal{H}^2} := \left(\int_{T_-}^{T_+} \int_{\|x\| \leq 1} \mathrm{Tr}(R(t,x) T_x R^*(t,x)) \mu(dx) dt \right)^{\frac{1}{2}} < \infty$$

Let $\mathcal{S}$ be the space of all functions $R \in \mathcal{H}^2$ such that

$$R = \sum_{i=0}^{n} \sum_{j=0}^{n} R_{ij}\ \chi_{(t_i, t_{i+1}]} \chi_{A_j}$$

where $T_- = t_0 < t_1 < ... < t_{n+1} = T_+$ for some $n \in \mathbb{N}$, where the $A_j \in \mathcal{R}_1$ are pair wisely disjoint and where each $R_{ij} \in \mathcal{L}(H, H')$.

For each $R \in \mathcal{S}$ define the stochastic integral as follows:

$$I(R) := \sum_{i=0}^{n} \sum_{j=0}^{n} R_{ij} M([t_i, t_{i+1}], A_j)$$

Remark 2.1.5 *We can also write $\| \cdot \|_{\mathcal{H}^2}$ as:*
$\|R\|_{\mathcal{H}^2} = \left(\int_{T_-}^{T_+} \int_{\|x\| \leq 1} \|R(t,x)x\|^2 \mu(dx) dt \right)^{\frac{1}{2}}$

Proposition 2.1.6 *The space $\mathcal{H}^2$ with inner product*

$$\langle R, U \rangle := \int_{T_-}^{T_+} \int_{\|x\| \leq 1} Tr(R(t,x) T_x U^*(t,x)) \mu(dx) dt \qquad R, U \in \mathcal{H}^2$$

is a Hilbert space.

Proposition 2.1.7 *The space $\mathcal{S}$ is dense in $\mathcal{H}^2$.*

Proposition 2.1.8 *We have for any $R \in \mathcal{S} : \mathbb{E}[I(R)] = 0$ and*

$$\mathbb{E}[\|I(R)\|^2] = \int_{T_-}^{T_+} \int_{\|x\| \leq 1} \mathrm{Tr}(R(t,x) T_x R^*(t,x)) \mu(dx) dt = \|R\|^2_{\mathcal{H}^2}$$

So, $I : \mathcal{S} \to L^2(\Omega, \mathcal{F}, P; H)$ is an isometry.

So we can isometrically extend the operator I from $\mathcal{S}$ to its closure $\mathcal{H}^2$.

Remark 2.1.9 *This approach to stochastic integration with respect to a Lévy process is closely related to the one developed in [40]. In particular the relevant isometry is the same in our case. While the approach used here allows for integrands depending on the jump size, the theory in [40] is more general in other aspects.*

2.2 Stochastic Convolution

We want to give meaning to the integral

$$X_{U,B} := \int_s^t U(t,r)B(r)dL(r)$$

which we will call a stochastic convolution. Here L is an H-valued Lévy process and we have $U(t,r) \in \mathcal{L}(H), B(r) \in \mathcal{L}(H)\ \forall\ s \le r \le t$. In anticipation of the assumptions in Section 4 we will pose the following conditions:

- $\sup_{r\in\mathbb{R}} \|B(r)\|_{\mathcal{L}(H)} < \infty$
- there are $M > 0, \omega > 0$ such that : $\|U(t,r)\|_{\mathcal{L}(H)} \le Me^{-\omega(t-r)}\ \forall\ t \ge r$
- $r \mapsto B(r)$ is measurable and $r \mapsto U(t,r)$ is measurable for any fixed t

Proposition 2.2.1 *If U and B are as above, the stochastic convolution exists in the following sense:*

$$\begin{aligned}
&\int_s^t U(t,r)B(r)dL(r) \\
&= \int_s^t U(t,r)B(r)b\,dr + \int_s^t \int_{\|x\|\ge 1} U(t,r)B(r)x\ N_r(dx)dr \qquad (2.2)\\
&+ \int_s^t U(t,r)B(r)dW_R(r) + \int_s^t \int_{\|x\|<1} U(t,r)B(r)x\ \tilde{N}_r(dx)dr
\end{aligned}$$

PROOF The proof is analogous to the one of Theorem 6 in [5] where $U(t,s) = S(t-s)$ for a strongly continuous semigroup S: the first term in (2.2) is well defined as a simple Bochner integral, and the second one as a finite random sum. The definition of the third one is well-known, and for the last one it is straightforward to check that for every $t \ge s\ \ r \mapsto U(t,r)B(r)$ is in the space $\mathcal{H}^2$, since U and B are bounded in operator norm. ∎

Chapter 3

Solving the generalized Ornstein-Uhlenbeck Equation

In this chapter we solve our generalized Ornstein-Uhlenbeck equation in the mild and the (analytically) weak sense. For the mild approach we need our time-dependent linear operators to generate a two-parameter semigroup with good smoothing properties. Therefore we will have to deal with a non-autonomous (deterministic) abstract Cauchy problem. Unlike in the autonomous case there is no easy characterization of well-posedness in the sense of the Hille-Yosida theorem available. There are different, yet technical, approaches (see [34] and the references therein for a recent overview), but since this subject is not in the primary interest of our work, we assume that the problem is well posed. This is closely related to the notion of *evolution semigroups*. Our definition is taken from [10].
We will, however, give the example of a time-dependent differential operator generating a two-parameter semigroup which satisfies all our abstract assumptions.

3.1 Existence of the Mild Solution

Let us lay out the framework we will be working in. We consider the following non-autonomous generalisation of the Langevin equation on a Hilbert space H:

$$\begin{cases} dX_t &= (A(t)X_t + f(t))dt + B(t)dL_t \\ X_s &= x \end{cases} \tag{3.1}$$

where $B : \mathbb{R} \to \mathcal{L}(H)$ is strongly continuous and bounded in operator norm, $f : \mathbb{R} \to H$ is uniformly Hölder continuous and bounded, L is an H-valued Lévy-process and where the $A(t)$ are linear operators on H with common and dense domain $D(A) \subset H$ and $A : \mathbb{R} \times D(A) \to H$ is such that we can solve the associated non-autonomous abstract Cauchy problem

$$\begin{cases} dX_t &= (A(t)X_t + f(t))dt \\ X_s &= x \end{cases} \tag{3.2}$$

according to the following definition:

Definition 3.1.1 *An exponentially bounded evolution family on H is a two parameter family $\{U(t,s)\}_{t\geq s}$ of bounded linear operators on H such that we have:*

(i) $U(s,s) = Id$ *and* $U(t,s)U(s,r) = U(t,r)$ *whenever* $r \leq s \leq t$

(ii) for each $x \in H$, $(t,s) \mapsto U(t,s)x$ *is continuous on* $s \leq t$

(iii) there are $M > 0$ *and* $\omega > 0$ *such that :* $\|U(t,s)\|_{\mathcal{L}(H)} \leq Me^{-\omega(t-s)}$, $s \leq t$

Assumption 3.1.2 *There is a unique solution to* (3.2) *given by an exponentially bounded evolution family $U(t,s)$ so that the solution takes the form:*

$$X_t = U(t,s)x + \int_s^t U(t,r)f(r)dr$$

Moreover, we assume that :

$$\frac{d}{dt}U(t,s)x = A(t)U(t,s)x \quad \text{whenever } t > s \quad \text{or } x \in D(A)$$

Concerning the adjoint operators, we require that $A^*(t)$ also have a common domain independent of t and dense in H, which we will denote by $D(A^*)$. Furthermore, we have strong continuity of $t \mapsto A(t)$ as well as of $t \mapsto A^*(t)$. These assumptions will be in force throughout the whole paper.

Remark 3.1.3 *Note that in the finite dimensional case, where each $A(t)$ is automatically bounded, we get the existence of an evolution family that solves* (3.2) *under the reasonable assumption that $t \mapsto A(t)$ is continuous and bounded in the operator norm, by solving the following matrix-valued ODE:*

$$\begin{cases} \frac{\partial}{\partial t}U(t,s) = A(t)U(t,s) \\ U(s,s) = Id \end{cases}$$

Existence and uniqueness are assured since $(t,M) \mapsto A(t)M$ is globally Lipschitz in M. This result even holds in infinite dimensions, see [13].

Example 3.1.4 *As a nontrivial example of a two parameter semigroup fulfilling Assumption 3.1.2 we consider the solution of a parabolic PDE with time-dependent coefficients.*
Let O be a bounded domain in $\mathbb{R}^n$, such that ∂O is of class C^2. Let $H := L^2(O,dx)$. Let $H^2 := H^{2,2}$ and $H_0^1 := H_0^{1,2}$ be the Sobolev spaces of order 2 and

of order 1 with Dirichlet boundary conditions, respectively. For $g \in H^2 \cap H_0^1$ *let* $A(t)$ *be given by second order differential operators in divergence form:*

$$A(t)g(x) := \sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ij}(t,x)\frac{\partial}{\partial x_j}g\right)(x)$$

We impose the following conditions on the coefficients:

- $(t,x) \mapsto a_{ij}(t,x)$ *is defined on* $\mathbb{R} \times \bar{O}$ *and is continuous* $\forall i,j$.
- $(t,x) \mapsto \frac{\partial}{\partial x_i}a_{ij}(t,x)$ *is defined on* $\mathbb{R} \times \bar{O}$ *and is continuous* $\forall i,j$.
- $t \mapsto a_{ij}(t,x)$ *and* $t \mapsto \frac{\partial}{\partial x_i}a_{ij}(t,x)$ *are Hölder continuous* $\forall i,j$ *uniformly in* x.
- *there is* $c > 0$ *s.t.* $\sum_{i,j=1}^{n} y_i a_{ij}(t,x) y_j \geq c\,\|y\|_2^2 \quad \forall y \in \mathbb{R}^n,\ t \in \mathbb{R},\ x \in \bar{O}$

By Theorem 9.1 in part 2 of [24] we obtain existence and uniqueness of a solution to the Cauchy problem (3.2)*. Moreover, the solution is given by a strongly continuous evolution family, say* U*. Strong continuity of* $t \mapsto A(t)$ *is clear by continuity of the coefficients. As the adjoints are explicitly given through integration by parts, the assumption on the* $A^*(t)$ *is also clearly fulfilled. It remains to show, that this family is also exponentially bounded. Hence we compute, for a fixed* $f \in H$ *with* $g := U(t,s)f$

$$\begin{aligned}\frac{d}{dt}\|U(t,s)f\|_H^2 &= 2\langle A(t)g,g\rangle_H = -2\int_O\left(\sum_{i,j=1}^{n} a_{ij}(t,x)\frac{\partial}{\partial x_i}g(x)\frac{\partial}{\partial x_j}g(x)\right)dx \\ &\leq -2c\int_O \langle\nabla g(x),\nabla g(x)\rangle_{\mathbb{R}^n}dx \leq -2c\int_O |g(x)|^2dx = -2c\|U(t,s)f\|_H^2\end{aligned}$$

where we employed the Poincaré inequality and absorbed its constant into c*. Thus, the result follows by a Gronwall type argument.*

Definition 3.1.5 *Given Assumption 3.1.2 we call the process:*

$$X(t,s,x) = U(t,s)x + \int_s^t U(t,r)f(r)dr + \int_s^t U(t,r)B(r)dL_r$$

a mild solution to equation (3.1).

3.2 Existence of the Weak Solution

We have called the above expression a mild solution, though there is no obvious relation to the equation yet. Now, we will show that our candidate solution actually solves our equation in the weak sense. The following definition makes this precise:

Definition 3.2.1 *An H-valued process X_t is called a weak solution to equation* (3.1) *if for every $y \in D(A^*)$ and P-almost every $\omega \in \Omega$ we have :*

$$\langle X_t, y\rangle = \langle x, y\rangle + \int_s^t \langle X_r, A^*(r)y\rangle dr + \int_s^t \langle f(r), y\rangle dr + \int_s^t B^*(r)y dL_r \quad , t > s$$

In particular $P-a.s.$ the integrals $\int_s^t \langle X_r, A^*(r)y\rangle dr$ must exist for every $t > s$. Here $(B^*(r)y)(h) := \langle B^*(r)y, h\rangle$ so that $B^*(r)y \in \mathcal{L}(H, \mathbb{R})$ and the integral is well defined, since
$\|B^*(r)y\|_H^2 \leq (T_+ - T_-) \sup_r \|B(r)\|^2_{\mathcal{L}(H)} \|y\|^2 \int_{\|x\|<1} \|x\|^2 \mu(dx) < \infty.$

Theorem 3.2.2 *Let Assumption 3.1.2 hold. Then the mild solution X_t from Definition 3.1.5 is also a weak solution for* (3.1).

Proof By calculating:

$$\int_s^t \langle U(r,s)x, A^*(r)y\rangle dr = \int_s^t \frac{d}{dr}\langle U(r,s)x, y\rangle dr = \langle U(t,s)x - x, y\rangle$$

$$\int_s^t \left\langle \int_s^r U(r,u)f(u)du, A^*(r)y \right\rangle dr = \int_s^t \int_s^r \frac{d}{dr} \langle U(r,u)f(u), y\rangle \, dudr$$
$$= \int_s^t \langle f(u), [U^*(t,u) - Id]y\rangle \, du = \int_s^t \langle U(t,u)f(u), y\rangle \, du - \int_s^t \langle f(u), y\rangle \, du$$

the problem is reduced to proving, for all $y \in D(A^*)$, the following equality:

$$\left\langle \int_s^t U(t,r)B(r)dL_r, y \right\rangle = \int_s^t \left\langle \int_s^r U(r,u)B(u)dL_u, A^*(r)y \right\rangle dr + \int_s^t B^*(u)y dL_u \tag{3.3}$$

To this end, let us first assume the case of a Lévy process where the second term in (2.2) is 0, so that the jumps are bounded in modulus by 1. We need to approximate U and B by functions which are simple in u.
Note that, thanks to continuity of U and B, we can choose our partition independent of r. Hence, for $n \in \mathbb{N}$ and $0 \leq k < 2^n$ let $u_k := s + (t-s)k2^{-n}$.
Define $U^{(n)}(r,u) := U(r,u_k)$ if $u_k \leq u < u_{k+1}$ and respectively:
$B^{(n)}(u) := B(u_k)$ if $u_k \leq u < u_{k+1}$. Hence:

$$\begin{aligned}
&\int_s^t \left\langle \int_s^r U^{(n)}(r,u)B^{(n)}(u)dL_u, A^*(r)y \right\rangle dr \qquad (*)\\
&= \int_s^t \sum_{k=0}^{2^n-1} \chi_{\{u_k<r\}} \langle U(r,u_k)B(u_k)(L_{u_{k+1}\wedge r} - L_{u_k}), A^*(r)y\rangle dr\\
&= \sum_{k=0}^{2^n-1} \int_s^t \chi_{\{u_k<r\}} \langle U(r,u_k)B(u_k)(L_{u_{k+1}\wedge r} - L_{u_k}), A^*(r)y\rangle dr\\
&= \sum_{k=0}^{2^n-1} \int_{u_k}^t \langle A(r)U(r,u_k)B(u_k)(L_{u_{k+1}} - L_{u_k}), y\rangle dr + R(n)\\
&= \sum_{k=0}^{2^n-1} \int_{u_k}^t \frac{d}{dr} \langle U(r,u_k)B(u_k)(L_{u_{k+1}} - L_{u_k}), y\rangle dr + R(n)\\
&= \sum_{k=0}^{2^n-1} \langle (U(t,u_k) - Id)B(u_k)(L_{u_{k+1}} - L_{u_k}), y\rangle + R(n)\\
&= \left\langle \int_s^t U^{(n)}(t,u)B^{(n)}(u)dL_u, y \right\rangle - \left\langle \int_s^t B^{(n)}(u)dL_u, y \right\rangle + R(n)
\end{aligned}$$

where $R(n) := \sum_{k=0}^{2^n-1} \int_{u_k}^{u_{k+1}} \langle U(r,u_k)B(u_k)(L_r - L_{u_{k+1}}), A^*(r)y\rangle dr$.

Taking the limit $n \to \infty$ in the equality above we will now establish (3.3). First, define $V^{(n)}(t,u) := U^{(n)}(t,u)B^{(n)}(u) - U(t,u)B(u)$ and observe that $\lim_{n\to\infty} \|V^{(n)}(t,u)x\| = 0$ for $t \geq u$ and $x \in H$ fixed. In order to show $\lim_{n\to\infty} \int_s^t V^{(n)}(t,u)dL_u = 0$ in $L^2(\Omega, P; H)$ we need to prove:

- $\lim_{n\to\infty} \int_s^t \|V^{(n)}(t,u)b\|du = 0$
- $\lim_{n\to\infty} \left(\int_s^t \sum_k \|V^{(n)}(t,u)(\sqrt{R})^* e_k\|^2 du\right)^{\frac{1}{2}} = 0$
- $\lim_{n\to\infty} \left(\int_s^t \int_{\|x\|\leq 1} \|V^{(n)}(t,u)x\|^2 \mu(dx)du\right)^{\frac{1}{2}} = 0$

where $(e_k)_{k\in\mathbb{N}}$ is an ONB of H and where b, R and μ are as in (2.1). In each case pointwise convergence of the integrand is clear and a majorizing expression is easily found remembering that U and B are uniformly bounded in Operator norm and $\sqrt{R}$ is Hilbert-Schmidt. The proof for $\lim_{n\to\infty} \int_s^t B^{(n)}(u)dL_u = \int_s^t B(u)dL_u$ is essentially the same.

For the expression in $(*)$ (understood as a Bochner integral in $L^2(\Omega, P; \mathbb{R})$) we have to show: $\lim_{n\to\infty} \int_s^t \left\| \left\langle \int_s^r V^{(n)}(t,u)dL_u, A^*(r)y \right\rangle \right\|_{L^2(\Omega,P;\mathbb{R})} dr = 0$ Pointwise convergence of the integrand in $L^2(\Omega, P; \mathbb{R})$ follows from the results above, and an integrable upper bound is given by:

$$\sqrt{(r-s)}\sup_u \|B(u)\|_{\mathcal{L}(H)}\left[\|b\|+\sqrt{\mathrm{Tr}(R)}+\left(\int_{\|x\|\le 1}\|x\|^2\mu(dx)\right)^{\frac{1}{2}}\right]\sup_u \|A^*(u)y\|$$

Finally, we show that R_n tends to 0 in $L^1(\Omega, P;\mathbb{R})$:

$$\mathbb{E}|R_n| \le \mathbb{E}\left[\sum_{k=0}^{2^n-1}\int_{u_k}^{u_{k+1}}\|U(r,u_k)B(u_k)(L_r - L_{u_{k+1}})\|\|A^*(r)y\|dr\right]$$

$$\le \underbrace{M\sup_{t\in\mathbb{R}}\|B(t)\|_{\mathcal{L}(H)}\sup_{t\in\mathbb{R}}\|A(t)y\|}_{:=C}\sum_{k=0}^{2^n-1}\int_{u_k}^{u_{k+1}}\mathbb{E}\left[\|L_r - L_{u_{k+1}}\|\right]dr$$

$$\le C\sum_{k=0}^{2^n-1}\int_{u_k}^{u_{k+1}}\left(\mathbb{E}[\|(L_{u_{k+1}-r}-(u_{k+1}-r)\mathbb{E}[L_1]\|]+\|\mathbb{E}[L_1]\|(u_{k+1}-r)\right)dr$$

Now we have $\lim_{t\to 0}\mathbb{E}[\|L_t - t\mathbb{E}[L_1]\|] = 0$. Convergence in probability is clear by definition of a Lévy process and uniform integrability follows by Doob's maximal inequality, since $\|L_t - t\mathbb{E}[L_1]\|$ is a right continuous positive submartingale. The result then follows by an easy $\varepsilon - \delta$ argument.

To complete the proof consider a Lévy process such that (2.2) consist only of the second term (pure and discrete jump case). Due to the discrete nature of the jumps, there is no need for the above approximation procedure and one can prove (3.3) ω by ω, repeating the steps following equation $(*)$. ■

Remark 3.2.3 *Due to the non-autonomous setting, the proof above is rather technically involved. The difficulties stem from the fact that $(t,s)\mapsto U^*(t,s)$ is in general only weakly continuous and weakly differentiable and $U^*(t,s)D(A^*)\subset D(A^*)$ does not hold. For the same reason, we were unable to prove the converse of Theorem 3.2.2. If U^* has the same properties as U (as in Example 5.1.4 where $U = U^*$), one can indeed show (following the proof in the autonomous case as in [19] Theorem 5.4) that the weak solution to* (3.1) *is also a mild solution and hence unique.*

Proposition 3.2.4 *Assume that $s\mapsto U^*(t,s)\xi \in C^1([0,T], D(A^*))$ for any $\xi \in D(A^*)$. Then the weak solution to* (3.1) *is unique.*

Proof For simplicity, let us assume that we have a weak solution X_t to (3.1) with $X(0) = 0$ and $f \equiv 0$.
Claim: Then for any function $\zeta \in C^1([0,T], D(A^*))$ we have

$$\langle X_t, \zeta(t)\rangle = \int_0^t \langle X_r, A^*(r)\zeta(r) + \zeta(r)\rangle dr + \int_0^t B^*(r)\zeta(r)dL_r.$$

By a density argument it is sufficient to show this for functions $\zeta(t) := \phi(t)\zeta_0$ with $\phi \in C^1([0,T])$, $\zeta_0 \in D(A^*)$.

Let us define the process

$$F_{\zeta_0}(t) := \int_0^t \langle X_s, A^*(s)\zeta_0\rangle ds + \langle B^*(s)\zeta_0, L_s\rangle.$$

Since X_s is a weak solution we have $F_{\zeta_0}(t) = \langle \zeta_0, X_s\rangle$ almost surely. Applying the Ito formula to the process $F_{\zeta_0}(s)\phi(s)$ we get

$$d\left[F_{\zeta_0}(s)\phi(s)\right] = \phi(s)dF_{\zeta_0}(s) + \phi'(s)F_{\zeta_0}(s)ds.$$

In other terms

$$F_{\zeta_0}(t)\phi(t) = \int_0^t \phi(s)\langle X_s, A^*(s)\zeta_0\rangle ds + \int_0^t \phi(s)\langle B^*(s)\zeta_0, dL_s\rangle + \int_0^t \phi'(s)F_{\zeta_0}(s)ds.$$

Replacing $F_{\zeta_0}(t)$ with $\langle \zeta_0, X_s\rangle$ and remembering that $\zeta(t) := \phi(t)\zeta_0$ we end up with

$$\langle \zeta(t), X_t\rangle = \int_0^t \langle X_s, A^*(s)\zeta(s) + \zeta'(s)\rangle ds + \int_0^t \langle B^*(s)\zeta(s), dL_s\rangle$$

and the claim is proved.
To prove the proposition, we apply the claim with $\zeta(s) := U^*(t,s)\zeta_0$. Since $\frac{d}{ds}U^*(t,s)\zeta_0 = -A^*(s)U^*(t,s)\zeta_0$ we obtain:

$$\langle X_t, \zeta_0\rangle = \int_0^t \langle U^*(t,s)\zeta_0, B(s)dL_s\rangle.$$

As $D(A^*)$ is assumed to be dense and $\zeta_0 \in D(A^*)$ is arbitrary, X must coincide with the mild solution. ■

The following is a counterexample (sketched in [10]) illustrating that – as mentioned in the remark above – good behavior of $U(t,s)$ is not necessarily inherited by $U^*(t,s)$.

Example 3.2.5 *Let $H = L^2(\mathbb{R}_+, dx)$ and for $f \in H$ and $t \geq s \geq 0$ let*

$$U(t,s)f(x) = u(t)u(s)^{-1}f(x)$$

where $u(t)f(x) := f(x) - \frac{1}{2}f(x+\frac{1}{t})1_{\{t>0\}}$.
Then $t \mapsto u(t)$ is strongly continuous since

$$\lim_{t\to 0}\int_0^\infty f^2\left(x+\frac{1}{t}\right)dx = \lim_{t\to 0}\int_{\frac{1}{t}}^\infty f^2(x)dx = 0$$

by dominated convergence. Moreover, for any $t \geq 0$, $u(t)$ is invertible as a Neumann series (since $\|\frac{1}{2}f(\cdot+\frac{1}{t})1_{\{t>0\}}\|_H = \frac{1}{2}\|f\|_H$), and the continuity of $t \mapsto u^{-1}(t)f$ can be seen first for continuous f in a similar spirit as above and

then for general f since $u^{-1}(t)$ is bounded for any $t \geq 0$.
Thus $(t,s) \mapsto U(t,s)$ is a strongly continuous evolution semigroup.
For the adjoints we have

$$u^*(t)f(x) = f(x) - 1_{\{xt\geq 1\}}\frac{1}{2}f(x-\frac{1}{t})$$

but since

$$\lim_{t\to 0}\int_0^\infty f^2(x-\frac{1}{t})dx = \int_0^\infty f^2(x)dx \neq 0$$

we do not have strong continuity for $t \mapsto u^*(t)$ and thus neither for

$$(t,s) \mapsto U^*(t,s) = u^*(s)^{-1}u^*(t)$$

.

Chapter 4

The Semigroup and a generalized Invariant Measure

In this chapter we show that the solution established in the last chapter induces a two-parameter semigroup of operators. We prove existence of an evolution system of measures for this semigroup in Section 4.2 and with its help define a generalized invariant measure for a corresponding one-parameter semigroup on an extended state space in Section 4.3. On the L^2 space with respect to this measure we then prove in the last section that the one-paramter semigroup is strongly continuous.

4.1 The semigroup property

Recall that the mild solution for (3.1) takes the following form

$$X(t,s,x) = U(t,s)x + \int_s^t U(t,r)f(r)dr + \int_s^t U(t,r)B(r)dL_r.$$

As opposed to the Gaussian case we are no longer able to give an easy representation of the law of $X(t,s,x)$, but we can calculate its Fourier transform. The following lemma will be an extremely helpful tool in the calculations to come.

Lemma 4.1.1 (characteristic function)

$$\mathbb{E}\left[\exp\left(i\left\langle h, X(t,s,x)\right\rangle\right)\right] = \\ \exp\left\{i\left\langle h, U(t,s)x + \int_s^t U(t,r)f(r)dr\right\rangle\right\}\exp\left\{\int_s^t \lambda(B^*(r)U^*(t,r)h)dr\right\}$$

where λ is the Lévy symbol of L.

Proof : We proceed by using the isometries to approximate the stochastic integral by a sum, and then using independence of increments and the Lévy-Khinchin formula.

First note that knowing how the Fourier transform acts on translations, it will be enough to show that:

$$\mathbb{E}\left[\exp\left(i\left\langle h, \int_s^t U(t,r)B(r)dL_r\right\rangle\right)\right] = \exp\left\{\int_s^t \lambda(B^*(r)U^*(t,r)h)dr\right\}$$

The strong continuity of U and B allows us to approximate the Lévy stochastic integral by a sequence of sums. More precisely, we have:

$$\int_s^t U(t,r)B(r)dL_r = P - \lim_{n\to\infty} \sum_{s_i\in\mathcal{P}_n} U(t,s_i)B(s_i)(L_{s_i} - L_{s_{(i-1)\vee 0}})$$

where the limit is taken in probability and $\mathcal{P}_n$ is a sequence of partitions $s = s_0 < ... < s_N = t$ of $[s,t]$ such that the mesh width tends to zero. The verification is a straightforward application of the respective isometries.

For the drift term which is a Bochner integral we have to show that:

$$\lim_{n\to\infty} \sum_{s_i\in\mathcal{P}_n} \int_{s_{i-1}}^{s_i} \|U(t,s_i)B(s_i)b - U(t,r)B(r)b\|\, dr = 0$$

but since $r \mapsto U_t(r)B(r)b$ is even uniformly continuous on $[s,t]$ we may find $\delta > 0$ such that $\|U_t(r)B(r)b - U_t(r')B(r')b\| < \frac{\varepsilon}{t-s}$ whenever $|r - r'| < \delta$, so that if we choose n such that the mesh width of $\mathcal{P}_n$ is smaller than δ we have

$$\sum_{s_i\in\mathcal{P}_n} \int_{s_{i-1}}^{s_i} \|U(t,s_i)B(s_i)b - U(t,r)B(r)b\|\, dr < \sum_{s_i\in\mathcal{P}_n} \int_{s_{i-1}}^{s_i} \frac{\varepsilon}{t-s} dr < \varepsilon$$

For the small jumps we make use of the isometry from 2.1.8, so we have to show that our piecewise approximation converges in the $\mathcal{H}^2$ norm, that is we need:

$$\lim_{n\to\infty} \sum_{k=1}^{\infty} \int_{\|x\|<1} \left(\sum_{s_i\in\mathcal{P}_n} \int_{s_{i-1}}^{s_i} \|[U_t(r)B(r) - U_t(s_i)B(s_i)]T_x^{\frac{1}{2}} e_k\|^2 dr \right) M(dx) = 0$$

For each k and x fixed the expression in round brackets converges to zero, for the same reasons as used for the drift term. So we only have to show that we may take the limit into the sum and the integral, but this follows by dominated convergence on considering the uniform integrable bound :

$$\|[U_t(r)B(r) - U_t(s_i)B(s_i)]T_x^{\frac{1}{2}} e_k\|^2 \le 2 \sup_{s\le r\le t} \|U_t(r)\|_{\mathcal{L}(H)} \sup_{s\le r\le t} \|B(r)\|_{\mathcal{L}(H)} \|T_x^{\frac{1}{2}} e_k\|^2$$

Thus we have convergence in L^2 of the approximating sums towards the integral.
The same argument works for the Brownian part, where there is even no dependence on x.

The big jumps, finally are quite simple to treat. Since the expression makes sense pointwise, we consider the approximation for ω fixed and we obtain:

$$\lim_{n\to\infty}\sum_{s_i\in\mathcal{P}_n}\sum_{s_{i-1}\le r\le s_i}[U_t(s_i)B(s_i)-U_t(r)B(r)]\Delta L_r(\omega)\chi_{\|\Delta L_r(\omega)\|>1}=0$$

again because of strong continuity.

So in any of the four cases we have at least convergence in probability and the claim is proved.

By convergence in distribution and independence of increments:

$$\begin{aligned}&\mathbb{E}\left[\exp\left(i\left\langle h,\int_s^t U(t,r)B(r)dL_r\right\rangle\right)\right]\\ =&\lim_{n\to\infty}\prod_{k\in\mathcal{P}_n}\mathbb{E}\left[\exp\left(i\left\langle h,U(t,s_k)B(s_k)(L_{s_k}-L_{s_{k-1}\vee 0})\right\rangle\right)\right]\\ &=\exp\left\{\int_s^t\lambda(B^*(r)U^*(t,r)h)dr\right\}\end{aligned}$$

where we employed the Lévy-Khinchine formula and the functional equation of the exponential. Note that the Riemannian sums converge to the integral because of strong continuity.

■

The following lemma is a straightforward generalization of the standard monotone class theorem.

Lemma 4.1.2 (complex monotone classes) *Let $\mathcal{H}$ be a complex vector space of complex-valued bounded functions, that contains the constants and is closed under componentwise monotone convergence. Let $\mathcal{M}\subset\mathcal{H}$ be closed under multiplication and complex conjugation. Then, all bounded functions measurable with respect to the σ-algebra generated by the functions in $\mathcal{M}$ belong to $\mathcal{H}$.*

The last and the next result in combination will be particularly useful:

Lemma 4.1.3 *The functions $\mathcal{M}:=\{e^{i\langle h,x\rangle},h\in H\}$ form a complex multiplicative system that generates the Borel σ-algebra of H.*

PROOF It is obvious that $\mathcal{M}$ is closed under multiplication and complex conjugation.
To show that indeed $\sigma(\mathcal{M})=\mathcal{B}(H)$ we make use of the following lemma: (see [44] page 108)

Lemma 4.1.4 *A countable family of real-valued functions on a Polish space X separating the points of X already generates the Borel-sigma-algebra of X.*

Our countable family will be $\{f_{n,k}(x):=\sin(\langle\frac{1}{n}e_k,x\rangle)\}_{k,n\in\mathbb{N}}\subset\mathcal{M}$ where $\{e_k\}$ is an orthonormal basis of H.

Since the sine function is injective in a neighborhood of zero, and the functions $\langle \frac{1}{n} e_k, x \rangle)$ separate the points of H, so do the $f_{n,k}$. As real and imaginary parts of functions in $\mathcal{M}$, it is clear, that the sigma-algebra generated by them is included in $\sigma(\mathcal{M})$. ∎

Now we will show that our solution induces a two-parameter transition evolution semigroup, defined as follows:

Definition 4.1.5 *Whenever $f : H \to \mathbb{C}$ is measurable and bounded, define :*

$$P(s,t)f(x) := \mathbb{E}[f(X(t,s,x))] \qquad (t \geq s)$$

$P(s,t)$ will be called the two-parameter semigroup (associated to the solution X).

Lemma 4.1.6 *For f as above, we have the following flow property, i.e. $P(s,t)$ satisfies the Chapman-Kolmogorov equation:*

$$P(r,s)P(s,t)f(x) = P(r,t)f(x) \qquad (t \geq s \geq r)$$

Moreover, $P(s,t)$ is Feller, mapping $C_b(H)$ into itself.

Proof We will show the equality for the functions $f_h(x) = e^{i\langle h,x \rangle}$ and extend it with the help of Lemma 4.1.2. First note, that by Lemma 4.1.1 we have

$$P(s,t)f_h(x) = \exp\left\{ i \left\langle h, U(t,s)x + \int_s^t U(t,r)f(r)dr \right\rangle + \int_s^t \lambda(B^*(r)U^*(t,r)h)dr \right\}$$

so that:

$$\begin{aligned} P(r,s)P(s,t)f_h(x) &= \mathbb{E}[P(s,t)f_h(X(s,r,x))] \\ &= \mathbb{E}\left[\exp\left\{ i \langle U^*(t,s)h, X(s,r,x) \rangle \right\}\right] \\ &\quad \times \exp\left\{ i \left\langle h, \int_s^t U(t,r)f(r)dr \right\rangle + \int_s^t \lambda(B^*(r)U^*(t,r)h)dr \right\} \end{aligned}$$

but again Lemma 4.1.1 gives us the Fourier transform of $X(s,r,x)$ this time evaluated at $U^*(t,s)h$:

$$\begin{aligned} &= \exp\left\{ i \langle U^*(t,s)h, U(s,r)x \rangle \right\} \\ &\quad \times \exp\left\{ i \left\langle U^*(t,s)h, \int_r^s U(s,q)f(q)dq \right\rangle \right\} \exp\left\{ i \left\langle h, \int_s^t U(t,r)f(r)dr \right\rangle \right\} \\ &\quad \times \exp\left\{ \int_r^s \lambda(B^*(q)U^*(s,q)U^*(t,s)h)dq \right\} \exp\left\{ \int_s^t \lambda(B^*(r)U^*(t,r)h)dr \right\} \end{aligned}$$

Interchanging $U(t,s)$ with the integral, as it is a bounded operator, making use of the semigroup property of U and U^* and combining the integrals yields the result for exponential f. By monotone convergence, it is easy to see that

the space of all bounded measurable f for which the flow property holds is a complex monotone vector space. Hence, the first assertion is proved.
For the second assertion we make use of a Mehler formula for the semigroup.

$$\begin{aligned} P(s,t)f(x) =& \mathbb{E}\left(\left(f(U(t,s)x + \int_s^t U(t,r)f(r)dr + \int_s^t U(t,r)B(r)dLr\right)\right) \\ =& \int_H f\left(U(t,s)x + \int_s^t U(t,r)f(r)dr + y\right)\mu_{s,t}(dy) \end{aligned}$$

where $\int_s^t U(t,r)B(r)dLr \sim \mu_{s,t}$. Since $U(t,s)$ is strongly continuous and since $f \in C_b(H)$ we have

$$P(s,t)f(x_n) \to P(s,t)f(x) \quad \text{as} \quad x_n \to x$$

by dominated convergence. ■

Remark 4.1.7 *Note that Lemma 4.1.6 is equivalent to the Markov property (with respect to the natural filtration) for our solution, but in our case its direct proof seems to be even more difficult. The Feller property will be strengthened in Corollary 5.2.7 as a consequence of a Harnack inquality for $P(s,t)$.*

4.2 Evolution Systems of Measures

Since our equation is non-autonomous we cannot hope for a single invariant measure. What one can still expect in our setting is a so called *evolution system of measures*, a whole family $\{\nu_t\}_{t\in\mathbb{R}}$ of probability measures such that for all $s < t$ and all bounded measurable ψ:

$$\int_H P(s,t)\psi(x)\nu_s(dx) = \int_H \psi(x)\nu_t(dx) \tag{4.1}$$

The use of such a system goes back at least to Dynkin [20] where it appears under the name of an entrance law. See also [17] and of course [14] which served as a direct motivation.

We want to assure the existence of such a system, therefore we require the following condition on the jump part of our driving Lévy process.

Assumption 4.2.1 *In addition to Assumption 3.1.2 we require that for the Lévy measure μ we have:*

$$\int_{\|x\|>1} \|x\|\mu(dx) < \infty$$

Remark 4.2.2 *Note that in the autonomous case the corresponding condition $\int_{\|x\|>1} \log(\|x\|)\mu(dx) < \infty$ (see [11]) is considerably weaker.*

The following lemma will provide a useful growth condition for the Lévy symbol that will allow us to construct limit measures. A proof is contained in the appendix.

Lemma 4.2.3 *Every Lévy symbol λ with a Lévy measure μ satisfying Assumption 4.2.1 is Fréchet differentiable. In particular such a λ is locally Lipschitz continuous.*

PROOF By the Lévy-Khinchine formula (2.1) we know that:

$$\lambda(u) = i\langle u, b\rangle - \frac{1}{2}\langle u, Ru\rangle + \int_{H\setminus\{0\}} \left(e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle\chi_{\{\|x\|\leq 1\}}\right)\mu(dx)$$

Clearly, it is enough to show that the integral expression is differentiable. We first show Gâteaux differentiability, hence we will need the directional derivatives to be uniformly integrable to obtain the result via dominated convergence. We have:

$$\frac{\partial}{\partial t}\left(e^{i\langle u+tv,x\rangle} - 1 - i\langle u + tv, x\rangle\chi_{\{\|x\|\leq 1\}}\right) = i\langle v, x\rangle e^{i\langle u+tv,x\rangle} - i\langle v, x\rangle\chi_{\{\|x\|\leq 1\}}$$

To see the uniform integrability in t (for, say, $t \in [0, 1]$) we split the integral in two parts:

$$\begin{aligned}
&\int_{\|x\|\leq 1} \sup_{t\in[0,1]} \left|i\langle v, x\rangle e^{i\langle u+tv,x\rangle} - i\langle v, x\rangle\chi_{\{\|x\|\leq 1\}}\right| \mu(dx)\\
&= \int_{\|x\|\leq 1} \left|i\langle v, x\rangle \sum_{k=0}^{\infty} \frac{(i\langle u, x\rangle)^k}{k!} - i\langle v, x\rangle\right| \mu(dx)\\
&\leq \int_{\|x\|\leq 1} \left(\|v\|\,\|x\| \sum_{k=1}^{\infty} \frac{|(i\langle u, x\rangle)|^k}{k!}\right) M(dx)\\
&\leq \int_{\|x\|\leq 1} \left(\|v\|\,\|x\| \sum_{k=1}^{\infty} \frac{\|u\|^k\|x\|^k}{k!}\right) M(dx)\\
&\leq \int_{\|x\|\leq 1} \sup_{t\in[0,1]} \left(\|v\|\,\|x\|^2\|u + tv\| \sum_{k=1}^{\infty} \frac{\|u + tv\|^{k-1}\|x\|^{k-1}}{k!}\right) \mu(dx)\\
&\leq \sup_{\|x\|\leq 1} \sup_{t\in[0,1]} \exp\{\|u + tv\|\|x\|\}\, \|u + tv\|\, \|v\| \int_{\|x\|\leq 1} \|x\|^2\mu(dx) = C(u, v) < \infty
\end{aligned}$$

since μ is a Lévy measure. Here $C(u, v)$ is a constant depending only on the fixed u and v. On the other hand, we have:

$$\int_{\|x\|>1} \left|i\langle v, x\rangle e^{i\langle u+tv,x\rangle} - i\langle v, x\rangle\chi_{\{\|x\|\leq 1\}}\right| \mu(dx) \leq \|v\| \int_{\|x\|>1} \|x\|\mu(dx) < \infty$$

■

by assumption.
Moreover, from the above, it is easy to see that the Gâteaux derivative is linear and bounded and depends continuously on u with respect to the operator norm. Thus λ is Fréchet differentiable and hence locally Lipschitz.

Theorem 4.2.4 *Let Assumption 4.2.1 hold. Denote by λ the Lévy symbol of L. Then the functions*

$$\hat{\nu_t}(h) := \exp\left\{ i\left\langle h, \int_{-\infty}^{t} U(t,r)f(r)dr \right\rangle \right\} \exp\left\{ \int_{-\infty}^{t} \lambda\{B^*(r)U^*(t,r)h\}dr \right\}$$

are the Fourier transforms of an evolution system of measures.

PROOF : We have to assure that the integrals above exist. Since U is stable and f is bounded on all of $\mathbb{R}$ we have:

$$\int_{-\infty}^{t} \|U(t,r)f(r)\|dr \leq \int_{-\infty}^{t} Me^{-\omega(t-r)}\|f\|_\infty dr = \frac{M}{\omega}\|f\|_\infty$$

As λ is Fréchet differentiable it has locally linear growth, so that with $\lambda(0) = 0$ we have $\|\lambda(u)\| \leq C\|u\|$ on the bounded range of the argument for some $C > 0$. Hence, as $\|B^*\|$ is bounded, we can treat the second integral like the first:

$$\int_{-\infty}^{t} \|\lambda\{B^*(r)U^*(t,r)h\}\|dr \leq C\sup_r \|B^*(r)\|\frac{M}{\omega}\|h\| < \infty \tag{4.2}$$

where we have used that $\|U^*\| = \|U\|$.

To show that these functions are indeed Fourier transforms of measures we can make use of Lévy's continuity theorem in the finite dimensional case. We have just proven pointwise convergence of the Fourier transforms of $P \circ [X(t,s,x)]^{-1}$, and that the limit function is continuous in 0 follows easily by dominated convergence. Pointwise convergence under the integral is clear by continuity of λ, U and B and a majorizing function is found by looking at (4.2) again.
In the infinite dimensional case, however, we cannot apply Lévy's continuity theorem, because we are unable to prove continuity in the Sazonov topology. For a better readability we postpone the somewhat technical alternative to the end of this proof, formulated as a claim.

In order to see that the respective measures constitute an evolution system of measures we will check (4.1) for exponential functions and then extend the result via monotone classes. So if we take $k(x) = e^{i\langle h,x\rangle}$ in (4.1) we get:

$$\int_H k(x)\nu_t(dx) = \hat{\nu_t}(h)$$

by the very definition of Fourier transformation.
On the other hand we have by Lemma 4.1.1:

$$P(s,t)k(x) = \exp\left\{ i\left\langle h, U(t,s)x + \int_s^t U(t,r)f(r)dr \right\rangle + \int_s^t \lambda\{B^*(r)U^*(t,r)h\}dr \right\}$$

Using the adjoint of U and the fact that Fourier transformation is only with respect to x we obtain by definition of $\hat{\nu_s}$:

$$
\begin{aligned}
&\int_H P(s,t)k(x)\nu_s(dx) \\
&= \hat{\nu_s}(U^*(t,s)h)\ \exp\left\{ i\left\langle h, \int_s^t U(t,r)f(r)dr \right\rangle + \int_s^t \lambda\{B^*(r)U^*(t,r)h\}dr \right\} \\
&= \exp\left\{ i\left\langle U^*(t,s)h, \int_{-\infty}^s U(s,r)f(r)dr \right\rangle + \int_{-\infty}^s \lambda\{B^*(r)U^*(s,r)U^*(t,s)h\}dr \right\} \\
&\qquad\qquad \times \exp\left\{ i\left\langle h, \int_s^t U(t,r)f(r)dr \right\rangle + \int_s^t \lambda\{B^*(r)U^*(t,r)h\}dr \right\} \\
&= \exp\left\{ i\left\langle h, \int_{-\infty}^s U(t,s)U(s,r)f(r)dr \right\rangle + \int_{-\infty}^s \lambda\{B^*(r)U^*(t,r)h\}dr \right\} \\
&\qquad\qquad \times \exp\left\{ i\left\langle h, \int_s^t U(t,r)f(r)dr \right\rangle + \int_s^t \lambda\{B^*(r)U^*(t,r)h\}dr \right\} \\
&= \exp\left\{ i\left\langle h, \int_{-\infty}^t U(t,r)f(r)dr \right\rangle \right\} \exp\left\{ \int_{-\infty}^t \lambda\{B^*(r)U^*(t,r)h\}dr \right\}
\end{aligned}
$$

but the last line equals $\hat{\nu_t}(h)$ and that is precisely what we had to show.

To prove the full assertion we have to show that (4.1) not only holds for functions of the form $k_h(x) := e^{i\langle h,x\rangle}$, but for any bounded measurable function. By Lemma 4.1.3 we can apply Lemma 4.1.2, because the bounded and measurable functions for which (4.1) holds, form a complex monotone vector space: for constant functions the equality is trivial and that (4.1) holds for monotone limits is seen by applying Beppo Levi's theorem on monotone convergence twice. Hence, the existence of an evolution system of measures is proved.

Claim: $\hat{\nu_t}$ is a characteristic function - Hilbert space case

The general idea is the following. In the Gaussian case, it is known that the limit distributions are Gaussian again. In the same manner we take advantage of the fact that, in the Lévy case, our limit distribution is infinitely divisible. We proceed here similarly as in [25] Chapter 3. First of all we show that our distributions $P \circ X(t,s,x)$ are infinitely divisible for any $t > s, x$. By the Lévy-Khinchine representation it is sufficient to prove that their characteristic functions have the form (2.1) for some triple $[b, R, \mu]$. Therefore, we calculate:

$$
\begin{aligned}
&\exp\left\{ \int_s^t \lambda(B^*(r)U^*(t,r)h)dr \right\} \\
&= \exp\left\{ \int_s^t i\langle b, B^*(r)U^*(t,r)h\rangle dr - \frac{1}{2}\int_s^t \langle B^*(r)U^*(t,r)h, RB^*(r)U^*(t,r)h\rangle dr \right. \\
&\left. + \int_s^t \left(\int_H \left(e^{i\langle x, B^*(r)U^*(t,r)h\rangle} - 1 - i\langle x, B^*(r)U^*(t,r)h\rangle \chi_{\{\|x\|\le 1\}} \right) \mu(dx) \right) dr \right\}
\end{aligned}
$$

For the jump part we have:

$$
\begin{aligned}
&\int_H \left(e^{i\langle x, B^*(r)U^*(t,r)h\rangle} - 1 - i\langle x, B^*(r)U^*(t,r)h\rangle \chi_{\{\|x\|\le 1\}}\right)\mu(dx) \\
&= \int_H \left(e^{i\langle U(t,r)B(r)x, h\rangle} - 1 - i\langle U(t,r)B(r)x, h\rangle \chi_{\{\|U(t,r)B(r)x\|\le 1\}}\right)\mu(dx) \\
&\quad + \int_H i\langle U(t,r)B(r)x, h\rangle \left(\chi_{\{\|U(t,r)B(r)x\|\le 1\}} - \chi_{\{\|x\|\le 1\}}\right)\mu(dx) \\
&= \int_H \left(e^{i\langle x, h\rangle} - 1 - i\langle x, h\rangle \chi_{\{\|x\|\le 1\}}\right)\mu\circ (U(t,r)B(r))^{-1}(dx) \qquad (4.3) \\
&\quad - \int_H i\langle U(t,r)B(r)x, h\rangle \left(\chi_{\{\|x\|\le 1\}} - \chi_{\{\|U(t,r)B(r)x\|\le 1\}}\right)\mu(dx) \qquad (4.4)
\end{aligned}
$$

Note that (4.3) is finite because of: (setting $C := \|U(t,r)B(r)\|_{\mathcal{L}(H)}$)

$$
\begin{aligned}
&\int_H (1\wedge \|x\|^2)\mu\circ (U(t,r)B(r))^{-1}(dx) = \int_H (1\wedge \|U(t,r)B(r)x\|^2)\mu(dx) \\
&\quad \le \int_H (1\wedge C^2\|x\|^2)\mu(dx) \le \max(1, C^2)\int_H (1\wedge \|x\|^2)\mu(dx) < \infty
\end{aligned}
$$

and only in that way we can argue that (4.4) must be finite as well. Thus, we obtain:

$$
\begin{aligned}
&\exp\left\{\int_s^t \lambda(B^*(r)U^*(t,r)h)dr\right\} \\
&= \exp\left\{i\left\langle \int_s^t U(t,r)B(r)b\,dr, h\right\rangle - \frac{1}{2}\left\langle h, \int_s^t U(t,r)B(r)RB^*(r)U^*(t,r)hdr\right\rangle\right. \\
&\quad + \int_s^t \left(\int_H \left(e^{i\langle x, h\rangle} - 1 - i\langle x, h\rangle\chi_{\{\|x\|\le 1\}}\right)\mu\circ (U(t,r)B(r))^{-1}(dx)\right)dr \\
&\quad \left. - i\left\langle \int_s^t \int_H U(t,r)B(r)x\left(\chi_{\{\|x\|\le 1\}} - \chi_{\{\|U(t,r)B(r)x\|\le 1\}}\right)\mu(dx)dr, h\right\rangle\right\}
\end{aligned}
$$

so that with:

- $b(t,s) := \int_s^t U(t,r)B(r)bdr$
  $- \int_s^t \int_H U(t,r)B(r)x[\chi_{\{\|x\|\le 1\}} - \chi_{\{\|U(t,r)B(r)x\|\le 1\}}]\mu(dx)dr$
- $Q(t,s) := \int_s^t U(t,r)B(r)RB^*(r)U^*(t,r)dr$
- $\mu_{t,s}(A) := \int_s^t \mu\circ (U(t,r)B(r))^{-1}(A)dr$ for $0 \notin A$

$\exp\left\{\int_s^t \lambda(B^*(r)U^*(t,r)h)dr\right\}$ is associated to the triple $[b(t,s), Q(t,s), \mu_{t,s}]$, where $Q(t,s)$ is still symmetric and nonnegative and we have :

$$
\begin{aligned}
&\operatorname{Tr} Q(t,s) = \sum_k \langle e_k, Q(t,s)e_k\rangle = \sum_k \int_s^t \|\sqrt{R}B^*(r)U^*(t,r)e_k\|^2 dr \\
&= \int_s^t \|\sqrt{R}B^*(r)U^*(t,r)\|_2^2 dr \le \int_s^t \|\sqrt{R}\|_2^2 \|B^*(r)U^*(t,r)\|^2 dr < \infty
\end{aligned}
$$

and $\mu_{t,s}$ is a Lévy measure, as we have [since $(1 \wedge \|x\|^2) \leq (\|x\| \wedge \|x\|^2)$]:

$$\begin{aligned}
&\int_s^t \int_H (1 \wedge \|x\|^2)\mu \circ (U(t,r)B(r))^{-1}(dx)dr \\
&\leq \int_s^t \int_H (\|x\| \wedge \|x\|^2)\mu \circ (U(t,r)B(r))^{-1}(dx)dr \\
&= \int_s^t \int_H (\|U(t,r)B(r)x\| \wedge \|U(t,r)B(r)x\|^2)\mu(dx)dr \\
&\leq \int_s^t \|U(t,r)B(r)\|_{\mathcal{L}(H)} \int_H (\|x\| \wedge \|U(t,r)B(r)\|_{\mathcal{L}(H)} \|x\|^2)\mu(dx)dr \\
&\leq \sup_{r\in\mathbb{R}} \|B(r)\|_{\mathcal{L}(H)} \frac{M}{\omega}(1 - e^{-\omega(t-s)}) \underbrace{\int_H (\|x\| \wedge \sup_{r\in\mathbb{R}} \|B(r)\|_{\mathcal{L}(H)} \|x\|^2)\mu(dx)}_{<\infty \text{ by Assumption 4.2.1}} < \infty
\end{aligned}$$

Moreover, we see that we can let $s \to -\infty$ and $Q(t, -\infty)$ will still be trace class as well as $\mu_{t,-\infty}$ will still be a Lévy measure, because of the exponential stability of U. Since we already know that the Fourier transform as a whole converges, convergence of the first part of $b(t, -\infty)$ (which is obvious) implies convergence of the second part. Hence the limit function is associated to a Lévy triple and thus the characteristic function of an infinitely divisible measure. ∎

Remark 4.2.5 *The condition that the Lévy symbol is of linear growth is actually stronger than necessary. To assure the existence of the integral in* (4.2) *it would be even sufficient to have a very weak estimate of the form* $|\lambda(u)| = \mathcal{O}(\sqrt{\|u\|})$*. But we were unable to find any other easy to check conditions to control the growth of a Lévy symbol around the origin. Moreover, to prove that* ν_t *is a charateristic function in the infinite-dimensional case, we really have to use Assumption 4.2.1 not just its implication Lemma 4.2.3.*

4.3 The Subinvariant Measure ν and the Space $L^2(\nu)$

Definition 4.3.1 *Let* $(P_\tau)_{\tau\geq 0}$ *be a semigroup of operators on a Hilbert space* X*. A measure* ν *on* X *is said to be subinvariant for* $(P_\tau)_{\tau\geq 0}$ *if we have*

$$\int_X P_\tau f(x)\nu(dx) \leq C(\tau) \int_X f(x)\nu(dx)$$

for all bounded, measurable and nonnegative functions f *on* X*, for all* $\tau \geq 0$ *and a (locally finite) real-valued function* $C(\tau)$*.*

To be able to use the powerful theory of one-parameter semigroups, we have to reduce our equation to the autonomous case.

Reduction of non-autonomous problems is a well-known method in the theory of ordinary differential equations (see e.g. [13]). We recall that the basic idea is to enlarge the state space, thus allowing to keep track of the elapsed time. The reduced problem then looks:

$$\begin{cases} dX_t = \{A(y(t))X(t) + f(y(t))\}dt + B(y(t))dL_t & X(0) = x \\ dy(t) = dt & y(0) = s \end{cases}$$

For measurable and bounded functions $u : \mathbb{R} \times H \to \mathbb{R}$ the one-parameter semigroup is then defined as follows:

$$P_\tau u(t,x) := P(t, t+\tau)u(t+\tau, \cdot)(x)$$

meaning that we apply the two-parameter semigroup to u as a function of x only. That the family $\{P_\tau\}_{\tau \geq 0}$ is indeed a semigroup, follows, of course, from the semigroup property of $\{P(s,t)\}_{s \leq t}$ and is a simple calculation:

$$\begin{aligned}(P_\sigma(P_\tau u))(t,x) &= P(t, t+\sigma)P(t+\sigma, t+\sigma+\tau)u(t+\sigma+\tau, \cdot)(x) \\ &= P(t, t+\sigma+\tau)u(t+\sigma+\tau, \cdot)(x) = P_{\tau+\sigma}u(t,x)\end{aligned}$$

Starting from our evolution system of measures, we will establish a subinvariant measure for the one-parameter semigroup. On the respective L^2-space the semigroup will then be bounded. To obtain our subinvariant measure we need the following lemma:

Lemma 4.3.2 *The function* $F : (t, A) \mapsto \nu_t(A) \quad t \in \mathbb{R},\ A \in \mathcal{B}(H)$ *is a transition kernel, that is* $t \mapsto \nu_t(A)$ *is measurable for any fixed Borel set* A *and for any fixed* t, ν_t *is a probability measure.*

PROOF By a monotone class argument. Let $\mathcal{M} := \{\exp(i\langle h, \cdot\rangle)\}_{h \in H}$ and $\mathcal{H} := \{f : H \to \mathbb{C} \mid t \mapsto \int_H f(x)\nu_t(dx) \text{ is measurable }\}$. Then $\mathcal{H}$ is a complex monotone vector space and $\mathcal{M}$ is a complex multiplicative system that generates $\mathcal{B}(H)$. $\mathcal{M} \subset \mathcal{H}$ is seen by checking that $\lim_{s \to t} \hat{\nu}_s = \hat{\nu}_t$ pointwisely. ■

Now, we introduce the space on which the semigroup will be strongly continuous and the subspace that will be the core for the generator of our semigroup.

Definition 4.3.3 *As* F *is a transition kernel we can form* $\nu(dt, dx) := \frac{1}{2}\nu_t(dx)e^{-|t|}dt$, *a measure on* $\mathbb{R} \times H$, *defined by*

$$\nu([s,t] \times A) := \frac{1}{2}\int_s^t \nu_t(A)e^{-|t|}dt$$

for $s < t \in \mathbb{R}$ *and* $A \in \mathcal{B}(H)$.

$$L^2(\nu) := \{f : \mathbb{R} \times H \to \mathbb{R} \text{ measurable} \mid \int_{\mathbb{R} \times H} |f(t,x)|^2 \nu(dt, dx) < \infty\}$$

$$\begin{aligned}\mathcal{M} := span_{\mathbb{C}}\{f : \mathbb{R} \times H \to \mathbb{C} \quad | \quad & f = \Phi(t)e^{i\langle U(t+\sigma,t)x, h\rangle}, \text{ where} \\ & \Phi \in C^1(\mathbb{R}, \mathbb{R}) \text{ and bounded}, h \in D(A^*), \sigma \geq 0\}\end{aligned}$$

$$K := \{\Re(f) \mid \quad f \in \mathcal{M}\}$$

That is, K *comprises the real parts of the functions in* $\mathcal{M}$.

Lemma 4.3.4 *K is dense in $L^2(\nu)$.*

Proof We will show density of $\mathcal{M}$ in $L^2(\nu;\mathbb{C})$. This implies density of the respective real vector spaces. We will use complex monotone classes again. The subsystem $\mathcal{M}_0 := \{\Phi_k \otimes \exp_h\}_{\{k\in\mathbb{N}_0, h\in D(A^*)\}}$, where $\Phi_k(t) := \exp(-k|t|)$ and $\exp_h(x) := \exp(i\langle x, h\rangle)$ is closed under multiplication and conjugation. Consider $\mathcal{H} := \bar{\mathcal{M}}_0$ as a subspace of $L^2(\nu;\mathbb{C})$ where we allow complex-valued integrable functions. By monotone convergence, applied separately to real and imaginary parts, $\mathcal{H}$ is seen to be a complex monotone vector space. Thus, $\mathcal{H}$ contains all $\sigma(\mathcal{M}_0)$-measurable functions. If we can show that $\sigma(\mathcal{M}_0) = \mathcal{B}(H\times\mathbb{R})$ we will have all step functions in $\mathcal{H}$, hence density will be obvious. Note that we want to show that functions of the form $\Phi_k \otimes \exp_h$ generate a product σ-algebra $\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(H)$.

Knowing that the Φ_k generate $\mathcal{B}([0,T])$ and that the $\exp_h$ generate $\mathcal{B}(H)$ (which follows again from Lemma 4.1.4), it is clear that $\sigma(\{\Phi_0\otimes\exp_h\}_{h\in D(A^*)}) = (\mathbb{R}\times\mathcal{C})_{\mathcal{C}\in\mathcal{B}(H)}$ and $\sigma(\{\Phi_k\otimes\exp_0\}_{k\in\mathbb{Z}}) = (\mathcal{A}\times H)_{\mathcal{A}\in\mathcal{B}(\mathbb{R})}$ and the result is obvious. ■

Remark 4.3.5 *Note that $\sigma = 0$ would be sufficient to prove density of K, but we will need $\sigma > 0$ later on to show that K is P_τ-invariant.*

Proposition 4.3.6 *The measure ν is a subinvariant measure for the semigroup P_τ.*

Proof Let u be a bounded, measurable and nonnegative function on $\mathbb{R}\times H$.

$$\int_{\mathbb{R}\times H} P_\tau u(t,x)\nu(dt,dx) = \frac{1}{2}\int_{\mathbb{R}}\int_H (P(t,t+\tau)u_{t+\tau})(x)\nu_t(dx)e^{-|t|}dt$$
$$= \frac{1}{2}\int_{\mathbb{R}}\int_H u(t+\tau,x)\nu_{t+\tau}(dx)e^{-|t|}dt = \frac{1}{2}\int_{\mathbb{R}}\int_H u(t,x)\nu_t(dx)e^{-|t-\tau|}dt$$
$$\leq \frac{e^\tau}{2}\int_{\mathbb{R}}\int_H u(t,x)\nu_t(dx)e^{-|t|}dt = e^\tau \int_{\mathbb{R}\times H} u(t,x)\nu(dt,dx)$$

where we used that $e^{-|t-\tau|} \leq e^\tau e^{-|t|}$ for any $\tau > 0$. ■

Proposition 4.3.7 *The semigroup P_τ is bounded on $L^2(\nu)$ with $\|P_\tau\|_{L^2(\nu)} \leq e^{\frac{\tau}{2}}$. Moreover, it is strongly continuous.*

Proof For the first assertion we have to show for a bounded and measurable function u: $\|P_\tau u\|_{L^2}^2 \leq e^\tau\|u\|_{L^2}^2$.
Using the Jensen inequality for the expectation and afterwards the subinvariance property for u^2:

$$\|P_\tau u\|_{L^2(\nu)}^2 = \int_{[0,T]\times H} \mathbb{E}[u(t+\tau, X(t+\tau,t,x))]^2\nu(dt,dx)$$
$$\leq \int_{[0,T]\times H} \mathbb{E}[u^2(t+\tau, X(t+\tau,t,x))]\nu(dt,dx) = \int_{[0,T]\times H} (P_\tau u^2)(t,x)\nu(dt,dx)$$

$$\leq e^{\tau} \int_{[0,T]\times H} u^2(t,x)\nu(dt,dx) = e^{\tau}\|u\|^2_{L^2(\nu)}$$

Hence, P_τ is continuous on a dense subset of $L^2(\nu)$. Thus, we can extend it to all of $L^2(\nu)$. Clearly, the subinvariance property also extends to any $u \in L^2(\nu)$.

The following proof of the strong continuity relies on an analysis of Proposition 4.3 from [33], which is not quite general enough for our needs.

First of all we need to show that $P_\tau g \to g$ as $\tau \to 0$ for any $g \in K$.
For $u(t,x) = \Phi(t)e^{i\langle U(t+\sigma,t)x,h\rangle}$ we have by Lemma 4.1.1:

$$\begin{aligned}(P_\tau u)(t,x) &= \exp\left\{\int_t^{t+\tau} \lambda(B^*(r)U^*(t+\tau+\sigma,r)h)dr\right\} \\ \times \quad &\Phi(t+\tau)\exp\left\{i\left\langle h, U(t+\tau+\sigma,t)x + \int_t^{t+\tau} U(t+\tau+\sigma,r)f(r)dr\right\rangle\right\}\end{aligned} \tag{4.5}$$

Taking $\tau \to 0$ we obtain the result, since all the integrals vanish and by strong continuity of U. Note that, by linearity, this extends to general $u \in K$.

For general functions f in $L^2(\nu)$ we first prove $P_\tau f \to f$ in measure using an approximating function g from K, the Chebychev inequality and the subinvariance property:

$$\begin{aligned}&\nu(\{|P_\tau f - f| > \varepsilon\}) \\ &\leq \nu(\{|P_\tau(f-g)| > \frac{\varepsilon}{3}\}) + \nu(\{|P_\tau g - g| > \frac{\varepsilon}{3}\}) + \nu(\{|f-g| > \frac{\varepsilon}{3}\}) \\ &\leq \frac{9}{\varepsilon^2}\int (P_\tau(f-g))^2 d\nu\} + \nu(\{|P_\tau g - g| > \frac{\varepsilon}{3}\}) + \frac{9}{\varepsilon^2}\int (f-g)^2 d\nu\} \\ &\leq \frac{9}{\varepsilon^2}(1+e^{\tau})\int (f-g)^2 d\nu\} + \nu(\{|P_\tau g - g| > \frac{\varepsilon}{3}\})\end{aligned}$$

The first term goes to zero by choosing g close to f the second by letting $\tau \to 0$.
For L^2-convergence we calculate for $f \geq 0$:

$$\begin{aligned}\int (P_\tau f - f)^2 d\nu &= \int (P_\tau f)^2 d\nu + \int f^2 d\nu - 2\int f P_\tau f d\nu \\ &\leq 2\left(\frac{1+e^{\tau}}{2}\int f^2 d\nu - \int f P_\tau f d\nu\right) \\ &= 2\left(\frac{1+e^{\tau}}{2}\int f^2 d\nu - \int f(P_\tau f - f)^+ d\nu - \int f(P_\tau f \wedge f) d\nu\right) \\ &\leq 2\left(\frac{1+e^{\tau}}{2}\int f^2 d\nu - \int f(P_\tau f \wedge f) d\nu\right)\end{aligned}$$

which turns to zero by dominated convergence. ■

4.4 Generator and Domain of Uniqueness

In this section we show that the generator G of P_τ is uniquely defined on the subspace K of its domain. The precise form we give in Remark 4.4.2 can however not be proven for all functions in K.

Lemma 4.4.1 *K is a core for the infinitesimal generator G of P_τ.*

PROOF Looking closely at (4.5) again, one notes that $(P_\tau u)(t,x)$ is again of the form $\Psi(t)e^{i\langle U(t+\tau+\sigma,t)x,h\rangle}$ with Ψ as follows:

$$\Psi(t) := \Phi(t+\tau)\exp\left\{i\left\langle h, \int_t^{t+\tau} U(t+\tau+\sigma,r)f(r)dr\right\rangle\right\}$$
$$\times \quad \exp\left\{\int_t^{t+\tau} \lambda(B^*(r)U^*(t+\tau+\sigma,r)h)dr\right\}$$

Ψ is indeed C^1: Several elementary calculations show that both integrands are continuous in r and differentiable in t. Differentiation under the integral can be justified by dominated convergence. Hence, K is invariant under P_τ. Furthermore, we have again by (4.5) and some simple computations:

$$Gu = \frac{d}{d\tau}P_\tau u\bigg|_{\tau=0} = \Phi'(t)\,e^{i\langle U(t+\sigma,t)x,h\rangle} + \lambda\left(B^*(t)U^*(t+\sigma,t)h\right)u(t,x) \quad (4.6)$$
$$+\{\, i\langle U(t+\sigma,t)x, A^*(t+\sigma)h\rangle + i\langle U(t+\sigma,t)f(t),h\rangle \,\}\, u(t,x)$$

Thus, we have $K \subset D(G)$ and we can apply a well-known result (see [6] page 47) to prove the assertion.
To give an idea of the computational details omitted here, we calculate:

$$\lim_{\tau\to 0}\frac{1}{\tau}\int_t^{t+\tau} \lambda(B^*(r)U^*(t+\tau+\sigma,r)h)dr =$$
$$\lim_{\tau\to 0}\frac{1}{\tau}\int_t^{t+\tau} [\lambda(B^*(r)U^*(t+\tau+\sigma,r)h) - \lambda(B^*(t)U^*(t+\tau+\sigma,t)h)]dr$$
$$+\lim_{\tau\to 0}\ [\lambda(B^*(t)U^*(t+\tau+\sigma,t)h) - \lambda(B^*(t)U^*(t+\sigma,t)h)]$$
$$+\lambda(B^*(t)U^*(t+\sigma,t)h) = \lambda(B^*(t)U^*(t+\sigma,t)h)$$

where both differences tend to zero because U^* is weakly continuous and λ is continuous with respect to the weak topology (as a characteristic exponent with Gaussian covariance operator of trace class).
Moreover, note that $\frac{d}{dt}\langle U(t,s)x,h\rangle\big|_{t=s} = \lim_{\tau\to 0}\langle\frac{U(s+\tau,s)x-x}{\tau},h\rangle = \langle x, A^*(s)h\rangle$ holds even if $x \notin D(A)$ by using the mean value theorem and strong continuity of U and A to establish the last equality above. ∎

Remark 4.4.2 *One would expect to have a realization of G as a pseudo-differential* *operator, however (with (b, R, M) as in (2.1)):*

$$
\begin{aligned}
Gu(t,x) = {} & u_t(t,x) + \langle f(t), \nabla_x u(t,x)\rangle + \langle x, A^*(t)\nabla_x u(t,x)\rangle \\
& + \langle B(t)b, \nabla_x u(t,x)\rangle + \frac{1}{2} Tr\{\sqrt{R}^* B^*(t)\nabla_{xx} u(t,x) B(t)\sqrt{R}\} \\
& + \int_H \{u(t, x + B(t)y) - u(t,x) - \langle B(t)y, \nabla_x u(t,x)\rangle \chi_{\|y\|\le 1}\}\mu(dy)
\end{aligned}
$$

is true in all of K only under quite restrictive assumptions (e.g. U^* should preserve $D(A^*)$ and we have continuity of $t \mapsto A^*(s_1+t)U^*(s_2+s_3+t, s_3+t)h$ for any fixed $h \in D(A^*)$, $s_1, s_2, s_3 \in \mathbb{R}$. Again, (see Remark 3.2.3) this result also holds if $(t,s) \mapsto U^*(t,s)$ is as regular as $(t,s) \mapsto U(t,s)$). The problem is that otherwise u_t might not exist. Note, however, that in the important case of the square field operator (to be defined later on) no such difficulties occur, since first order terms do not appear by construction.

Chapter 5

Asymptotic Behaviour of the Semigroup

In this chapter we calculate the precise form of the square field operator. Under some strong compatibility assumptions (for which we give a valid example) on the semigroup in connection with the noise we obtain a gradient estimate on the square field operator, which leads in Section 5.2 to a Poincaré and a Harnack inequality.

5.1 The Square Field Operator and an Estimate

In the following we will introduce the square field operator. Its importance lies in the crucial role that it will play in the proof of the following functional inequalities.

Definition 5.1.1 $\Gamma(u,v) := G(uv) - uGv - vGu$ *will be called the square field operator.*

Lemma 5.1.2 (square field operator) *On K we have:*

$$\Gamma(u,u) = \left\langle \sqrt{R}^* B^*(t)\nabla_x u, \sqrt{R}^* B^*(t)\nabla_x u \right\rangle + \int_H \left[u(t,x+B(t)y) - u(t,x)\right]^2 \mu(dy)$$

where μ is as in (2.1).

Proof As Γ is not linear we have to establish the formula not only for functions $u_i(t,x) = \Phi_i(t) e^{i\langle U(t+\sigma_i,t)x, h_i\rangle}$ but also for sums of such functions. Since $\Gamma(u+v, u+v) = \Gamma(u,u) + \Gamma(v,v) + 2\Gamma(u,v)$ the essential part of the proof is to show the following claim:

$$\begin{aligned}\Gamma(u_1,u_2) = &\left\langle \sqrt{R}^* B^*(t)\nabla_x u_1, \sqrt{R}^* B^*(t)\nabla_x u_2 \right\rangle \\ &+ \int_H \left[u_1(t,x+B(t)y) - u_1(t,x)\right]\left[u_2(t,x+B(t)y) - u_2(t,x)\right] \mu(dy)\end{aligned}$$

Though u_1u_2 may not be in K, Lemma 4.1.1 again allows to derive $P_\tau u_1 u_2$, thus we can compute $G(u_1u_2)$ explicitly as:

$$\begin{aligned} G(u_1u_2) =& [\Phi_1'(t)\Phi_2(t) + \Phi_1(t)\Phi_2'(t)]e^{i\langle x, U^*(t+\sigma_1,t)h_1 + U^*(t+\sigma_2,t)h_2\rangle} \\ &+ i\langle x, U^*(t+\sigma_1,t)A^*(t+\sigma_1)h_1 + U^*(t+\sigma_2,t)A^*(t+\sigma_2)h_2\rangle u_1u_2 \\ &+ i\langle f(t), U^*(t+\sigma_1,t)h_1 + U^*(t+\sigma_2,t)h_2\rangle u_1u_2 \\ &+ \lambda[B^*(t)U^*(t+\sigma_1,t)h_1 + U^*(t+\sigma_2,t)h_2]u_1u_2 \end{aligned}$$

As $u_1G(u_2)$ and $u_2G(u_1)$ are likewise multiples of $e^{i\langle x, U^*(t+\sigma_1,t)h_1 + U^*(t+\sigma_2,t)h_2\rangle}$ most terms in $\Gamma(u_1,u_2)$ cancel out. In fact we are left with:

$$\begin{aligned} \Gamma(u_1,u_2) &= \lambda[B^*(t)U^*(t+\sigma_1,t)h_1 + B^*(t)U^*(t+\sigma_2,t)h_2]u_1u_2 \\ &- \lambda[B^*(t)U^*(t+\sigma_1,t)h_1]u_1u_2 - \lambda[B^*(t)U^*(t+\sigma_2,t)h_2]u_1u_2 \\ &= \left\langle \sqrt{R}^* B^*(t)U^*(t+\sigma_1,t)h_1u_1, \sqrt{R}^* B^*(t)U^*(t+\sigma_2,t)h_2u_2 \right\rangle \\ &+ \int_H \Big(u_1(t,x+B(t)y)u_2(t,x+B(t)y) + u_1(t,x)u_2(t,x) \\ &\qquad - u_1(t,x+B(t)y)u_2(t,x) - u_2(t,x+B(t)y)u_1(t,x) \Big) \mu(dy) \end{aligned}$$

and from this the claim easily follows. ∎

Assumption 5.1.3 *In addition to Assumptions 3.1.2 and 4.2.1 we will need the following conditions to hold for the rest of the paper:*

(i) For every $t > s$: $U(t,s)RH \subset \sqrt{R}H$ and there is a strictly positive function $C_1 \in C[0,\infty)$ such that, denoting by $\sqrt{R}^{-1}x$ the solution z with minimal norm of $\sqrt{R}z = x$:

$$\|\sqrt{R}^{-1}U(t,s)Rx\| \le C_1(t-s)\|\sqrt{R}x\| \quad x \in H,\ t > s$$

(ii) There is a strictly positive function $C_2 \in C[0,\infty)$ such that:

$$\mu \circ U(t,s)^{-1} \le C_2(t-s)\mu \quad t > s$$

that is $C_2(t-s)\mu - \mu \circ U(t,s)^{-1}$ is a positive measure.

Example 5.1.4 *Assumption 5.1.3 (i) is easily seen to be fulfilled in the case where H is finite dimensional and R has full rank. Let us now give an infinite-dimensional example. Let H be a real separable Hilbert space and let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal basis of H. For $x \in D(A) \subset H$, $t \in \mathbb{R}$ let*

$$A(t)x := \sum_{n\in\mathbb{N}} \lambda_n(t)\langle x, e_n\rangle e_n$$

where $D(A) := \{x \in H : \sum_{n\in\mathbb{N}} n^4\ |\langle x, e_n\rangle|^2 < \infty\}$ *is the common and dense domain of the operators* $A(t)$ *and each* λ_n *is a continuous,* T*-periodic, real-valued function satisfying* $-\frac{n^2}{\Lambda} \leq \lambda_n(t) \leq -\Lambda n^2$ *for some* $0 < \Lambda < 1$. *For* $x \in H$, $t \geq s \in \mathbb{R}$ *let*

$$U(t,s)x := \sum_{n\in\mathbb{N}} \exp\left(\int_s^t \lambda_n(r)dr\right)\langle x, e_n\rangle e_n.$$

Then it is easy to check that $U(t,s)_{t\geq s}$ *define an exponentially bounded evolution family as in Definition 3.1.1, e.g. we have* $\|U(t,s)x\| \leq \exp(-\Lambda(t-s))\|x\|$. *Moreover one can verify easily that Assumption 3.1.2 is fulfilled. For* $x \in H$ *let* $Rx := \sum_{n\in\mathbb{N}} \frac{1}{n^2}\langle x, e_n\rangle e_n$. *Then* R *is a positive, symmetric operator of trace class and we have* $\sqrt{R}^{-1}z = \sum_{n\in\mathbb{N}} n\langle z, e_n\rangle e_n$ *for every* z *in the image of* $\sqrt{R}$ *i.e.* $\sum_{n\in\mathbb{N}} n^2|\langle z, e_n\rangle e_n|^2 < \infty$. *Thus, we obtain:*

$$\left\|\sqrt{R}^{-1}U(t,s)Rx\right\|^2 = \left\|\sum_{n\in\mathbb{N}} n\exp\left(\int_s^t \lambda_n(r)dr\right)\frac{1}{n^2}\langle x, e_n\rangle e_n\right\|^2$$
$$\leq \exp(-2\Lambda(t-s))\sum_{n\in\mathbb{N}}\frac{1}{n^2}|\langle x, e_n\rangle|^2 = \exp(-2\Lambda(t-s))\left\|\sqrt{R}x\right\|^2$$

and Assumption 5.1.3 (i) is fulfilled with $C_1(t) = \exp(-\Lambda t)$.

Now we show, that in this setting, Assumption 5.1.3 (ii) will be fulfilled for finite-dimensional α*-stable Lévy noise. Let* $H_N := span(e_1, ..., e_N) \setminus \{0\}$. *For* $A \in \mathcal{B}(H_N)$ *define the measure* μ_N *as* $\mu_N(A) = \int_A \|x\|^{-N-\alpha}d^Nx$, *where* d^Nx *is N-dimensional Lebesgue measure and* $1 < \alpha < 2$. *It is well known that* $\mathcal{B}(H_N) = \mathcal{B}(H\setminus\{0\}) \cap H_N$ *so we can extend the measure* μ_N *by zero to a measure* $\tilde{\mu_N}$ *on all of* $H \setminus \{0\}$. *It is easy to see that* μ_N *(and thus* $\tilde{\mu_N}$*) is indeed a Lévy measure that also fulfills Assumption 4.2.1. Since for every* $t \geq s$ $U(t,s)$ *leaves* H_N *invariant, it will be sufficient to check 5.1.3 (ii) on measurable subsets of* H_N. *First let* O *be an open subset in* H_N. *Then we have by the transformation rule:*

$$\int\limits_{U(t,s)^{-1}(O)} \|x\|^{-(N+\alpha)}d^Nx = \int\limits_O \|U(t,s)^{-1}x\|^{-(N+\alpha)}\underbrace{\exp\left(-\int\limits_s^t\sum_{n=1}^N \lambda_n(r)dr\right)}_{\det U(t,s)^{-1}_{|H_N}} d^Nx$$

$$\leq \exp\{(t-s)[\Lambda^{-1}N^3 - (N+\alpha)\Lambda]\}\int\limits_O \|x\|^{-(N+\alpha)}d^Nx$$

since we have $\|U(t,s)^{-1}x\| \geq e^{\Lambda(t-s)}\|x\|$ *and* $-\lambda_1(r) \leq ... \leq -\lambda_N(r) \leq \Lambda^{-1}N^2$. *Moreover, as* μ_N *is finite on compacts it is outer regular and we obtain the inequality on all of* $\mathcal{B}(H_N)$. *Thus, Assumption 5.1.3 (ii) holds with* $C_2(t) = \exp(\,t\,[\Lambda^{-1}N^3 - (N+\alpha)\Lambda])$.

For a concrete realization of this abstract example let $H := L^2(0,\pi)$ *and let* $(e_n)_{n\in\mathbb{N}}$ *be the eigenfunctions of the Dirichlet Laplacian* Δ. *For a* T*-periodic, continuous and strictly positive function* f *let* $\lambda_n(t) := -f(t)n^2$. *Then we have* $A(t) = f(t)\Delta$, *since it can be verified that* $D(A) = H^2 \cap H_0^1$. *In turn,* R *is given by* $(-\Delta)^{-1}$.

Remark 5.1.5 *Note that under Assumption 5.1.3 we can only prove Proposition 5.2.4. For Corollary 6.2.3 the function* C_2 *needs to be integrable over* $\mathbb{R}_+$. *In our example this is only true if we restrict* $A(t)$ *to be a time-dependent multiple of the identity. That is* $\lambda_n(t) := -f(t)$ *with* f *as above results in* $C_2(t) = \exp\left(-t\alpha \inf_{s\in[0,T]} f(s)\right)$ *which is clearly integrable.*

Lemma 5.1.6 (estimate of the square field operator) *If* $B = Id$, *we have for* $u \in K$:

$$\|\sqrt{R}\nabla_x P_\tau u\| \le C_1(\tau)\, P_\tau\left(\|\sqrt{R}\nabla_x u\|\right)(t,x) \tag{5.1}$$

$$\int_H [P_\tau u(t,x+y) - P_\tau u(t,x)]^2\,\mu(dy) \le C_2(\tau) P_\tau\left(\int_H [u(\cdot+y)-u(\cdot)]^2\mu(dy)\right)(t,x) \tag{5.2}$$

So that combining the two estimates, we have:

$$\Gamma(P_\tau u, P_\tau u) \le \max(C_1, C_2)(\tau) P_\tau \Gamma(u,u)$$

PROOF Recall that $P_\tau u(t,x) = \mathbb{E}[u(t+\tau, X(t+\tau,t,x))]$. Let $z \in H$ and $u(t,x) = \Phi(t)e^{i\langle U(t+\sigma,t)x,h\rangle}$, then (using Lemma 4.1.1 in the first equality):

$$\begin{aligned}
\langle \nabla_x P_\tau u(t,x), Rz\rangle &= \langle iU^*(t+\tau+\sigma,t)h\, P_\tau u(t,x), Rz\rangle \\
&= \int_H \langle iU^*(t+\tau+\sigma,t)h\, u(t+\tau,y), Rz\rangle\, P\circ X(t+\tau,t,x)^{-1}(dy) \\
&= P_\tau\left(\langle \nabla_x u, \sqrt{R}\sqrt{R}^{-1} U(\cdot,\cdot-\tau)Rz\rangle\right)(t,x) \qquad (+) \\
&\le P_\tau\left(\|\sqrt{R}\nabla_x u\|\, \|\sqrt{R}^{-1}U(\cdot,\cdot-\tau)Rz\|\right)(t,x) \\
&\le P_\tau\left(\|\sqrt{R}\nabla_x u\|\right)(t,x)\, C_1(\tau)\, \|\sqrt{R}z\|
\end{aligned}$$

Now, for every pair (t,x) choosing $z = \nabla_x P_\tau u(t,x)$ we obtain:

$$\langle \nabla_x P_\tau u(t,x), R\nabla_x P_\tau u(t,x)\rangle \le \sqrt{C_1(\tau)}\, \|\sqrt{R}\nabla_x P_\tau u(t,x)\|\;\; P_\tau\left(\|\sqrt{R}\nabla_x u\|\right)(t,x)$$

which is equivalent to (5.1).
Note that we have used the special form of u only up to equation $(+)$, but by

linearity of P and ∇_x it is clear that this also holds for sums. So we obtain (5.1) on all of K.

Recall for the proof of (5.2), that we can write $P_\tau u(t,\xi)$ as $P_\tau u(t,\xi) = \int_H u(t+\tau, U(t+\tau,t)\xi+\eta)\ P\circ X(t+\tau,t,0)^{-1}(d\eta)$.
Thus, setting $\tilde{P} := P\circ X(t+\tau,t,0)^{-1}$ and $\tilde{\mu} := \mu\circ U(t+\tau,t)^{-1}$ we have for general $u\in K$: (setting $\tilde{\tau} := t+\tau$ for brevity)

$$\begin{aligned}
&\int_H |P_\tau u(t,x+y) - P_\tau u(t,x)|^2\,\mu(dy)\\
\le&\int_H\left(\int_H |u(\tilde{\tau},U(\tilde{\tau},t)(x+y)+z) - u(\tilde{\tau},U(\tilde{\tau},t)x+z)|^2\tilde{P}(dz)\right)\mu(dy)\\
=&\int_H\left(\int_H |u(\tilde{\tau},U(\tilde{\tau},t)x+y+z) - u(\tilde{\tau},U(\tilde{\tau},t)x+z)|^2\tilde{\mu}(dy)\right)\tilde{P}(dz)\\
\le&C_2(\tau)\int_H\left(\int_H |u(\tilde{\tau},U(\tilde{\tau},t)x+y+z) - u(\tilde{\tau},U(\tilde{\tau},t)x+z)|^2\mu(dy)\right)\tilde{P}(dz)\\
=&C_2(\tau)P_\tau\left(\int_H |u(\cdot+y)-u(\cdot)|^2\mu(dy)\right)(t,x)
\end{aligned}$$
■

Corollary 5.1.7

$$\sqrt{\langle\nabla_x P_\tau u(t,x), \nabla_x P_\tau u(t,x)\rangle} \le \|U(t+\tau,t)\|P_\tau\left(\|\nabla_x u\|\right)(t,x) \tag{5.3}$$

PROOF Reconsidering the proof above and setting $R = Id$ yields the result. ■

5.2 Functional Inequalities

Following [42], we will now prove a Poincaré and a Harnack inequality.

Definition 5.2.1

$$\overline{u}_t := \int_H u(t,x)\nu_t(dx), \quad u\in L^2_*(\nu)$$

Lemma 5.2.2 *The members of the evolution family of measures $(\nu_t)_{t\in\mathbb{R}}$ have a uniformly bounded first moment:* $\sup_{t\in[0,T]}\left\{\int_H \|x\|\nu_t(dx)\right\} < \infty$.

PROOF Let us write $[b_t, R_t, \mu_t]$ instead of $[b_{t,-\infty}, R_{t,-\infty}, \mu_{t,-\infty}]$ for the Lévy triple associated to ν_t. By the Lévy-Ito-decomposition and Theorem 3.25 in [2] we obtain (using also Jensen's inequality) :

$$\int_H \|x\|\nu_t(dx) \le \|b_t\| + (\mathrm{Tr}(R_t))^{\frac{1}{2}} + \left(\int_{\|x\|\le 1}\|x\|^2\mu_t(dx)\right)^{\frac{1}{2}} + \int_{\|x\|>1}\|x\|\mu_t(dx)$$

We have to estimate this expression uniformly in t. According to the last chain of inequalities in the proof of Theorem 4.2.4 we get for the last two terms:

$$\sup_{t\in\mathbb{R}}\left\{\int_{\|x\|\le 1}\|x\|^2\mu_t(dx)+\int_{\|x\|>1}\|x\|\mu_t(dx)\right\}=\sup_{t\in\mathbb{R}}\int_H(\|x\|\wedge\|x\|^2)\mu_t(dx)<\infty.$$

To estimate $\|b_t\|$ let us for simplicity assume that $B=Id$, hence:

$$\begin{aligned}\|b_t\| &= \left\|\int_{-\infty}^t U(t,r)b\,dr-\int_{-\infty}^t\int_H U(t,r)x[\chi_{\{\|x\|\le 1\}}-\chi_{\{\|U(t,r)x\|\le 1\}}]\mu(dx)dr\right\| \\ &\le \int_{-\infty}^t Me^{-\omega(t-r)}dr\|b\|+\int_{-\infty}^t Me^{-\omega(t-r)}dr\int_H\|x\|\chi_{\{\|x\|\ge M^{-1}\}}\mu(dx)\end{aligned}$$

The last expression is finite by Assumption 4.2.1 and obviously independent of t.
Finally, as in the proof of Theorem 4.2.4 we have:

$$(\mathrm{Tr}\,R_t)^{\frac12}\le\sup_r\|R(r)\|_{\mathcal{L}(H)}\int_{-\infty}^t\|U(t,r)\|_{\mathcal{L}(H)}dr\,(\mathrm{Tr}\,R)^{\frac12}$$

and using once more the exponential stability of U this bound is seen to be independent of t, as well. ■

Proposition 5.2.3 *We have for all $u\in K$:*

$$\lim_{\tau\to\infty}\left(\sup_t|P_\tau u(t,x)-\overline{u}_{t+\tau}|\right)=0\quad \textit{for every fixed } x$$

PROOF We have, since $\overline{u}_{t+\tau}:=\int_H u(t+\tau,y)\nu_{t+\tau}(dy)=\int_H P_{t,t+\tau}u(t+\tau,\cdot)(y)\nu_t(dy)$ by the property of the evolution system :

$$\begin{aligned}&|P_\tau u(t,x)-\overline{u}_{t+\tau}|=\left|\int_H[P_{t,t+\tau}u(t+\tau,\cdot)(x)-P_{t,t+\tau}u(t+\tau,\cdot)(y)]\nu_t(dy)\right|\\ &\le\|\nabla_x P_{t,t+\tau}u(t+\tau,\cdot)\|_\infty\int_H\|x-y\|\nu_t(dy)\\ &\le Me^{-\omega\tau}\|\nabla_x u(t+\tau,\cdot)\|_\infty\int_H\|x-y\|\nu_t(dy)\overset{\tau\to\infty}{\longrightarrow}0\end{aligned}$$

since the integral is bounded by Lemma 5.2.2 and $\|\nabla_x u(t,x)\|$ is easily seen to be bounded as well. ■

Proposition 5.2.4 (Poincaré Inequality) *Given Assumption 5.1.3 and $B=Id$, we have for $C(\tau):=\max\left(\int_0^\tau C_1(s)ds,\int_0^\tau C_2(s)ds\right)$:*

$$P_\tau u^2-(P_\tau u)^2\le C(\tau)P_\tau\Gamma(u,u)\qquad \textit{for all}\quad \tau>0,u\in K \tag{5.4}$$

PROOF Set $f(s) := P_{\tau-s}(P_s u)^2$. Then we have by the product rule:

$$\begin{aligned}\frac{d}{ds}f(s) &= -P_{\tau-s}G(P_s u)^2 + P_{\tau-s}2P_s uGP_s u\\ &= -P_{\tau-s}[G(P_s u)^2 - 2P_s uGP_s u] = -P_{\tau-s}\Gamma(P_s u, P_s u)\end{aligned}$$

Hence, using Lemma 5.1.2 and Lemma 5.1.6 :

$$\begin{aligned}-\frac{d}{ds}f(s) &= P_{\tau-s}\Gamma(P_s u, P_s u)\\ &= P_{\tau-s}\left\langle \sqrt{R}^*\nabla_x P_s u, \sqrt{R}^*\nabla_x P_s u\right\rangle\\ &\quad + P_{\tau-s}\int_H [P_s u(t,x+y) - P_s u(t,x)]^2\,\mu(dy)\\ &\le C_1(s)P_{\tau-s}P_s\left\langle \sqrt{R}^*\nabla_x u, \sqrt{R}^*\nabla_x u\right\rangle \qquad \text{by (5.1)}\\ &\quad + C_2(s)P_{\tau-s}P_s\int_H [u(t,\cdot+y) - u(t,\cdot)]^2\,\mu(dy) \qquad \text{by (5.2)}\end{aligned}$$

Integrating with respect to s and noting that $f(0) = P_\tau u^2$ and $f(\tau) = (P_\tau u)^2$ we obtain:

$$\begin{aligned}P_\tau u^2 - (P_\tau u)^2 \le \left(\int_0^\tau C_1(s)ds\right) P_\tau\left\langle \sqrt{R}^*\nabla_x u, \sqrt{R}^*\nabla_x u\right\rangle\\ + \left(\int_0^\tau C_2(s)ds\right) P_\tau\int_H [u(t,\cdot+y) - u(t,\cdot)]^2\,\mu(dy)\end{aligned}$$

and the result is proved. ∎

For the following Harnack inequality we need a definition.

Definition 5.2.5

$$\rho(x,y) := \inf\{\ \|z\|\ :\ \sqrt{R}z = x - y\}$$

with the usual convention that $\inf\emptyset = \infty$ *, so* ρ *may take the value infinity if* $(x-y) \notin Im\sqrt{R}$.

ρ just describes the distance induced by the Cameron-Martin norm. Heuristically this means we can only compare $P_\tau u$ and $P_\tau u^2$ at points x and y which are connected via the (Gaussian) noise. In particular, if he have no Gaussian part, the Harnack inequality gives no information.

For much more information on Harnack inequalities (including other ways of proving them) see the thesis [37] which is entirely dedicated to this topic.

Proposition 5.2.6 (Harnack Inequality) *Let* $C(t) := C_1(t)$ *be the strictly positive function introduced in Assumption 5.1.3 (i). Then we have:*

$$|P_\tau u(t,y)|^2 \le P_\tau u^2(t,x)\exp\left(\frac{\rho^2(x,y)}{\int_0^\tau \frac{1}{C(s)}ds}\right) \quad \text{for all } u \in C_b(H),\ x,y \in H \tag{5.5}$$

PROOF First, let $u \in K$ be such that u is strictly positive. Since $P_{\tau-s}(P_s u)^2(t,x)$ will then also be strictly positive we can define:

$$\Phi(s) := \log[P_{\tau-s}(P_s u)^2(t,x_s)]$$

where x_s is given by $x_s := x + \frac{(y-x)\int_0^s \frac{1}{C(\tau-u)}du}{\int_0^\tau \frac{1}{C(u)}du}$. Differentiating Φ we obtain:

$$\frac{d}{ds}\Phi(s) = \frac{\frac{d}{ds}P_{\tau-s}(P_s u)^2(t,x_s)}{P_{\tau-s}(P_s u)^2(t,x_s)}. \tag{5.6}$$

Using that $(P_s u)^2 \in K \subset D(G)$ we have for the numerator:

$$\begin{aligned}\frac{d}{ds}[P_{\tau-s}(P_s u)^2(t,x_s)] &= \frac{d}{ds}[P_{\tau-s}(P_s u)^2](t,x_s) + \left\langle \nabla_x[P_{\tau-s}(P_s u)^2](t,x_s), \frac{dx_s}{ds}\right\rangle \\ &= -GP_{\tau-s}(P_s u)^2(t,x_s) + P_{\tau-s}[2P_s u G P_s u](t,x_s) \\ &\quad + \frac{1}{C(\tau-s)\int_0^\tau \frac{1}{C(u)}du}\left\langle \nabla_x[P_{\tau-s}(P_s u)^2](t,x_s),(y-x)\right\rangle \\ &= -P_{\tau-s}\Gamma(P_s u, P_s u) \\ &\quad + \frac{1}{C(\tau-s)\int_0^\tau \frac{1}{C(u)}du}\left\langle \nabla_x[P_{\tau-s}(P_s u)^2](t,x_s),(y-x)\right\rangle\end{aligned} \tag{5.7}$$

We will now estimate $\left\langle \nabla_x[P_{\tau-s}(P_s u)^2](t,x_s),(y-x)\right\rangle$:

$$\begin{aligned}&\left\langle \nabla_x[P_{\tau-s}(P_s u)^2](t,x_s),(y-x)\right\rangle \\ &= \inf_{\{z:\sqrt{R}z=x-y\}}\left\langle \nabla_x[P_{\tau-s}(P_s u)^2](t,x_s),\sqrt{R}z\right\rangle && (x-y)\in \mathrm{Im}\sqrt{R} \\ &\le \sqrt{\langle R\nabla_x[P_{\tau-s}(P_s u)^2](t,x_s), \nabla_x[P_{\tau-s}(P_s u)^2](t,x_s)\rangle}\,\rho(x,y) && \text{Cau.-Schw.} \\ &\le \rho(x,y)\sqrt{C(\tau-s)}P_{\tau-s}\left(\sqrt{\langle R\nabla_x(P_s u)^2, \nabla_x(P_s u)^2\rangle}\right)(t,x_s) && \text{by (5.1)} \\ &\le 2\rho(x,y)\sqrt{C(\tau-s)}P_{\tau-s}\left(P_s u\sqrt{\langle R\nabla_x(P_s u), \nabla_x(P_s u)\rangle}\right)(t,x_s) && \text{chain rule}\end{aligned} \tag{5.8}$$

Combining (5.6),(5.7) and (5.8) and using Lemma 5.1.2 we obtain:

$$
\begin{aligned}
&\frac{d}{ds}\Phi(s) \leq \quad \frac{-P_{\tau-s}\Gamma(P_s u, P_s u)}{P_{\tau-s}(P_s u)^2(t,x_s)} \\
&+ \frac{\frac{1}{C(\tau-s)\int_0^\tau \frac{1}{C(u)}du} 2\rho(x,y)\sqrt{C(\tau-s)}P_{\tau-s}\left(P_s u\sqrt{\langle R\nabla_x(P_s u), \nabla_x(P_s u)\rangle}\right)(t,x_s)}{P_{\tau-s}(P_s u)^2(t,x_s)} \\
&\leq \frac{-P_{\tau-s}\left(\langle R\nabla_x(P_s u), \nabla_x(P_s u)\rangle\right)(t,x_s)}{P_{\tau-s}(P_s u)^2(t,x_s)} \\
&+ \frac{\frac{1}{\sqrt{C(\tau-s)}\int_0^\tau \frac{1}{C(u)}du} 2\rho(x,y)P_{\tau-s}\left(P_s u\sqrt{\langle R\nabla_x(P_s u), \nabla_x(P_s u)\rangle}\right)(t,x_s)}{P_{\tau-s}(P_s u)^2(t,x_s)} \\
&= \frac{1}{P_{\tau-s}(P_s u)^2(t,x_s)} \\
&\times P_{\tau-s}\left((P_s u)^2\left[2H\frac{\sqrt{\langle R\nabla_x(P_s u), \nabla_x(P_s u)\rangle}}{P_s u} - \frac{\langle R\nabla_x(P_s u), \nabla_x(P_s u)\rangle}{(P_s u)^2}\right]\right)(t,x_s)
\end{aligned}
$$

where we have set $H := \frac{\rho(x,y)}{\sqrt{C(\tau-s)}\int_0^\tau \frac{1}{C(u)}du}$ for brevity.

Furthermore, setting $G := \frac{\sqrt{\langle R\nabla_x(P_s u), \nabla_x(P_s u)\rangle}}{P_s u}$:

$$
\begin{aligned}
\frac{d}{ds}\Phi(s) &\leq \frac{1}{P_{\tau-s}(P_s u)^2(t,x_s)} P_{\tau-s}\left((P_s u)^2\left[-G^2 + 2HG\right]\right)(t,x_s) \\
&= \frac{1}{P_{\tau-s}(P_s u)^2(t,x_s)} P_{\tau-s}\left((P_s u)^2\left[-G^2 + 2HG - H^2 + H^2\right]\right)(t,x_s) \\
&\leq \frac{1}{P_{\tau-s}(P_s u)^2(t,x_s)} P_{\tau-s}\left((P_s u)^2\left[H^2\right]\right)(t,x_s) \\
&= H^2
\end{aligned}
$$

since H depends neither on x_s nor on t. Integration over s yields:

$$
\begin{aligned}
&\log[(P_\tau u)^2(t,y)] - \log[(P_\tau u^2)(t,x)] = \Phi(\tau) - \Phi(0) \\
&\leq \int_0^\tau H^2(s)ds = \int_0^\tau \frac{\rho^2(x,y)}{C(\tau-s)(\int_0^\tau \frac{1}{C(u)}du)^2}ds = \frac{\rho^2(x,y)}{\int_0^\tau \frac{1}{C(u)}du}
\end{aligned}
$$

Hence, applying the exponential yields: $(P_\tau u)^2(t,y) \leq P_\tau u^2(t,x)\ \exp\left(\frac{\rho^2(x,y)}{\int_0^\tau \frac{1}{C(u)}du}\right)$ and the proof is complete for positive functions. To obtain the result for general u, note first, that it is sufficient to have it for $|u|$, since then we get: $|P_\tau u(t,y)|^2 \leq [P_\tau |u|(t,y)]^2 \leq P_\tau u^2(t,x)\exp\left[\rho^2(x,y)(\int_0^\tau \frac{1}{C(s)}ds)^{-1}\right]$
Of course, we cannot take modulus without leaving K, but we may take the

square of our functions. Thus, let $u \in C_b(H)$ and $\varepsilon > 0$. Then $f := \sqrt{|u|} \in C_b(H)$. Now, by Lemma 5.2.8 we can approximate f pointwisely by functions f_n from K. Then $f_n^2 + \varepsilon$ is strictly positive, it will approach $|u| + \varepsilon$ and since the approximating functions are uniformly bounded, we can take limits in (5.5) and obtain the result via dominated convergence and then letting $\varepsilon \to 0$. ∎

The following consequence of the Harnack inequality concerning smoothing properties of the two-parameter semigroup is the time-dependent generalization of Proposition 4.1 from [18]. Since we do not have full Gaussian White Noise, we can expect continuity only along the directions of the Cameron-Martin space.

Corollary 5.2.7 *For any $s < t$ the operator $P(s,t)$ is strong Feller, in the following sense: Let $H_0 := \sqrt{R}H$ be the Cameron-Martin space equipped with the norm $\|x\|_{H_0} := \rho(x,0)$. Then, for any y in a set of full μ_s-measure we have $P(s,t)f \in C_b(y+H_0)$ for any function $f \in L^p(H;\mu_t)$, where μ_t is the member of the evolution system of measures established in Theorem 4.2.4. Here $C_b(y+H_0)$ is meant to be the continuous (and bounded) functions with respect to the norm $\|x\|_{H_0}$.*

Proof By the definition of an evolution system of measures, we have for bounded and measurable functions f:

$$\int_H P(s,t)f(x)\mu_s(dx) = \int_H f(x)\mu_t(dx). \tag{*}$$

Thus, by Jensen's inequality we obtain

$$\int_H |P(s,t)f(x)|^p \mu_s(dx) \leq \int_H P(s,t)|f|^p(x)\mu_s(dx) = \int_H |f|^p(x)\mu_t(dx)$$

and, hence, the operator $P(s,t)$ extends to a contraction from $L^p(\mu_t)$ to $L^p(\mu_s)$. The equality (*) thus extends to functions f from $L^p(\mu_t)$ as well.
Now for some function $f \in L^p(\mu_t)$ let f_n be an approximation of f in $L^p(\mu_t)$ by bounded and continuous functions. We have using (*):

$$\lim_{m,n\to\infty} \int_H P(s,t)|f_m - f_n|^p(x)\mu_s(dx) = 0$$

In particular we have $\lim_{m,n\to\infty} P(s,t)|f_m - f_n|^p = 0$ μ_s- a.e. along a subsequence. Hence, there is $y \in H$ such that $\lim_{m,n\to\infty} P(s,t)|f_m - f_n|^p(y) = 0$. We will now show that the sequence $P(s,t)f_n$ is even Cauchy in some uniform norm. By the Feller property of $P(s,t)$ (proven in Lemma 4.1.6) , it is clear that $P(s,t)f_n \in C_b(H)$. In particular, we have $P(s,t)f_n \in C_b(H_0)$ since the corresponding norm is stronger. Moreover, spaces of continuous functions are complete under the uniform norm. Thus, let $B_N := \{x \in H \mid \frac{\rho^2(x,y)}{\int_0^{t-s} \frac{1}{C(r)}dr} \leq N\}$. Then, by using the Harnack inequality (5.5) for the time-independent function $u(t,x) := (f_n - f_m)(x)$, we obtain

$$\begin{aligned}\sup_{x\in B_N} |P(s,t)f_n(x) - P(s,t)f_m(x)|^p &\leq \sup_{x\in B_N} (P(s,t)|f_n - f_m|(x))^p \\ &\leq P(s,t)|f_n - f_m|^p(y)e^{pN}.\end{aligned}$$

Hence, $P(s,t)f$ will be bounded and continuous (in the norm $\|\cdot\|_{H_0}$) for every $x \in B_N$ and letting $N \to \infty$ we retain continuity of f everywhere in $y + H_0$. ∎

The following Lemma was needed in the proof of the harnack inequality 5.2.6.

Lemma 5.2.8 *For every $f \in C_b(H)$ we can find a sequence $f_n \in K$ such that:*

- *$f_n \to f$ pointwisely*
- $\sup_{x\in H, t\in\mathbb{R}, n\in\mathbb{N}} |f_n(t,x)| \le 1 + \sup_{x\in H, t\in\mathbb{R}} |f(t,x)|$

PROOF Since $D(A^*)$ is dense in H we can find an ONB $(e_n)_{n\in\mathbb{N}}$ of H such that every e_n is in $D(A^*)$.

Let $P_n : \mathbb{R} \times H \to \mathbb{R} \times \text{span}(e_1, ..., e_n)$, $(t,x) \mapsto (t, \sum_{i=1}^n \langle e_i, x\rangle e_i)$.
For technical purposes we need to consider $f \circ P_n$ as a function on $\mathbb{R}\times\mathbb{R}^n$, hence let $g_n : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}$ be such that $f \circ P_n(t,x) = g_n(t, \langle x, e_1\rangle, ..., \langle x, e_n\rangle)$.
It is clear that g_n is also continuous. Hence, by the Theorem of Stone-Weierstrass, we can approximate g_n on $[-n,n]\times[-n,n]^n$ uniformly by linear combinations of functions $F^n_{\Phi,\alpha}(t,x) := \Phi(t)\exp(i\frac{\Pi}{n}\langle\alpha, x\rangle_{\mathbb{R}^n})$ where Φ is differentiable and $2n$-periodic and $\alpha \in \mathbb{Z}^n$. Let us call $\tilde{g_n}$ such an approximation of g_n with

$$\sup_{t\in[-n,n], x\in[-n,n]^n} |g_n(t,x) - \tilde{g_n}(t,x)| < \frac{1}{n}.$$

Let $f_n(t,x) := \tilde{g_n}(t, \langle x, e_1\rangle, ..., \langle x, e_n\rangle)$. Then we have

$$|f(t,x) - f_n(t,x)| \le |f(t,x) - f\circ P_n(t,x)| + |f\circ P_n(t,x) - f_n(t,x)|.$$

For $n \to \infty$ the first summand goes to 0 by continuity of f. The second summand is by definition equal to $|g_n(t, \langle x, e_1\rangle, ..., \langle x, e_n\rangle) - \tilde{g_n}(t, \langle x, e_1\rangle, ..., \langle x, e_n\rangle)|$ and if $\max(\|x\|_H, |t|) < n$ this expression is smaller than $\frac{1}{n}$. Thus pointwise convergence is proved. Furthermore we have:

$$\sup_{n,t,x} |f_n(t,x)| \le \sup_{n,t,x} |g_n(t, \langle x, e_1\rangle, ..., \langle x, e_n\rangle)| + \frac{1}{n} \le \sup_{t,x} |f(t,x)| + 1$$

where the first inequality holds by periodicity of $\tilde{g_n}$ in t and x.
Finally, it is not difficult to see that indeed $f_n \in K$, since the e_i were chosen from $D(A^*)$.

Chapter 6

The Case of Periodic Coefficients

In the case of periodic coefficients, we can work on a different state space, switching from $\mathbb{R} \times H$ to $I \times H$ where I is a bounded intervall whose length equals the period. This allows in the definition of the measure ν to drop the exponential weight function.(We will be abusing notation by denoting this generalized invariant measure still by ν.) Translation invariance of (normalized) Lebesgue measure then guarantees full invariance for ν. This invariance property, in turn, will allow for stronger statements of some results obtained so far.

6.1 A Fully Invariant Measure for the Reduced Equation

In order to obtain a fully invariant measure let us henceforth assume that we deal with periodic coefficents.

Assumption 6.1.1 *There is $T > 0$ such that the coefficients A, f and B in equation* (3.1) *are T-periodic.*

As a first consequence, we can prove that the evolution system of measures established in Theorem 4.2.4 is periodic as well. We also get a uniqueness property. The evolution system found is the only *periodic* one.

Proposition 6.1.2 *Let $(\nu_t)_{t\in\mathbb{R}}$ be the evolution system of measures defined in Theorem 4.2.4. Then we have $\nu_t = \nu_t + T$ for any $t \in \mathbb{R}$. Any other T-periodic evolution system of measures coincides with the above.*

PROOF Recall the form of $\hat{\nu}_t$

$$\hat{\nu}_t(h) := \exp\left\{ i \left\langle h, \int_{-\infty}^{t} U(t,r)f(r)dr \right\rangle \right\} \exp\left\{ \int_{-\infty}^{t} \lambda\{B^*(r)U^*(t,r)h\}dr \right\}.$$

To establish T-periodicity, first note that we have
$U(t,s) = U(t+T, s+T)$ for any $s < t$, which follows easily from its defining differential equation and the assumption that $t \mapsto A(t)$ is T-periodic. Hence, we get

$$\int_{-\infty}^{t+T} U(t+T,r)f(r)dr = \int_{-\infty}^{t} U(t+T,r+T)f(r+T)dr = \int_{-\infty}^{t} U(t,r)f(r)dr$$

and for the other integral in the definition of $\hat{\nu}_t$ the argument is analogous.

To prove uniqueness, let $\{\mu_s\}$ be another T-periodic family satisfying (4.1), then we get using (4.1) with $\psi(x) = e^{i\langle h,x\rangle}$ and Lemma 4.1.1 :

$$\begin{aligned}\hat{\mu_s}(h) = \hat{\mu}_{s+T}(h) = \quad & \hat{\mu_s}(U^*(s+T,s)h) \\ \times \exp & \left\{ i \left\langle h, \int_s^{s+T} U(s+T,r)f(r)dr \right\rangle + \int_s^{s+T} \lambda\{B^*(r)U^*(s+T,r)h\}dr \right\}\end{aligned}$$

Furthermore, with the help of the following easy to check relations (recall that $U(s,r-T) = U(s+T,r)$ by periodicity):

$$\int_s^{s+T} U(s+T,r)f(r)dr = \int_{-\infty}^{s} U(s,r)f(r)dr - U(s+T,s)\int_{-\infty}^{s} U(s,r)f(r)dr$$

$$\begin{aligned}&\int_s^{s+T} \lambda\{B^*(r)U^*(s+T,r)h\}dr = \\ &\qquad \int_{-\infty}^{s} \lambda\{B^*(r)U^*(s,r)h\}dr - \int_{-\infty}^{s} \lambda\{B^*(r)U^*(s+T,r)h\}dr\end{aligned}$$

we arrive at:

$$\begin{aligned}&\hat{\mu_s}(h) = \\ &\hat{\mu_s}(U^*(s+T,s)h) \exp\left\{ i \left\langle h, \int_{-\infty}^{s} U(s,r)f(r)dr \right\rangle + \int_{-\infty}^{s} \lambda\{B^*(r)U^*(s,r)h\}dr \right\} \\ &\times \exp\left\{ i \left\langle h, -U(s+T,s)\int_{-\infty}^{s} U(s,r)f(r)dr \right\rangle - \int_{-\infty}^{s} \lambda\{B^*(r)U^*(s+T,r)h\}dr \right\}\end{aligned}$$

or equivalently:

$$\begin{aligned}&\hat{\mu_s}(h) \left[\exp\left\{ i \left\langle h, \int_{-\infty}^{s} U(s,r)f(r)dr \right\rangle \right\} \exp\left\{ \int_{-\infty}^{s} \lambda\{B^*(r)U^*(s,r)h\}dr \right\} \right]^{-1} \\ &\quad = \hat{\mu_s}(U^*(s+T,s)h) \left[\exp\left\{ i \left\langle U^*(s+T,s)h, \int_{-\infty}^{s} U(s,r)f(r)dr \right\rangle \right\} \right. \qquad (*) \\ &\qquad\qquad \left. \times \exp\left\{ \int_{-\infty}^{s} \lambda\{B^*(r)U^*(s,r)U^*(s+T,s)h\}dr \right\} \right]^{-1}\end{aligned}$$

Now the right-hand side of the last equation is the same as the left-hand side, only h is replaced by $U^*(s+T,s)h$. But as $h \in H$ was arbitrary, we can rerun the argument with $U^*(s+T,s)h$ replacing h. Iterating this procedure n times, we will obtain (*) with $U^*(s+T,s)h$ replaced by $[U^*(s+T,s)]^n h$. But, as $\|U^*(s+T,s)\| < 1$ holds by our stability assumption, all the factors on the right hand side will tend to 1, if we take n to infinity. Thus, $\hat{\mu_s}$ must have the desired form. ∎

The uniqueness we just proved for the evolution system of measures will extend also to the generalized invariant measure on the extended state space which we set out to define now.

Definition 6.1.3 *As in the non-periodic case, by Lemma 4.3.2 we can form* $\nu := \frac{1}{T}dt \otimes F$, *a measure on* $\mathbb{R} \times H$, *defined by* $\nu([s,t] \times A) := \frac{1}{T}\int_s^t \nu_r(A)dr$ *for* $s < t \in \mathbb{R}$ *and* $A \in \mathcal{B}(H)$.

$$L^2_*(\nu) := \{f : \mathbb{R} \times H \to \mathbb{R} \text{ measurable} \mid f(t+T,x) = f(t,x) \quad \nu - a.e.$$
$$\int_{[0,T]\times H} |f(y)|^2 \nu(dy) < \infty\}$$

$$\mathcal{M}_* := span_{\mathbb{C}}\{f : \mathbb{R} \times H \to \mathbb{C} \quad | \quad f = \Phi(t)e^{i\langle U(t+\sigma,t)x,h\rangle}, \text{ where}$$
$$\Phi \in C^1(\mathbb{R},\mathbb{R}) \text{ and } T\text{-periodic}, h \in D(A^*), \sigma \geq 0\}$$

$$K_* := \{\Re(f) \mid \quad f \in \mathcal{M}\}$$

That is, K_* *comprises the real parts of the functions in* $\mathcal{M}$.

Remark 6.1.4 *It is not hard to see that* $L^2_*(\nu)$ *is a Hilbert space. Because of the periodicity it is clear, that* $(\int_{[0,T]\times H} \|f\|^2(y)\nu(dy))^{\frac{1}{2}}$ *is a norm(where we introduce* ν *a.e.-equivalence classes as usual). Given a Cauchy-sequence* f_n *we consider* f_n^z *the restriction of* f_n *to the interval* $I_z := [zT,(z+1)T], z \in \mathbb{Z}$. *By Riesz-Fischer we obtain a limit* f^0 *of* f_n^0 *on* $[0,T]$, *and because of periodicity it is clear that the other restrictions form the same Cauchy sequences, that is:* $\lim f_n^z = f^z = f^0 \; \forall z \in \mathbb{Z}$. *Hence, the limit function is periodic, and the space is complete.*

Lemma 6.1.5 K_* *is dense in* $L^2_*(\nu)$.

PROOF Note that by periodicity, we can think of our functions to be defined on $[0,T] \times H$ and in the following we will do so without changing notation.
We will show density of $\mathcal{M}$ in $L^2_*(\nu;\mathbb{C})$. This implies density of the respective real vector spaces. We will use complex monotone classes again. The subsystem $\mathcal{M}_0 := \{\Phi_k \otimes \exp_h\}_{\{k\in\mathbb{Z},h\in D(A^*)\}}$, where $\Phi_k := \exp(k\,\cdot\,)$ and $\exp_h := \exp(i\langle\,\cdot\,,h\rangle)$ is closed under multiplication and conjugation. Consider $\mathcal{H} := \bar{\mathcal{M}}_0$ as a subspace of $L^2_*(\nu;\mathbb{C})$ where we allow complex-valued integrable and periodic functions. By monotone convergence, applied separately to real and imaginary

parts, $\mathcal{H}$ is seen to be a complex monotone vector space. Thus, $\mathcal{H}$ contains all $\sigma(\mathcal{M}_0)$-measurable functions. If we can show that $\sigma(\mathcal{M}_0) = \mathcal{B}(H \times [0,T])$ we will have all step functions in $\mathcal{H}$, hence density will be obvious. Note that we want to show that functions of the form $\Phi_k \otimes \exp_h$ generate a product σ-algebra $\mathcal{B}([0,T]) \otimes \mathcal{B}(H)$.

Knowing that the Φ_k generate $\mathcal{B}([0,T])$ and that the $\exp_h$ generate $\mathcal{B}(H)$ (which follows again from Lemma 4.1.4), it is clear that $\sigma(\{\Phi_0 \otimes \exp_h\}_{h\in D(A^*)}) = ([0,T] \times \mathcal{C})_{\mathcal{C}\in\mathcal{B}(H)}$ and $\sigma(\{\Phi_k \otimes \exp_0\}_{k\in\mathbb{Z}}) = (\mathcal{A} \times H)_{\mathcal{A}\in\mathcal{B}([0,T])}$ and the result is obvious. ■

Proposition 6.1.6 *Let Assumption 4.2.1 hold. Then the measure ν is the unique invariant probability measure for the semigroup P_τ. That is we have for every bounded measurable function u such that $u(t+T,x) = u(t,x)$ $\forall t > 0$, $x \in H$:*

$$\int_{[0,T]\times H} P_\tau u(t,x)\nu(dt,dx) = \int_{[0,T]\times H} u(t,x)\nu(dt,dx) \quad \forall \tau > 0$$

and if this holds for another probability measure μ then $\mu = \nu$. Furthermore, the semigroup P_τ is a contraction on $L^2_*(\nu)$.

Proof Let u be measurable, bounded and T-periodic in its first component. Let us write $u_t(x) := u(t,x)$, then we have $(P_\tau u)(t,x) = (P(t,t+\tau)u_{t+\tau})(x)$. Taking into account (4.1) we have:

$$\int_{[0,T]\times H} P_\tau u(t,x)\nu(dt,dx) = \frac{1}{T}\int_{[0,T]}\int_H (P(t,t+\tau)u_{t+\tau})(x)\nu_t(dx)dt$$

$$= \frac{1}{T}\int_{[0,T]}\int_H u_{t+\tau}(x)\nu_{t+\tau}(dx)dt = \frac{1}{T}\int_{[\tau,T+\tau]}\int_H u_t(x)\nu_t(dx)dt$$

$$= \frac{1}{T}\int_{[0,T]}\int_H u_t(x)\nu_t(dx)dt = \int_{[0,T]\times H} u(t,x)\nu(dt,dx)$$

because of translation invariance of dt and T-periodicity of u and ν_t.

For the contraction property we have to show for u as above: $\|P_\tau u\|_{L^2_*} \leq \|u\|_{L^2_*}$,
Using the Jensen inequality for the expectation and afterwards the invariance property for u^2:

$$\|P_\tau u\|^2_{L^2_*} = \int_{[0,T]\times H} \mathbb{E}[u(t+\tau, X(t+\tau,t,x))]^2\nu(dt,dx)$$

$$\leq \int_{[0,T]\times H} \mathbb{E}[u^2(t+\tau, X(t+\tau,t,x))]\nu(dt,dx) = \int_{[0,T]\times H} (P_\tau u^2)(t,x)\nu(dt,dx)$$

$$= \int_{[0,T]\times H} u^2(t,x)\nu(dt,dx) = \|u\|^2_{L^2_*}$$

Hence, P_τ is a contraction on a dense subset of $L^2_*(\nu)$. Thus, we can extend it to a contraction on all of $L^2_*(\nu)$. Clearly, the invariance property also extends to any $u \in L^2_*(\nu)$.

To show uniqueness, let μ be another invariant probability measure for P_τ, so that we have for all measurable, bounded and T-periodic functions u:

$$\int_{[0,T]\times H} P_\tau u(t,x)\mu(dt,dx) = \int_{[0,T]\times H} u(t,x)\mu(dt,dx) \quad \forall \tau > 0 \tag{6.1}$$

By [21] : Corollary 10.2.8 , we can disintegrate μ as follows:

$$\int_{[0,T]\times H} u(t,x)\mu(dt,dx) = \int_{[0,T]} \left(\int_H u(t,x)\mu_t(dx) \right) \mu_1(dt) \tag{6.2}$$

for the marginal $\mu_1(dt) = \mu \circ Pr^{-1}$ where Pr is the Projection on the t-component, and $\{\mu_t\}_{t\in\mathbb{R}}$ is a family of probability measures on H. Choosing $u(t,x) = f(t)$ independent of x in (6.1) we have by (6.2):

$$\int_{[0,T]} f(t+\tau)\mu_1(dt) = \int_{[0,T]} f(t)\mu_1(dt)$$

Since f is T-periodic, μ_1 is translation invariant (note, that we need here a similar monotone class argument as in Lemma 4.3.4). So μ_1 must be Lebesgue measure.

To show $\mu_t = \nu_t$, we will make use of the uniqueness proof from Theorem 4.2.4. Choosing $u(t,x) = f(t)g(x)$ and $\tau = T$ in (6.1) yields:

$$\int_{[0,T]} f(t) \left(\int_H P(t,t+T)g(x)\mu_t(dx) \right) \mu_1(dt) = \int_{[0,T]} f(t) \left(\int_H g(x)\mu_t(dx) \right) \mu_1(dt)$$

Clearly, if this holds for a fixed, bounded g and arbitrary bounded f, we must have

$$\int_H P(t,t+T)g(x)\mu_t(dx) = \int_H g(x)\mu_t(dx) \quad dt - a.e. \tag{$*$}$$

The dependence of the above null set on g can be overcome by choosing a countable and multiplicative family of functions separating the points of H (e.g. all functions of the form $g(x) = e^{\langle x, \sum_n r_n e_n \rangle}$, $r_n \in \mathbb{Q}$, $(e_n)_{n\in\mathbb{N}}$ form an ONB of H) and applying a monotone class argument . Now it can be checked that the uniqueness proof of Theorem 4.2.4 still works - though $(*)$ is weaker than (4.1) - and we obtain $\nu_t = \mu_t \quad dt - a.e.$ which, of course, implies $\nu = \mu$. ■

6.2 Additional Results through Invariance

In the case of an invariant measure we can obtain interesting results on the spectrum of our generator. Moreover by integrating the Poincare inequality 5.2.4 with the invariant measure we obtain its stronger form 6.2.3 as a corollary.

Lemma 6.2.1 *For all $u \in D(G)$ we have*

$$\int_{[0,T]\times H} Gu(t,x)\nu(dt,dx) = 0 \tag{6.3}$$

PROOF By invariance of P_τ with respect to ν, we have for all $u \in D(G)$:

$$0 = \int_{[0,T]\times H} \frac{1}{\tau}\left(P_\tau u(t,x) - u(t,x)\right)\nu(dt,dx)$$

Letting $\tau \to 0$ we obtain the result, since the integrand converges to Gu in $L^2_*(\nu)$. ∎

Exactly as in [14] we obtain the following result on the spectrum of G:

Corollary 6.2.2 *For any $z \in \sigma(G)$ and $k \in \mathbb{Z}$ we have $z + 2\frac{\pi}{T}ki \in \sigma(G)$. Moreover 0 is a simple eigenvalue of G.*

PROOF Analogous to [14] Corollary 5.5 .

For fixed $k \in \mathbb{Z}$ consider the operator $T_k u(t,x) = e^{2k\frac{\pi}{T}it}u(t,x)$. Since T is unitary the spectrum of G is equal to the spectrum of $T_k^{-1}GT_k = G + (2ki\frac{\pi}{T})Id$. where the equality holds, because the factors cancel out everywhere, except for the derivative with respect to t where the product rule applies. This proves the first statement.
Since every unique invariant measure is ergodic, we also have the equivalent property (see [16]):
If $u \in L^2_*$ fulfills $P_\tau u = u$ for every $\tau > 0$ then u is equal to a constant in L^2_*.
Now, let $u \in L^2_*(\nu)$ such that $Gu = 0$. Hence, $P_\tau u - u = \int_0^\tau P_s Gu ds = 0$ and thus $P_\tau u = u$ which in turn implies (by ergodicity of ν) that u is a constant function. Finally, if $G^2 u = 0$ then Gu must be equal to a constant, but by Lemma 6.2.1 this constant must be 0, hence $\operatorname{Ker} G^2 = \operatorname{Ker} G$. ∎

Corollary 6.2.3 *In the situation of Proposition 5.2.4 let $C(\infty) < \infty$. Then we have for all $u \in K$:*

$$\int_{[0,T]\times H} \left(u(t,x) - \overline{u}_t\right)^2 \nu(dt,dx) \le C(\infty)\int_{[0,T]\times H} \Gamma(u,u)\nu(dt,dx)$$

PROOF Integrating (5.4) with respect to ν yields, because of invariance:

$$\int_{[0,T]\times H} \left(u^2 - (P_\tau u)^2\right)\nu(dt,dx) \le C(\tau)\int_{[0,T]\times H} \Gamma(u,u)\nu(dt,dx)$$

Letting $\tau \to \infty$, using Proposition 5.2.3 together with dominated convergence and employing $\int_{[0,T]\times H} \overline{u}^2_{t+\tau} d\nu = \int_{[0,T]\times H} \overline{u}^2_t d\nu$ we obtain:

$$\int_{[0,T]\times H} \left(u^2 - (\overline{u}_t)^2\right)\nu(dt,dx) \le C(\infty)\int_{[0,T]\times H} \Gamma(u,u)\nu(dt,dx)$$

Since $\overline{u}_t$ does not depend on x, we have:

$$\int_{[0,T]\times H} \left(u(t,x) - \overline{u}_t\right)^2 \nu(dt,dx) = \int_{[0,T]\times H} \left(u^2(t,x) - 2u(t,x)\overline{u}_t + \overline{u}_t^2\right) \nu(dt,dx)$$

$$= \int_{[0,T]\times H} \left(u^2(t,x) + \overline{u}_t^2\right) \nu(dt,dx) - 2\int_{[0,T]} \overline{u}_t \underbrace{\int_H u(t,x)\nu_t(dx)}_{\overline{u}_t} dt$$

■

Chapter 7

The Variational Interpretation of the associated Fokker-Planck Equation

In this chapter we will study Ornstein-Uhlenbeck equations in a different context. As a consequence of our study of the Kolmogorov equations related to our SPDE, it is trivial to see that also the Fokker-Planck equations – the dual equations for the distributions – can be solved. For simplicity let us go back to the time-independent case for a moment

$$dX_t = AX_t dt + dW_t.$$

In this setting the evolution system of measures collapses into a unique invariant measure, say μ and we have a one-parameter transition semigroup P_t with generator, say L. For a given function f, $P_t f$ then solves the Kolmogorov equation $\frac{d}{dt}u(t,x) = Lu(t,x)$, whereas for a given measure ρ, $P_t^* \rho$ solves the Fokker-Planck equation $\frac{d}{dt}\int_H h(x)\rho_t(dx) = \int_H Lh(x)\rho_t(dx)$. Moreover, in our setting, it is clear that we will have $\lim_{t\to\infty} \rho_t = \mu$.

In [27], for the first time, insight was provided into the dynamics of attraction towards the invariant measure μ by making a connection between the Fokker-Planck equation and a certain variational problem.(Actually, the framework of [27] is more general and may also apply if no invariant measure exists.) The flow of measures generated by the Fokker-Planck equation was seen to be directed towards the invariant measure μ in such a way as to minimize a certain functional generated by the measure μ. The *relative entropy* functional with respect to μ. More precisely, the entropy functional generates a *gradient flow* in a certain metric space of probability measures, the *Wasserstein* space. The theory of gradient flows in general metric spaces is covered extensively in [3]. See also [12] for a nice introduction. Using this theory, the variational interpretation of the Fokker-Planck equation introduced in [27] has recently been extended (in [22] and [32]) to infinite-dimensional situations. But one always works with a fixed reference measure, the unique invariant measure of some underlying equation. In our non-autonomous setting, it turns out that we must deal with a family of reference measures. But unlike in the previous part of this work, this family is not given by an evolution system of measures. It is obtained in the following way: at every instance of time, we freeze the coefficients of our equation and use the invariant measure of this equation as a reference measure. This takes us outside the theory of [3] where the energy functional must not depend on time. Indeed we will not establish a real gradient flow, we will just show the existence of a so-called (see [3] page 279) generalized minimizing movement scheme as used e.g. in [27]. That is, we will set up a time-dependent and time-discrete variational approximation scheme which generates curves of measures, prove the existence of a limit curve and show that this limit solves a Fokker-Planck equation.

7.1 The Framework

Let us consider a Gaussian Ornstein-Uhlenbeck Equation on a separable Hilbert space $(H, \langle\cdot,\cdot\rangle)$ with time-dependent coefficients

$$dX_t = -A(t)X_t + dW_t. \tag{7.1}$$

where W_t is a R-Brownian motion on H with trace class covariance operator R and where $A(t)$ are (possibly unbounded) linear operators, such that for each fixed $t_0 \in \mathbb{R}$, $A(t_0)$ generates a strongly continuous contraction semigroup of bounded operators $(S_s^{t_0})_{s\geq 0}$, such that $\|S_s^{t_0}\|_{\mathcal{L}(H)} \leq e^{-\omega s}$ for some $\omega > 0$. The maps $t \mapsto A(t)$ and $t \mapsto A^*(t)$ are assumed to be strongly continuous and the domains $D(A(t)) =: D(A)$ and $D(A^*(t)) =: D(A^*)$ are assumed to be independent of t. Since we are only interested in a solution of the Fokker-Plank equation, in this chapter, we do not need that the operators $A(t)$ generate an exponentially bounded evolution family as in Assumption 3.1.2. However, Assumption 7.1.4 (i) and (ii) implicitly pose some restrictions on the map $t \mapsto A(t)$.

The following definition introduces a family of reference measures which, as in the first part, will serve as a replacement for a single invariant measure.

Definition 7.1.1 *In equation* (7.1)*, for every fixed time t_0 we consider the corresponding autonomous equation*

$$dX_t = -A(t_0)X_t + dW_t.$$

By our assumptions on the coefficients it admits a unique invariant measure. We will denote it by ν_{t_0} and $\{\nu_{t_0}\}_{t_0\in\mathbb{R}}$ will be the collection of all these 'invariant measures' obtained by freezing the coefficients. The covariance operator of the Gaussian measure ν_{t_0} will be denoted by R_{t_0}. Moreover, G_{t_0} will be the Generator associated to the equation with coefficients frozen at t_0.

Remark 7.1.2 *The covariance operators R_{t_0} have the form*

$$R_{t_0} = \int_0^\infty (S_s^{t_0})^* R S_s^{t_0} ds$$

and they are of trace class as well. It is also well known that on suitable test functions f we have

$$G_{t_0} f(x) = \langle x, A^*(t_0) \nabla_x f(x) \rangle + \mathrm{Tr}(\sqrt{R}^* \nabla_{xx} f(x) \sqrt{R}).$$

The reference measures just introduced above, will induce a family of entropy functionals.

Definition 7.1.3 *For probability measures μ and ν on H, we will denote by $\mathrm{Ent}_\nu(\mu)$ the entropy of μ relative to ν, i.e.*

$$\mathrm{Ent}_\nu(\mu) := \int_H \log\left(\frac{d\mu}{d\nu}(x)\right) \mu(dx).$$

If the Radon-Nikodym derivative $\frac{d\mu}{d\nu}$ does not exist we will set the entropy to be $+\infty$.
We will say that $\mu \in D(\mathrm{Ent}_\nu)$ if $\mathrm{Ent}_\nu(\mu) < \infty$.

It is not difficult to see that the we have $\mathrm{Ent}_\nu(\mu) \geq 0$ for all probability measuresν and μ.

Assumption 7.1.4 *The coefficients in (7.1) are such that*

(i) there is a constant $K_1 > 0$ such that

$$\sup_{s,t \in \mathbb{R}} \mathrm{Ent}_{\nu_s}(\nu_t) < K_1.$$

(ii) there is a constant $K_2 > 0$ and a bounded function $f : \mathbb{R}^2 \to \mathbb{R}$ with $f(r,s)f(s,t) = f(r,t)$, $r \leq s \leq t$ such that:

$$\left| \frac{d\nu_s}{d\nu_t}(x) \right| \leq \exp(K_2 |t-s| \|x\|^2) f(s,t),$$

where $\frac{d\nu_s}{d\nu_t}(x)$ is the Radon-Nikodym derivative which exists since ν_s and ν_t are equivalent by (i).

(iii) For every $t \in \mathbb{R}$ and $s \geq 0$ we have:

$$S_t(s) R = R S_t^*(s),$$

where $(S_t(s))_{s \geq 0}$ is the semigroup generated by the operator $A(t)$.

(iv) There is $\beta > 0$ such that:

$$\langle R_t x, x \rangle \leq \beta \langle R x, x \rangle \quad \forall x \in H,\, t \in \mathbb{R}$$

where R_t is the covariance operator introduced in Definition 7.1.1.

Remark 7.1.5 *Let us elaborate on these rather technical assumptions.*

(i) implies that the reference measures $(\nu_t)_{t\in\mathbb{R}}$ are all equivalent. It ensures that the domains of the different entropy functionals are very similar allowing the construction of a non-autonomous discrete gradient flow. It is also vital for proving tightness of the resulting measures in Proposition 7.3.2.

In finite dimensions (i) would be satisfied if we required the (negative) eigenvalues of the matrices $A(t)$ to be bounded uniformly away from 0 and $-\infty$.

(ii) is best seen as a technical description of smoothness of the map $t \mapsto \nu_t$. It is required for Lemma 7.4.1. In one dimension it would be sufficient to have $t \mapsto A(t)$ continuously differentiable.

(iii) ensures that for every fixed t the respective Ornstein-Uhlenbeck semigroup is self-adjoint. This assumption is needed to show that the gradient flow solves a Fokker-Planck equation.

(iv) is a compatibility condition for the noise and the reference measures in terms of their covariance operators. It is necessary to establish convexity of the entropy functionals (induced by the reference measures) with respect to the distance (induced by the noise) in Theorem 7.2.11.

Assumptions (iii) and (iv) stem from the autonomous case and seem thus unavoidable.

If (i) and (ii) are interpreted as ensuring the 'boundedness' and smoothness of the map $t \mapsto \nu_t$ then these seem rather natural conditions for the existence of a continuous non-autonomous flow.

7.2 Entropy Minimizing Movements

We want to set up a discrete scheme governed by relative Entropy and the Wasserstein metric and show that it will converge to the solution of the Fokker-Planck equation. In the autonomous case, it is well known that relative entropy with respect to the invariant measure generates a gradient flow that solves the Fokker-Planck equation. This was proven in [22] and [32] in slightly different settings.

7.2.1 The Wasserstein Distance and Optimal Transport

The Wasserstein distance will play a prominent role in the definition of the discrete scheme. For a nice introduction into optimal transportation and Wasserstein distances see the recent book [46]. It also gives a good account on convexity and gradient flows in Wasserstein spaces.

Let $(H, \|\cdot\|_H)$ be a Hilbert space. Let $\mathcal{P}(H)$ be the space of all Borel probability measures on H. Let us equip this space with the topology of weak convergence, where as test functions we consider all bounded functions which are continuous with respect to the norm $\|\cdot\|_H$. Although we will talk about different norms for the definition of the Wasserstein distance in an instant, we stress that the weak topology is fixed as the one introduced above.

The following definition of the Wasserstein distance on $\mathcal{P}(H)$ will depend on the distance to be fixed on the underlying Hilbert space. This need not necessarily be the original Hilbertian norm. Thus, fix a (possibly different) norm $\|\cdot\|_N$ on H.

Definition 7.2.1 *For $\mu, \nu \in \mathcal{P}(H)$ one can define the 2-Wasserstein distance $d_N(\mu, \nu)$ between μ and ν as follows:*

$$d_N^2(\mu, \nu) := (d_N(\mu, \nu))^2 = \inf_{\Sigma \in \Gamma(\mu,\nu)} \int_{X \times X} \|x - y\|_N^2 \Sigma(dx, dy),$$

where Γ is the set of all Borel probability measures on $X \times X$ with marginals μ and ν.

For the notation, let us stress that we denote the dependence on the underlying norm $\|\cdot\|_N$ by writing d_N instead of just d. Besides the 2-Wasserstein distance one can also define a p-Wasserstein distance by changing the exponent of the norm. We will always work with the 2-Wasserstein distance only, so we refrain from including a 2 in the notation. Moreover we will often work with the squared Wasserstein distance, so a superscript 2 will always mean a plain square as stressed in the above definition.

For the intuition, very loosely speaking, in order to compute the Wasserstein distance, one has to look for a coupling of ν and μ whose mass is concentrated on pairs of points which are close with respect to $\|\cdot\|_N$.

The first question raised by this definition is: Why should this expression be finite? Indeed in general this will not be true, but if we require both ν and μ to have finite second moments (with respect to $\|\cdot\|_N$) we get immediately for the trivial coupling $\mu \otimes \nu$:

$$d_N^2(\mu, \nu) \leq \int_{X \times X} (2\|x\|_N^2 + 2\|y\|_N^2)\mu(dx)\nu(dy) < \infty$$

But in general the assumption of second moments is too strong and we will continue to work on the space $\mathcal{P}(H)$, living with the fact that the distance d_N may take the value infinity.

Next, one can establish that, if the infimum is finite, it is indeed a minimum. Recall that every Hilbert space is in particular also a Radon space, so every single probability measure on H is tight. Thus, since μ and ν are both tight so is $\Gamma(\mu, \nu)$ (it is also weakly closed) and so we have a minimizing convergent sequence of measures. As long as the cost functional $\|\cdot\|_N$ is lower semicontinuous with respect to the original Hilbertian norm H defining the weak topology, its limit point is then indeed a minimizer. (See e.g. [32] Proposition 6.6) Such a minimizer will be called an *optimal plan* or an *optimal coupling*.

Let us now specify the norm we will work with in order to find solutions to the Fokker-Planck equation. As for the Harnack inequality 5.2.6, on H define the Cameron-Martin norm

$$\|x\|_R := \inf_{z \in H} \{\|z\|_H \mid \sqrt{R}z = x\}$$

which will be infinite if x does not lie in the image of $\sqrt{R}$.

It is an easy consequence of the definition that the Cameron-Martin norm is stronger than the Hilbertian norm of H with

$$\|\cdot\|_H \leq \|\sqrt{R}\|_{\mathcal{L}(H)} \|\cdot\|_R.$$

Since the compact operator $\sqrt{R}$ is not boundedly invertible, we cannot have the converse, that is the norm $\|\cdot\|_R$ cannot be continuous with respect to $\|\cdot\|_H$. Nevertheless, one can show that it is lower semicontinuous with respect to $\|\cdot\|_H$, as proved for example in [32] Lemma 6.5. This is important since it makes sure that optimal plans always exist as mentioned above.

The Wasserstein distance is normally defined with respect to a true distance. But also in our case of a pseudo distance which can be infinite, all nice properties of the Wasserstein distance are conserved. In particular we have:

Proposition 7.2.2 *The Wasserstein distance d_R is a complete pseudo metric on $\mathcal{P}(H)$, that is, we have for any $\mu, \nu, \rho \in \mathcal{P}(H)$:*

- $\mu = \nu \iff d_R(\mu, \nu) = 0$,
- $d_R(\mu, \nu) \leq d_R(\mu, \rho) + d_R(\rho, \nu)$,
- *any Cauchy sequence with respect to d_R coverges to a limit in $\mathcal{P}(H)$.*

Proof See Proposition 6.8 in [32]. ■

It will often be useful to compare Wasserstein distances induced by different norms.

Lemma 7.2.3 *Given two lower semicontinuous norms $\|\cdot\|_A, \|\cdot\|_B$ on H, then $\|\cdot\|_A \leq C\|\cdot\|_B$ implies $d_A(\mu, \nu) \leq C\ d_B(\mu, \nu)$ for all $\mu, \nu \in \mathcal{P}(H)$ and some $C > 0$.*

Proof If $d_B(\mu, \nu)$ is infinite there is nothing to prove, hence let Σ be an optimal coupling of μ and ν with respect to the cost $\|\cdot\|_B$. Then we have:

$$\begin{aligned} W_A^2(\mu, \nu) &\leq \int_H \|x - y\|_A^2 \Sigma(dx, dy) \\ &\leq C^2 \int_H \|x - y\|_B^2 \Sigma(dx, dy) \\ &= C^2\ W_B^2(\mu, \nu) \end{aligned}$$

■

Bounds on the Wasserstein distance can be transferred into bounds on the second moments.

Lemma 7.2.4 *Let $\mu, \nu \in \mathcal{P}(H)$. If $\int_H \|x\|_H^2 \nu(dx) < \infty$ and $d_R(\mu, \nu) < \infty$ then we have $\int_H \|x\|_H^2 \mu(dx) < \infty$ as well.*

PROOF Let Σ be an optimal coupling of μ and ν with respect to the norm $\|\cdot\|_R$. Then we have

$$\begin{aligned}\int_H \|x\|_H^2 \mu(dx) &= \int_H \|x-y+y\|_H^2 \Sigma(dx,dy) \\ &\le 2\int_H \|x-y\|_H^2 \Sigma(dx,dy) + 2\int_H \|y\|_H^2 \Sigma(dx,dy) \\ &\le 2\int_H \|\sqrt{R}\|_{\mathcal{L}(H)}^2 \|x-y\|_R^2 \Sigma(dx,dy) + 2\int_H \|y\|_H^2 \nu(dy) < \infty \quad \blacksquare\end{aligned}$$

The following important inequality allows to bound the Wasserstein distance by the relative entropy. As shown in [32] it follows readily from the Talagrand inequality in Wiener spaces from [23] in connection with Assumption 7.1.4 (iv).

Lemma 7.2.5 (Talagrand Inequality) *Let β be the constant from Assumption 7.1.4 (iv). Then we have for any $t \in \mathbb{R}$*

$$d_R^2(\nu_t, \mu) \le 2\beta \operatorname{Ent}_{\nu_t}(\mu). \tag{7.2}$$

7.2.2 The Discrete Scheme and Convexity in Wasserstein Spaces

Having defined the Wasserstein distance, we can now describe how to obtain a sequence of measures which we will use later on to construct a solution to the Fokker-Planck equation. The Wasserstein distance employed in the definition of the scheme will be the one induced by $\|\cdot\|_R$. This choice is crucial to establish tightness of the measures generated by this scheme.

We choose a starting distribution $\rho_0 \in D(\operatorname{Ent}_{\nu_0})$ and a time step $\tau > 0$. For notational simplicity, and without loss of generality, let us assume that the starting time is 0. The arbitrary yet finite time horizon will be denoted by T. For a given time step τ denote by N^τ the largest integer smaller than $\frac{T}{\tau}$. Denote $t_k := k\tau$, then we can define the measures $\rho_k := \rho_k^\tau$, $k = 1, ..., N^\tau$ recursively by the following scheme:

$$\rho_{k+1}^\tau := J_\tau^{t_k} \rho_k^\tau, \tag{7.3}$$

where the operator J is defined as follows:

Definition 7.2.6

$$\Phi_t(\tau, \rho; \mu) := \frac{1}{2\tau} d_R^2(\rho, \mu) + \operatorname{Ent}_{\nu_t}(\mu),$$

$$J_\tau^t \rho = \operatorname{argmin}_{\mu \in \mathcal{P}(H)} \quad \Phi_t(\tau, \rho; \mu),$$

$$\Phi_t^\tau(\rho) := \frac{1}{2\tau} d_R^2(\rho, J_\tau^t \rho) + \operatorname{Ent}_{\nu_t}(J_\tau^t \rho) \qquad \left[= \inf_\mu \Phi_t(\tau, \rho; \mu)\right].$$

$\Phi_t^\tau(\rho)$ is called the Moreau-Yosida approximation of the functional Ent_{ν_t} and J_τ^t is the corresponding resolvent.

In the autonomous case, once it is known that the entropy functional with respect to the unique invariant measure fulfills a certain convexity condition, general theory from [3] ensures that this scheme is well defined and converges to a continuous gradient flow with nice regularizing properties. In the time-dependent case we are not aware of any analogous theory. But since each 'frozen' invariant measure induces an entropy functional which enjoys the above mentioned convexity called *convexity along generalized geodesics* we can use the autonomous theory to make sure that each minimization problem is well-posed and in order to obtain a crucial *discrete evolution variational inequality* which will help to prove convergence towards a continuous flow.

In order to use the framework of [3], we need to check that the functionals $\mu \mapsto \mathrm{Ent}_{\nu_t}(\mu)$ are lower semicontinuous on $(\mathcal{P}(H), d_R)$. But Lemma 9.4.3 in [3] states that $\mu \mapsto \mathrm{Ent}_{\nu_t}(\mu)$ is lower semi continuous in the weak topology. And since d_R induces an even finer topology (see e.g. Theorem 6.10 in [32]) we are done.

In order for $J_\tau^t\rho$ to be well-defined we have to make sure that there is always a unique minimizer for the functional $\Phi_t(\tau, \rho; \cdot)$.

As was shown in [3] (see Assumption 4.0.1 there) the right condition on Φ_t is a rather intricate notion of convexity.

Proposition 7.2.7 *There is $\lambda > 0$ such that, for every t, $\Phi_t(\tau, \rho; \cdot)$ is $(\tau^{-1} + \lambda)$-convex. That is, for every choice of $\mu_0, \mu_1, \nu \in D(\mathrm{Ent}_{\nu_t})$ there is a curve $(\mu_\alpha^\nu)_{\alpha\in[0,1]}$ joining μ_0 and μ_1, such that for every $\alpha \in [0,1]$:*

$$\Phi_t(\tau, \nu, \mu_\alpha^{nu}) \leq (1-\alpha)\Phi_t(\tau, \nu, \mu_0) + t\Phi(\tau, \nu, \mu_1) - \frac{1+\lambda\tau}{2\tau}\alpha(1-\alpha)d_R^2(\mu_0, \mu_1)$$

PROOF This will follow as a consequence of general theory from [3] together with a result from [32]. ■

Let us give some intuition for this notion of convexity. The nature of the curves μ_α will be given in the next definition. $\Phi_t(\tau, \nu; \mu) = \frac{1}{2\tau} d_R^2(\nu, \mu) + \mathrm{Ent}_{\nu_t}(\mu)$ consists of two parts. The entropy part has good convexity properties, but the Wasserstein part is only convex for a very specific choice of connecting curves. The important observation in [3] was, that one can expect convexity of the functional $\mu \mapsto d_R^2(\nu, \mu)$ only along curves which explicitly depend on the parameter μ.

The following definition (Definition 9.2.2 in [3]) describes the form of such curves. For probability measures ρ_1, ρ_2 we denote here by by $\Gamma_0(\rho_1, \rho_2)$ all optimal couplings between ρ_1 and ρ_2 with respect to the cost $\|\cdot\|_R$. Moreover, for a function $f : X \to Y$ and a measure μ on X we denote by $f_\sharp\mu$ the image measure $\mu \circ f^{-1}$ on Y, where X and Y are measurable spaces.

Definition 7.2.8 (Generalized Geodesics) *Let $\mu_0, \mu_1, \nu \in \mathcal{P}(H)$. A generalized geodesic from μ_0 to μ_1 based in ν is a path $(\mu_t)_{t\in[0,1]}$ in $\mathcal{P}(H)$ which can* *be represented as*

$$\mu_t := (\pi_t^{2\to 3})_\sharp \Sigma$$

for some $\Sigma \in \Gamma(\nu, \mu_0, \mu_1)$ *satisfying also*

$$(\pi^{1,2})_\sharp \Sigma \in \Gamma_0(\nu, \mu_0), \qquad (\pi^{1,3})_\sharp \Sigma \in \Gamma_0(\nu, \mu_1)$$

Here,

$$\pi^{1,2} : H^3 \to H^2 \ (x, y, z) \mapsto (x, y) \qquad \pi^{1,3} : H^3 \to H^2 \ (x, y, z) \mapsto (x, z)$$

$$\pi_t^{2\to 3} : H^3 \to H^2 \ (x, y, z) \mapsto (1-t)y + tz$$

Then, $d_R^2(\nu, \cdot)$ is convex along any such generalized geodesic *based in* ν as stated in the next proposition.

Proposition 7.2.9 *Let* $(\mu_t)_{t\in[0,1]}$ *be a generalized geodesic based in* ν*. Then we have:*

$$d_R^2(\nu, \mu_t) \le (1-t) d_R^2(\nu, \mu_0) + d_R^2(\nu, \mu_1) - t(1-t) d_R^2(\mu_0, \mu_1).$$

PROOF See Proposition 8.7 in [32]. ∎

Now we need the entropy part to be compatible with these curves as well, according to the following definition:

Definition 7.2.10 (Convexity along Generalized Geodesics) *Let* $\lambda \in \mathbb{R}$. *A functional* $F : \mathcal{P}(H) \to \mathbb{R} \cup \{\infty\}$ *is called* λ*-convex along generalized geodesics if for any* $\mu_0, \mu_1, \nu \in D(F)$ *with* $d_R(\mu_0, \nu) < \infty$, $d_R(\mu_1, \nu) < \infty$, *there is a generalized geodesic from* μ_0 *to* μ_1 *based in* ν*, such that*

$$F(\mu_t) \le (1-t)F(\mu_0) + tF(\mu_1) - \frac{\lambda}{2} t(1-t) d_R^2(\mu_0, \mu_1), \quad t \in [0,1].$$

The corresponding result in [32] (Theorem 10.10) states:

Theorem 7.2.11 *Let* β *be the positive constant from Assumption 7.1.4 (iv). The relative entropy functional* Ent_{ν_t} *is* (β^{-1})*-convex along generalized geodesics in* $(\mathcal{P}(H), d_R)$.

Together with the good behavior of $\rho \mapsto d_R^2(\rho, \nu)$ along generalized geodesics based in ν, this result implies (according to Lemma 9.2.7 from [3]) the sought for convexity property of Proposition 7.2.7 with $\lambda = \beta^{-1}$.

As a consequence, we have the following result from [3], assuring the viability of our scheme.

Lemma 7.2.12 *If* $\rho \in D(\mathrm{Ent}_{\nu_t})$*, then for every* $\tau > 0$ $J_\tau^t \rho$ *is well-defined, i.e. the corresponding minimization problem is uniquely solvable.*

PROOF See [3] Theorem 4.1.2. ∎

To apply the above lemma successively, we just have to make sure that the measures generated by the scheme are always in the right domain of definition.

Proposition 7.2.13 *For every starting point $\rho_0 \in D(\mathrm{Ent}_{\nu_0})$ the scheme given in* (7.3) *is well-defined, that is, there is in every step a unique minimizer.*

Proof According to the above Lemma 7.2.12 we only need to make sure that, for every $k = 1, ..., N^\tau$, we have

$$\rho_k^\tau \in D(\mathrm{Ent}_{\nu_{k\tau}}).$$

We will show how $\rho_0 \in D(\mathrm{Ent}_{\nu_0})$ will imply $\rho_1 \in D(\mathrm{Ent}_{\nu_\tau})$. For general k the result is then clear by iteration. First of all we have $\rho_1 = J_\tau^0 \rho_0 \in D(\mathrm{Ent}_{\nu_0})$ since the defining minimization scheme enforces $\mathrm{Ent}_{\nu_0}(\rho_1) \leq \mathrm{Ent}_{\nu_0}(\rho_0)$. In particular we have $\rho_1 \ll \nu_0$ and thus also $\rho_1 \ll \nu_\tau$ since all reference measures are equivalent by Assumption 7.1.4 (i). Thus, the Radon-Nykodim derivative $\frac{d\rho_1}{d\nu_\tau}$ exists and moreover we can write it as

$$\frac{d\rho_1}{d\nu_\tau} = \frac{d\rho_1}{d\nu_0}\frac{d\nu_0}{d\nu_\tau}.$$

Thus, we have

$$\begin{aligned}
&\mathrm{Ent}_{\nu_\tau}(\rho_1) - \underbrace{\mathrm{Ent}_{\nu_0}(\rho_1)}_{<\infty} \\
&= \int\limits_{\{\frac{d\rho_1}{d\nu_\tau}>0\}} \log\left(\frac{d\rho_1}{d\nu_\tau}(x)\right)\frac{d\rho_1}{d\nu_\tau}(x)\nu_\tau(dx) \quad - \int\limits_{\{\frac{d\rho_1}{d\nu_0}>0\}} \log\left(\frac{d\rho_1}{d\nu_0}(x)\right)\frac{d\rho_1}{d\nu_0}(x)\nu_0(dx) \\
&= \int\limits_{\{\frac{d\rho_1}{d\nu_0}>0\}} \log\left(\frac{d\nu_0}{d\nu_\tau}(x)\frac{d\rho_1}{d\nu_0}(x)\left(\frac{d\rho_1}{d\nu_0}\right)^{-1}(x)\right)\rho_1(dx) \\
&= \int\limits_{\{\frac{d\rho_1}{d\nu_0}>0\}} \log\left(\frac{d\nu_0}{d\nu_\tau}(x)\right)\rho_1(dx) \\
&\leq \int_H K_2\tau f(0,\tau)\|x\|_H^2\rho_1(dx)
\end{aligned}$$

where we used Assumption 7.1.4 for the last inequality. If we can show that $\int_H \|x\|_H^2\rho_1(dx) < \infty$ the proof is finished. Therefore it is enough to remark the following. By Talagrands inequality 7.2.5 we have $d_R(\rho_0, \nu_0) < \infty$ and ν_0 has second $\|\cdot\|_H$ moment as a Gaussian measure with trace class covariance operator. Hence Lemma 7.2.4 assures that ρ_0 also has the second $\|\cdot\|_H$ moment. By the minimization property we have $d_R(\rho_0, \rho_1) < \infty$ and another application of Lemma 7.2.4 yields the result. ■

Another consequence of the convexity property is the following discrete variational inequality, which will be useful in the next section. See [3] Theorem 4.1.2. for the proof.

Proposition 7.2.14 (Discrete Evolution Variational Inequality) *Let $\rho \in D(\mathrm{Ent}_{\nu_t})$ and $\rho_\tau = J_\tau^t \rho$. Then, for each $\mu \in D(\mathrm{Ent}_{\nu_t})$ we have:*

$$\frac{1}{2\tau} d_R^2(\rho_\tau, \mu) - \frac{1}{2\tau} d_R^2(\rho, \mu) + \frac{1}{2\beta} d_R^2(\rho_\tau, \mu) \leq \mathrm{Ent}_{\nu_t}(\mu) - \mathrm{Ent}_{\nu_t}(\rho_\tau) - \frac{1}{2\tau} d_R^2(\rho_\tau, \rho) \tag{7.4}$$

where β is the constant from Assumption 7.1.4.

7.3 Compactness of the Discrete Trajectories

In this section we will establish the existence of some limiting curve of measures, generated by the discrete flows from the last section. To this end, we will find a bound on the Wasserstein distance for the elements of the discrete scheme and connect this with compactness properties of level sets of the Wassersteinian.

The following Lemma is Proposition 6.12 from [32]. It states that the Wasserstein distance with respect to the norm $\|\cdot\|_R$ enjoys compact level sets. This is not so surprising, as, due to the fact that the Gaussian covariance operator R is trace class, we have that $\{x \mid \|x\|_R \leq C\}$ is compact.

Lemma 7.3.1 *The Wasserstein distance has compact level sets in the following sense:*

For every $C > 0$ and every fixed measure $\nu \in \mathcal{P}(H)$ the set

$$K_C := \{\rho \in H \mid d_R(\rho, \nu) \leq C\}$$

is compact in the weak topology.

We will now obtain a bound on the Wasserstein distance with the help of the discrete evolution variational inequality. In comparison to the autonomous case, where the evolution is confined to a sublevel of the (single) entropy functional, it is not quite so clear how the flow could be restricted in our setting with varying entropies. The intuition is the following: Since the reference measures are all close to each other (formalized in Assumption 7.1.4 (i)), the attraction to one of them will also shorten the distance to each one, provided we are far enough away from the set of reference measures, so that we do not start 'in between' two of them.

Proposition 7.3.2 *Let $\rho_0 \in D(\mathrm{Ent}_{\nu_0})$ with $d_R(\rho_0, \nu_0) < \infty$. Then there is a constant C_1 depending only on ρ_0 and the time horizon T, such that we have*

$$\sup_{\tau>0, t \leq T, k \leq N^\tau} d_R(\rho_k^\tau, \nu_t) \leq C_1 \tag{7.5}$$

for all $\rho_k^\tau := J_\tau^0 \cdots J_\tau^{\tau(k-1)} \rho_0$ appearing in the discrete scheme.

PROOF We will show that if ρ_k^τ fulfills $\sup_t d_R(\rho_k^\tau, \nu_t) \le C_1$, so does ρ_{k+1}^τ accordingly. Thus, assume the contrary. Then, for some $s \in [0, T]$, we must have

$$d_R(\rho_{k+1}^\tau, \nu_s) \ge d_R((\rho_k^\tau, \nu_s),$$

otherwise it is clear that ρ_{k+1}^τ fulfills $\sup_t d_R(\rho_{k+1}^\tau, \nu_t) \le C_1$. Using the discrete evolution variational inequality (7.4) with $\mu = \nu_s$ we obtain:

$$\frac{1}{2\tau} d_R^2(\rho_{k+1}^\tau, \nu_s) - \frac{1}{2\tau} d_R^2(\rho_k^\tau, \nu_s) + \frac{1}{2\beta} d_R^2(\rho_{k+1}^\tau, \nu_s) \\ \le \operatorname{Ent}_{\nu_{k\tau}}(\nu_s) - \operatorname{Ent}_{\nu_{k\tau}}(\rho_{k+1}^\tau) - \frac{1}{2\tau} d_R^2(\rho_{k+1}^\tau, \rho_k) \quad (7.6)$$

and since $\beta > 0$ it follows that

$$\operatorname{Ent}_{\nu_{k\tau}}(\rho_{k+1}^\tau) \le \operatorname{Ent}_{\nu_{k\tau}}(\nu_s) \le K_1,$$

where K_1 is the constant from Assumption 7.1.4 which is determined by the reference measures only. By Talagrands inequality (7.2.5) we obtain a bound

$$d_R(\rho_{k+1}^\tau, \nu_{k\tau}) \le \sqrt{2\beta \operatorname{Ent}_{\nu_{k\tau}}(\rho_{k+1}^\tau)} \le \sqrt{2\beta K_1}.$$

Let us use the triangle inequality to obtain

$$\sup_{r\in\mathbb{R}} d_R(\rho_{k+1}^\tau, \nu_r) \le d_R(\rho_{k+1}^\tau, \nu_{k\tau}) + \sup_{r\in\mathbb{R}} d_R(\nu_{k\tau}, \nu_r) \le 2\sqrt{2\beta K_1}.$$

So, we can choose $C_1 := \max(2\sqrt{2\beta^2 K_1}, \sup_t d_R(\rho_0, \nu_t))$ and the proof is finished, since $\sup_t d_R(\rho_0, \nu_t)$ is finite by a similar reasoning with the triangle inequality. ■

As a consequence, in connection with Lemma 7.3.1 we obtain tightness of the measures appearing in our entropy minimizing movement.

Corollary 7.3.3 *The family of measures* $\{\rho_k^\tau \mid \tau > 0,\ k \le N^\tau\}$ *is tight.*

PROOF According to Lemma 7.3.1 the set $\{\rho \mid d_R(\rho, \nu_t) \le C\}$ is compact for any measure ν_t. ■

From Proposition 7.3.2 we can also deduce, that all measures produced by the minimizing scheme have uniformly bounded second moments with respect to the original norm $\|\cdot\|_H$ of H. This is not true for the stronger norm $\|\cdot\|_R$ but nevertheless this estimate will be of use in the proof of the Fokker-Planck equation.

Corollary 7.3.4 *We have*

$$\sup_{k,\tau} \int_H \|x\|_H^2 \rho_k^\tau(dx) < \infty.$$

PROOF Fix some v_t among the reference measures and let Σ be an optimal coupling between ρ_k^τ and ν_t with respect to the norm $\|\cdot\|_H$. Then we have

$$\int_H \|x\|_H^2 \rho_k^\tau(dx) = \int_H \|x-y+y\|^2 \Sigma(dx,dy)$$
$$\leq 2d_H^2(\rho_k^\tau,\nu_t) + 2\int_H \|y\|_H^2 \nu_t(dy) \leq 2\|\sqrt{R}\|_{\mathcal{L}(H)}^2 d_R^2(\rho_k^\tau,\nu_t) + 2\int_H \|y\|_H^2 \nu_t(dy),$$

which is uniformly bounded by Proposition 7.3.2. Note that t is fixed and that ν_t as a Gaussian measure with trace class covariance has finite second $\|\cdot\|_H$ moment. Moreover we could apply Lemma 7.2.3 since $\|\cdot\|_H \leq \|\sqrt{R}\|_{\mathcal{L}(H)}\|\cdot\|_R$.■

7.4 Convergence to the Solution of the Fokker-Planck Equation

In this section we will construct a continuous function $t \mapsto \rho(t)$ from our discrete approximations and show that it solves a Fokker-Planck equation.

The first lemma sheds some light on the regularity of the approximating curves ρ^τ in dependence on the time step τ. With its help we will be able to prove the continuity of the limit curve in the next proposition. Moreover, it will allow for essential estimates in the proof of the Fokker-Planck equation.

Lemma 7.4.1 *For any fixed starting point $\rho_0 \in D(\mathrm{Ent}_{\nu_0})$ and any fixed time horizon T there is a constant $C = C(\rho_0, T) > 0$ such that for any $\tau > 0$ we have:*

$$\frac{1}{2\tau}\sum_{k=0}^{N^\tau} d_R^2(\rho_{k+1}^\tau, \rho_k^\tau) \leq C$$

PROOF By the minimization property of $\rho_{k+1}^\tau = \operatorname{argmin} \frac{1}{2\tau} d_R^2(\rho_k^\tau,\cdot) + \mathrm{Ent}_{\nu_{k\tau}}(\cdot)$ we can estimate the Wasserstein distance between two consecutive measures by the decrease in the current entropy:

$$\begin{aligned}
&\frac{1}{2\tau}\sum_{k=0}^{N^\tau} d_R^2(\rho_{k+1}^\tau,\rho_k^\tau)\\
&\leq \sum_{k=0}^{N^\tau} \mathrm{Ent}_{\nu_{\tau k}}(\rho_k^\tau) - \mathrm{Ent}_{\nu_{\tau k}}(\rho_{k+1}^\tau)\\
&= \sum_{k=0}^{N^\tau} \mathrm{Ent}_{\nu_{(k+1)\tau}}(\rho_{k+1}^\tau) - \mathrm{Ent}_{\nu_{\tau k}}(\rho_{k+1}^\tau) + \mathrm{Ent}_{\nu_0}(\rho_0)\underbrace{- \mathrm{Ent}_{\nu_{N\tau}}(\rho_{N^\tau+1})}_{\leq 0}\\
&\leq \sum_{k=0}^{N^\tau}\int_H \left(\log\left(\frac{d\rho_{k+1}^\tau}{d\nu_{(k+1)\tau}}(x)\right) - \log\left(\frac{d\rho_{k+1}^\tau}{d\nu_{k\tau}}\right)(x)\right)\rho_{k+1}^\tau(dx) + \mathrm{Ent}_{\nu_0}(\rho_0)\\
&= \sum_{k=0}^{N^\tau}\int_H \log\left(\frac{d\nu_{k\tau}}{d\nu_{(k+1)\tau}}(x)\right)\rho_{k+1}^\tau(dx) + \mathrm{Ent}_{\nu_0}(\rho_0)\\
&\leq \sum_{k=0}^{N^\tau}\int_H \log\left(\exp(K_2|(k+1)\tau - k\tau|\|x\|^2) f(k\tau,(k+1)\tau)\right)\rho_{k+1}^\tau(dx) + \mathrm{Ent}_{\nu_0}(\rho_0)\\
&\leq \sum_{k=0}^{N^\tau} K_2\tau\int_H \|x\|^2\rho_{k+1}^\tau(dx) + \sum_{k=0}^{N^\tau}\log(f(k\tau,(k+1)\tau)) + \mathrm{Ent}_{\nu_0}(\rho_0)\\
&\leq K_2 T\sup_k \int_H \|x\|^2\rho_{k+1}^\tau(dx) + \log(f(0,(N^{\tau+1)}\tau)) + \mathrm{Ent}_{\nu_0}(\rho_0)\\
&< \infty
\end{aligned}$$

where we used in the fourth line that

$$\frac{d\rho_{k+1}^\tau}{d\nu_{k\tau}}\frac{d\nu_{k\tau}}{d\nu_{(k+1)\tau}} = \frac{d\rho_{k+1}}{d\nu_{(k+1)\tau}} \quad \nu_{(k+1)\tau}a.e.$$

since $\rho_{k+1} \ll \nu_{k\tau} \ll \nu_{(k+1)\tau}$. See also the proof of Proposition 7.2.13 for a more careful justification of this step. In the fifth line we used Assumption 7.1.4 (ii). The supremum in the last expression is finite by Corollary 7.3.4. ■

Let us define a piece-wisely constant curve of measures in $\mathcal{P}(H)$

$$\rho^\tau(t) := \sum_{k=0}^{N}\rho_{k+1}^\tau(dx)1_{[k\tau,(k+1)\tau)}(t).$$

In the next proposition, we will show, that the above approximations converge to a curve of measures which is continuous in the Wasserstein metric. Note that we need to prove that the non-continuous approximations ρ^τ converge to a continuous limit. The proof is based on Proposition 3.3.1 from [3].

Proposition 7.4.2 *There is a sequence $\tau_n \to 0$ and a continuous curve $t \mapsto \rho(t)$ in $(\mathcal{P}(H), d_R)$ such that for every $t \in [0,T]$, $\rho^{\tau_n}(t) \to \rho(t)$ weakly as $\tau_n \to 0$.*

PROOF By Corollary 7.3.3 for every $t \in [0,T]$, $\tau > 0$ we have $\rho^\tau(t) \in K$ where K is a compact set in the weak topology. Hence for every t fixed, $\rho^\tau(t)$ admits a convergent subsequence. By a diagonal argument, we can find a common subsequence τ_n such that $\rho^{\tau_n}(t)$ converges for every $t \in [0,T] \cap \mathbb{Q}$. Thus we obtain a limit function $\rho : [0,T] \cap \mathbb{Q} \to K$.

In order to show that this function is continuous we need the following two claims. The first estimate is a condition of equicontinuity in the limit.
claim: with the constant C from Lemma 7.4.1 for every $0 \le s < t \le T$ we have

$$\limsup_{\tau \to 0} \, d_R^2(\rho^\tau(s), \rho^\tau(t)) \le |t-s| 2C.$$

Let k_τ and l_τ be the indices such that $k_\tau \tau \le s < (k_\tau + 1)\tau$ and $l_\tau \tau \le t < (l_\tau + 1)\tau$. Then using the triangle inequality, the Jensen inequality and Lemma 7.4.1 we have:

$$\begin{aligned}
\limsup_{\tau \to 0} \, d_R^2(\rho^\tau(s), \rho^\tau(t)) &\le \limsup_{\tau \to 0} \left(\sum_{j=k_\tau}^{l_\tau - 1} d_R(\rho^\tau(s), \rho^\tau(t)) \right)^2 \\
&\le \limsup_{\tau \to 0} \, (l_\tau - k_\tau) \sum_{j=k_\tau}^{l_\tau - 1} d_R^2(\rho^\tau(s), \rho^\tau(t)) \\
&\le \limsup_{\tau \to 0} \, (l_\tau - k_\tau) \tau 2C \\
&\le \limsup_{\tau \to 0} \, (t - s + \tau) 2C \\
&= (t-s) 2C
\end{aligned}$$

and the claim is proved.

The next claim states that the Wasserstein distance d_R is lower semicontinuous with respect to the weak topology.

claim: If $\mu_n \to \mu$ and $\nu_n \to \nu$ weakly as $n \to \infty$, then

$$\liminf_{n \to \infty} d_R(\mu_n, \nu_n) \ge d_R(\mu, \nu).$$

Let Σ_n be an optimal coupling between μ_n and ν_n, that is we have

$$d_R^2(\mu_n \nu_n) = \int_{H \times H} \|x - y\|_R^2 \Sigma_n(dx, dy).$$

Let us show that the sequence $(\Sigma_n)_{n \in \mathbb{N}}$ of measures on $H \times H$ is tight.
Since the convergent sequence μ_n is tight, there is $K_1 \subset H$ compact with $\inf_n \mu_n(K_1) \ge 1 - \varepsilon$. Let K_2 be a respective compact for the measures ν_n with $\inf_n \nu_n(K_2) \ge 1 - \varepsilon$. Then we have $\inf_n \Sigma_n(K_1 \times \mathcal{K}_2) \ge 1 - 2\varepsilon$. Assume there is some $n \in \mathbb{N}$ with $\Sigma_n(K_1 \times \mathcal{K}_2) < 1 - 2\varepsilon$.
Since $(K_1 \times K_2)^c \subset (K_1^c \times H) \cup (H \times K_2^c)$ we must have either $\Sigma_n(K_1^c \times H) = \mu_n(K_1) > \varepsilon$ or $\Sigma_n(H \times K_2^c) = \nu_n(K_2) > \varepsilon$. Both lead to a contradiction and

hence Σ_n admits a weakly convergent subsequence with limit say Σ. Since we have for any $f \in C_b(H)$:

$$\int_H f(x)\mu(dx) = \lim_{n\to\infty} \int_H f(x)\mu_n(dx)$$
$$= \lim_{n\to\infty} \int_{H\times H} f(x)\Sigma_n(dx, dy) = \int_{H\times H} f(x)\Sigma(dx, dy),$$

it is clear that Σ is a coupling of μ and ν, whence the first inequality in the following display:

$$\begin{aligned} d_R^2(\mu,\nu) \le \int_{H\times H} \|x-y\|_R^2 \Sigma(dx,dy) &= \sup_k \int_{H\times H} (\|x-y\|_R^2 \wedge k)\Sigma(dx,dy) \\ &= \sup_k \liminf_{n\to\infty} \int_{H\times H} (\|x-y\|_R^2 \wedge k)\Sigma_n(dx,dy) \\ &\le \liminf_{n\to\infty} \int_{H\times H} \|x-y\|_R^2 \Sigma_n(dx,dy) \\ &= \liminf_{n\to\infty} d_R^2(\mu_n,\nu_n) \end{aligned}$$

and the claim is proved.

We have for $s, t \in [0,T] \cap \mathbb{Q}$ by using the two claims:

$$\begin{aligned} d_R(\rho(s),\rho(t)) &\le \liminf_{n\to\infty} \; d_R(\rho^{\tau_n}(s),\rho^{\tau_n}(t)) \\ &\le \limsup_{n\to\infty} \; d_R(\rho^{\tau_n}(s),\rho^{\tau_n}(t)) \le \sqrt{2c|t-s|}. \end{aligned} \tag{7.7}$$

Hence, we can extend the function $\rho : [0,T] \cap \mathbb{Q} \to K$ continuously to all of $[0,T]$ as follows. For any $t \in [0,T] \setminus \mathbb{Q}$ let t_n be a sequence of rational numbers tending to t. Then by equation (7.7) the sequence $\rho(t_n)$ is d_R-Cauchy and we define $\rho(t)$ to be its limit. (Remember that $(\mathcal{P}(H), d_R)$ is complete according to Proposition 7.2.2.) Moreover $\rho(t) \in K$ as it is weakly closed and convergence in Wasserstein distance induces weak convergence.

Now we have to show that $\rho^{\tau_n}(t) \to \rho(t)$ for all $t \in [0,T]$ and the same subsequence as chosen above. For this it suffices to show that $\rho(t)$ is the only accumulation point of the subsequence $\rho^{\tau_n}(t)$. Thus, let μ be a cluster point of $\rho^{\tau_n}(t)$ along a subsequence τ_{n_k}. Then for $s \in [0,T] \cap \mathbb{Q}$:

$$d(\rho(s),\mu) \le \liminf_{k\to\infty} d(\rho^{\tau_{n_k}}(s),\rho^{\tau_{n_k}}(t)) \le \sqrt{|t-s|2C}.$$

Letting $s \to t$ along the rationals we obtain $\mu = \rho(t)$. ■

The following Proposition is Proposition 10.16 from [32] where the autonomous case is discussed. It is a key result in order to prove the next theorem, which is a generalization of Theorem 10.17 from [32]. Moreover its proof contains most of the intuition as to why the entropy gradient flow solves the Fokker-Planck equation since it relates the generator of the Ornstein-Uhlenbeck equation to the resolvent operator governing the entropy minimizing movement.

Proposition 7.4.3 *Fix $t > 0$. Let $\rho \in D(\mathrm{Ent}_{\nu_t})$ and let $p(dx, dy)$ be an optimal coupling between ρ and $J_t^\tau \rho$. Then we have for $f : H \to \mathbb{R}$ which are linear combinations of real and imaginary parts of functions of the form $x \mapsto \exp(i\langle x, h)$ for some $h \in D(A^*)$*

$$\frac{1}{\tau}\int_{H\times H} \langle \nabla_R f(y), x - y\rangle_R p(dx, dy) = 2\int_H G_t f(x) J_t^\tau \rho(dx),$$

where $\langle x, y\rangle_R := \langle R^{-\frac{1}{2}}x, R^{-\frac{1}{2}}y\rangle_H$ is the pseudo scalar product induced by the pseudo norm $\|\cdot\|_R$. Here, $R^{-\frac{1}{2}}$ is the pseudo inverse of $R^{\frac{1}{2}}$. Recall that G_t is the generator corresponding to the equation $dX_s = A(t)X_s ds + dW_R(s)$. On exponential test functions f it takes the form

$$G_t f(x) = \langle A^*(t)\nabla_x f(x), x\rangle + \frac{1}{2}\operatorname{Tr}(\sqrt{R}^*\nabla_{xx} f(x)\sqrt{R}).$$

In the following proof of the fact that our minimizing movements satisfy a Fokker-Planck equation we adopt a change of notation and write also $t \mapsto \rho_t$ instead of $t \mapsto \rho(t)$. In the same way we will also write $t \mapsto \rho_t^\tau$ instead of $t \mapsto \rho^\tau(t)$ for the discrete approximations.

Theorem 7.4.4 *Let $\rho_0 \in D(\mathrm{Ent}_{\nu_0})$ and let $t \mapsto \rho(t)$ be the entropy minimizing movement associated to $\{\mathrm{Ent}_{\nu_t}\}_{t\geq 0}$ with starting point ρ_0, that we obtained as a weak limit point of the approximating discrete flows ρ^τ in Proposition 7.4.2. It satisfies the Fokker-Planck equation:*

$$\partial_t \rho_t = -2G_t^* \rho_t$$

in the following weak sense: for all $\alpha \in \mathcal{C}_c([0, T])$ and all $f : H \to \mathbb{R}$ which are linear combinations of real and imaginary parts of functions of the form $x \mapsto \exp(i\langle x, h)$ for some $h \in D(A^*)$ we have

$$-\int_0^T \alpha'(t)\int_H f(x)\rho_t(dx)dt = \\ -\int_0^T \alpha(t)\int_H 2G_t f(x)\rho_t(dx)dt + \alpha(0)\int_H f(x)\rho_0(dx)$$

Proof We start out by rewriting the equation in terms of the approximating measure flows ρ^τ with additional error terms. Then we show that it converges to the equation for the limiting measure flow if we let τ tend to zero along a sequence such that $\rho^\tau \to \rho$. Let $p_k^\tau(dx, dy)$ be an optimal coupling between ρ_k^τ and ρ_{k+1}^τ, then:

$$
\begin{aligned}
&\int_0^T \alpha'(t) \int_H f(x)\rho_t^\tau(dx)dt \\
&= \sum_{k=0}^{N^\tau} [(\alpha(k+1)\tau) - \alpha(k\tau)] \int_H f(x)\rho_{k+1}^\tau(dx) \\
&= \sum_{k=0}^{N^\tau} \alpha(k\tau) \left(\int_H f(x)\rho_k^\tau(dx) - \int_{\mathbb{R}} f(x)\rho_{k+1}^\tau(dx) \right) - \alpha(0) \int_H f(x)\rho_0(dx) \\
&= \sum_{k=0}^{N^\tau} \alpha(k\tau) \left(\int_H \int_H [f(x) - f(y)] p_k^\tau(dx, dy) \right) - \alpha(0) \int_H f(x)\rho_0(dx)
\end{aligned}
$$

On the other hand, for p_k^τ as above, we have:

$$
\begin{aligned}
&\int_0^T \alpha(t) \int_H 2G_t f(x)\rho_t^\tau(dx)dt \\
&= \sum_{k=0}^{N^\tau} \int_{k\tau}^{(k+1)\tau} \alpha(t) \int_H 2G_t f(x)\rho_{k+1}^\tau(dx)dt \\
&= \sum_{k=0}^{N^\tau} \int_{k\tau}^{(k+1)\tau} \alpha(t) \int_H (2G_t - 2G_{k\tau}) f(x)\rho_{k+1}^\tau(dx)dt \\
&\quad + \sum_{k=0}^{N^\tau} \int_{k\tau}^{(k+1)\tau} \alpha(t)dt \int_H 2G_{k\tau} f(x)\rho_{k+1}^\tau(dx) \\
&= \sum_{k=0}^{N^\tau} \int_{k\tau}^{(k+1)\tau} \alpha(t) \int_H (2G_t - 2G_{k\tau}) f(x)\rho_{k+1}^\tau(dx)dt \\
&\quad + \sum_{k=0}^{N^\tau} \int_{k\tau}^{(k+1)\tau} \alpha(t)dt \frac{1}{\tau} \int_{H\times H} \langle \nabla_R f(y), x - y\rangle_R p_k^\tau(dx, dy).
\end{aligned}
$$

where we used Proposition 7.4.3 in the last equation. Combining the last two identities, we obtain:

$$\begin{aligned}
&\int_0^T \alpha'(t)\int_H f(x)\rho_t^\tau(dx)dt - \int_0^T \alpha(t)\int_H 2G_t f(x)\rho_t^\tau(dx)dt + \alpha(0)\int_H f(x)\rho_0(dx)\\
&= -\sum_{k=0}^{N^\tau}\int_{k\tau}^{(k+1)\tau}\alpha(t)\int_H (2G_t - 2G_{k\tau})f(x)\rho_{k+1}^\tau(dx)dt\\
&\quad + \sum_{k=0}^{N^\tau}\alpha(k\tau)\left(\int_{H\times H}(f(x)-f(y)-\langle\nabla_R f(y), x-y\rangle_R)\,p_k^\tau(dx,dy)\right) \qquad (7.8)\\
&\quad + \sum_{k=0}^{N^\tau}\left(\frac{1}{\tau}\int_{k\tau}^{(k+1)\tau}\alpha(t)dt - \alpha(k\tau)\right)\int_{H\times H}\langle\nabla_R f(y), x-y\rangle_R p_k^\tau(dx,dy)\\
&= I_1 + I_2 + I_3.
\end{aligned}$$

We will show that all three summands go to 0 as $\tau \to 0$. For I_1 remember that we have $G_t f(x) = \langle\nabla_x f(x), A(t)x\rangle + \frac{1}{2}\operatorname{Tr}(\sqrt{R}^*\nabla_{xx}f(x)\sqrt{R})$. If $f(x) = e^{i\langle x,h\rangle}$ for some $h \in D(A^*)$, then $t \mapsto \langle A(t)x, \nabla f(x)\rangle = \langle x, A^*(t)h\rangle i f(x)$ is continuous by strong continuity of $t \mapsto A^*(t)$. Thus:

$$\begin{aligned}
&\sum_{k=0}^{N^\tau}\int_{k\tau}^{(k+1)\tau}\alpha(t)\int_H (2G_t - 2G_{k\tau})f(x)\rho_{k+1}^\tau(dx)dt\\
&\le 2\sup_{x\in H}|f(x)| \sup_{k\le N^\tau}\sup_{t\in[k\tau,(k+1)\tau]}\|(A^*(t)-A^*(k\tau))h\| \sum_{k=0}^{N^\tau}\tau\|\alpha\|_\infty \int_H \|x\|\rho_{k+1}^\tau.
\end{aligned}$$

Now, $t \mapsto A^*(t)h$ is even uniformly continuous on $[0,T]$. Moreover, the second (and thus also the first) moment of the ρ_k^τ is uniformly bounded by Corollary 7.3.4 and f is bounded by definition of the test functions. Thus, this expression goes to 0.

For I_2 we use the Taylor formula for f to obtain:

$$\begin{aligned}
I_2 &\le \|\alpha\|_\infty \|\nabla_R^2 f\|_\infty \sum_{k=0}^{N^\tau}\int_{H\times H}\|x-y\|_R^2 p_k^\tau(dx,dy)\\
&= \|\alpha\|_\infty \|\nabla_R^2 f\|_\infty \sum_{k=0}^{N^\tau} d_R^2(\rho_{k+1}^\tau, \rho_k^\tau) = \mathcal{O}(\tau)
\end{aligned}$$

by Lemma 7.4.1. Finally, since we have $|\frac{1}{\tau}\int_{k\tau}^{(k+1)\tau}\alpha(t)dt - \alpha(k\tau)| \le \|\alpha'\|_\infty\tau$ by the mean value theorem, we obtain

$$\begin{aligned}
I_3 &\leq \tau \|\alpha'\|_\infty \sum_{k=0}^{N^\tau} \int_{H\times H} \langle \nabla_R f(y), x-y\rangle_R p_k^\tau(dx,dy) \\
&\leq \tau \|\alpha'\|_\infty \|\nabla_R f\|_\infty (N^\tau+1) \left(\sum_{k=0}^{N^\tau} \int_{H\times H} \|x-y\|_R \frac{1}{N^\tau+1} p_k^\tau(dx,dy) \right)^{2^{\frac{1}{2}}} \\
&\leq \tau \|\alpha'\|_\infty \|\nabla_R f\|_\infty (N^\tau+1) \left(\sum_{k=0}^{N^\tau} \int_{H\times H} \|x-y\|_R^2 \frac{1}{N^\tau+1} p_k^\tau(dx,dy) \right)^{\frac{1}{2}} \\
&\leq \tau \|\alpha'\|_\infty \|\nabla_R f\|_\infty \sqrt{N^\tau+1} \left(\sum_{k=0}^{N^\tau} d^2(\rho_{k+1},\rho_k) \right)^{\frac{1}{2}} \\
&= \mathcal{O}(\tau)
\end{aligned}$$

again by Lemma 7.4.1 and since $N^\tau = \mathcal{O}(\tau)$.

Having proven that the right hand side of (7.8) goes to zero we will now investigate the limits of the terms on the left hand side, thereby finding the Fokker-Planck equation. For the first term we easily have:

$$\int_0^T \alpha'(t) \int_H f(x) \rho_t^\tau(dx) dt \overset{\tau\to 0}{\longrightarrow} \int_0^T \alpha'(t) \int_H f(x)\rho_t(dx)dt$$

by weak convergence of $\rho_t^\tau \to \rho_t$ for each fixed t in connection with the dominated convergence theorem and the fact that α' as well as f is bounded.

In order to make use of weak convergence for the other term, let us show that $(t,x) \mapsto \alpha(t) G_t f(x)$ is continuous. Clearly, it is enough to show that $(t,x) \mapsto \langle \nabla_x f(x), A(t)x\rangle$ is continuous. For $f(x) = \exp(i\langle x,h\rangle)$ we calculate

$$\begin{aligned}
&|if(x)\langle x, A^*(t)h\rangle - if(y)\langle y, A^*(s)h\rangle| \\
&\leq |f(x)||\langle x,(A^*(t){-}A^*(s))h\rangle| {+} |f(y)|\,|\langle y{-}x, A^*(s)h\rangle| {+} |f(y){-}f(x)|\,|\langle x, A^*(s)h\rangle|.
\end{aligned}$$

Now, if $x \to y$ and $t \to s$ all three summands tend to zero by strong continuity of $t \mapsto A^*(t)$ and continuity of $x \mapsto f(x)$.

Thus $(t,x) \mapsto G_t f(x)$ is continuous, yet it is not bounded so we have to use a cut-off argument to make use of weak convergence.

We have $G_t f(x) = \langle \nabla_x f(x), A(t)x\rangle + \frac{1}{2}\operatorname{Tr}(\sqrt{R}^* \nabla_{xx} f(x)\sqrt{R})$. Since f is a test function all of its derivatives are bounded, so we only have to worry about the term of order $\|x\|$. Hence, choose a continuous cut-off function $0 \leq \chi_M \leq 1$ which is 1 if $\|x\| \leq M$ and 0 if $\|x\| \geq 2M$. Then

$$
\begin{aligned}
&\int_0^T \alpha(t) \int_H 2\langle A(t)x, \nabla_x f(x)\rangle (1-\chi_M(x))\rho^\tau(dx,dt) \\
&= \sum_{k=0}^{N^\tau} \int_{k\tau}^{(k+1)\tau} \alpha(t)dt \int_H 2\langle x, A^*(k\tau)h\rangle if(x)(1-\chi_M(x))\rho_{k+1}^\tau(dx) \\
&\leq 2\|\alpha\|_\infty \tau \|f\|_\infty \sup_{t\in[0,T]} \|A^*(t)h\| \sum_{k=0}^{N^\tau} \int_{\|x\|\geq M} \|x\|\rho_{k+1}^\tau(dx) \overset{M\to\infty}{\longrightarrow} 0\,.
\end{aligned}
$$

as we have

$$
\int_{\|x\|\geq M} \|x\|\rho_k^\tau(dx) \leq \int_H \frac{1}{M}\|x\|^2 \rho_k^\tau(dx) \overset{M\to\infty}{\longrightarrow} 0
$$

since the second moments of the measures ρ_k are uniformly bounded by Corollary 7.3.4. Knowing this, we can prove (setting $\Delta_R := \frac{1}{2}\operatorname{Tr}(\sqrt{R}^* \nabla_{xx} f(x)\sqrt{R})$

$$
\lim_{\tau\to 0} \underbrace{\int_{[0,T]\times H} \alpha(t) G_t f(x)\rho_t^\tau(dx)dt}_{I_\tau} = \underbrace{\int_{[0,T]\times H} \alpha(t) G_t f(x)\rho_t(dx)dt}_{I}
$$

by showing $\limsup_{\tau\to 0} I_\tau \leq I \leq \liminf_{\tau\to 0} I_\tau$.

$$
\begin{aligned}
&\limsup_{\tau\to 0} \int_{[0,T]\times H} \alpha(t) G_t f(x)\rho_t^\tau(dx)dt \\
&= \limsup_{M\to\infty} \limsup_{\tau\to 0} \Bigg(\int_{[0,T]\times H} \alpha(t)(\Delta_R f(x) + \langle A(t)x, \nabla f(x)\rangle \chi_M(x))\rho_t^\tau(dx)dt \\
&\qquad + \int_{[0,T]\times H} \langle A(t)x, \nabla f(x)\rangle (1-\chi_M(x))\rho_t^\tau(dx)dt \Bigg) \\
&\leq \limsup_{M\to\infty} \limsup_{\tau\to 0} \left(\int_{[0,T]\times H} \alpha(t)(\Delta_R f(x) + \langle A(t)x, \nabla f(x)\rangle \chi_M(x))\rho_t^\tau(dx)dt \right) \\
&\leq \limsup_{M\to\infty} \int_{[0,T]\times H} \alpha(t)(\Delta_R f(x) + \langle A(t)x, \nabla f(x)\rangle \chi_M(x))\rho_t(dx)dt \\
&\leq \int_{[0,T]\times H} \alpha(t)(\Delta_R f(x) + \langle A(t)x, \nabla f(x)\rangle)\rho_t(dx)dt \\
&= \int_{[0,T]\times H} \alpha(t) G_t f(x)\rho_t(dx)dt
\end{aligned}
$$

Here we could apply Fatou's Lemma because we have

$$
\sup_{t\in[0,T]} \int_H \|x\|^2 \rho_t(dx) < \infty
$$

as an easy consequence of the uniform moment bound on the approximating measures ρ_k. The argument for the lim inf is analogous and thus the proof is finished. ∎

7.5 Example

If we want to make use of our method in the infinite-dimensional case, our assumptions require amongst others that all reference measures are absolutely continuous with respect to each other. This forces us to impose severe restrictions on the time-dependent drift part of our Ornstein-Uhlenbeck equation. The following proposition gives a feasible choice where the drift is minus identity up to a (time-dependent) trace class operator.

Proposition 7.5.1 *In our setting let $(e_n)_n$ be an orthonormal basis of H diagonalizing the covariance matrix R of our Brownian noise. Let us denote its Eigenvalues by r_n. Let then $-A(t)$ be diagonalizable in the same basis with Eigenvalues $\lambda_n(t) = 1 + f_n(t)r_n$ where the real-valued functions f_n are smooth and bounded away from zero and infinity, that is we have a positive constant C such that:*

- $0 < \frac{1}{C} \leq f_n(t) \leq C < \infty$ *for all* $n \in \mathbb{N},\ t \in \mathbb{R}$,
- $\sup_{n,t} \left|\frac{d}{dt} f_n(t)\right| \leq C.$

Then Assumption 7.1.4 is fulfilled for the reference measures (ν_t).

Proof Let us denote by R_t the covariance operator of ν_t. We have $R_t = \int_0^\infty e^{sA(t)} R e^{sA^*(t)} ds$. Since all operators are diagonal in the same basis we obtain the Eigenvalues $r_n(t)$ of R_t as one-dimensional integrals, in a simple calculation:

$$R_t e_n = \int_0^\infty e^{sA(t)} R e^{sA^*(t)} e_n ds = r_n \int_0^\infty e^{2s\lambda_n(s)} e_n ds = \frac{1}{2} \frac{r_n}{1 + f_n(t) r_n} e_n.$$

Thus we have $r_n(t) = \frac{1}{2} \frac{r_n}{1+f_n(t)r_n}$.

For the calculation of the Entropy, note that we have the well-known property for product measures:

$$\mathrm{Ent}_{\otimes_{n=1}^N \mu_n}(\otimes_{n=1}^N \nu_n) = \sum_{n=1}^N \mathrm{Ent}_{\mu_n}(\nu_n).$$

By Lemma 7.5.3 this is still true in our infinite-dimensional setting so that we have:

$$\mathrm{Ent}_{\nu_s}(\nu_t) = \sum_{n=1}^\infty \mathrm{Ent}_{\mathcal{N}(0, r_n(s))}(\mathcal{N}(0, r_n(t))).$$

So we obtain, using Lemma 7.5.2 for the calculation of the one-dimensional entropies:

$$\begin{aligned}
\mathrm{Ent}_{\nu_s}(\nu_t) &= \frac{1}{2}\sum_{n=1}^{\infty}\left(\log\left(\frac{1+f_n(t)r_n}{1+f_n(s)r_n}\right)+\frac{1+f_n(t)r_n}{1+f_n(s)r_n}-1\right)\\
&= \frac{1}{2}\sum_{n=1}^{\infty}\left(\log\left(1+f_n(t)r_n\right)-\log\left(1+f_n(s)r_n\right)+\frac{(f_n(t)-f_n(s))r_n}{1+f_n(s)r_n}\right)\\
&\leq \frac{1}{2}\sum_{n=1}^{\infty}\left(f_n(t)r_n-f_n(s)r_n+\frac{(f_n(t)-f_n(s))r_n}{1+f_n(s)}\right)\\
&\leq C\,\mathrm{Tr}(R)<\infty,
\end{aligned}$$

where we have used boundedness and positivity of the f_n and the mean value theorem for the logarithm. Thus, Assumption 7.1.4 (i) is fulfilled.
In this setting we can also calculate the Radon-Nikodym derivatives $\frac{d\nu_t}{d\nu_s}$ according to Lemma 7.5.3.
Denoting $x_n = \langle x, e_n\rangle$, we obtain:

$$\begin{aligned}
\frac{d\nu_t}{d\nu_s}(x) &= \prod_{n=1}^{\infty}\exp\left(\frac{x_n^2}{2r_n(s)}-\frac{x_n^2}{2r_n(t)}\right)\sqrt{\frac{r_n(s)}{r_n(t)}}\\
&= \exp\left(\sum_{n=1}^{\infty}\frac{1}{2}x_n^2(f_n(t)-f_n(s))\right)\prod_{n=1}^{\infty}\sqrt{\frac{r_n(s)}{r_n(t)}}\\
&\leq \exp(C|t-s|\|x\|^2)\prod_{n=1}^{\infty}\sqrt{\frac{r_n(s)}{r_n(t)}}
\end{aligned}$$

where we used the mean value theorem and the uniform bound on the derivatives of the functions f_n.
Hence, Assumption 7.1.4 (ii) is fulfilled, since $f(s,t) := \prod_{n=1}^{\infty}\frac{r_n(s)}{r_n(t)}$ is multiplicative as required and convergence of the infinite product is seen easily.
Assumption 7.1.4 (iii) follows easily since the relevant operators commute and Assumption 7.1.4 (iv) holds since we have $r_n(t) \leq r_n$ for all $n \in \mathbb{N}$. ■

Lemma 7.5.2 *Let $\nu_1 := \nu_{\Sigma_1}$ and $\nu_2 := \nu_{\Sigma_2}$ be centered Gaussian measures on $\mathbb{R}^n$ with covariance matrices Σ_1 and Σ_2 respectively. Let μ be any measure absolutely continuous with respect to Lebesgue measure and admitting a second moment. Then we have*

$$(i)\,\mathrm{Ent}_{\nu_1}(\nu_2) = \frac{1}{2}\left(\log\left(\frac{\det(\Sigma_1)}{\det(\Sigma_2)}\right)+\mathrm{Tr}(\Sigma_2\Sigma_1^{-1})-n\right),$$

$$(ii)\,\mathrm{Ent}_{\nu_1}(\mu)-\mathrm{Ent}_{\nu_2}(\mu) = \frac{1}{2}\left(\log\left(\frac{\det(\Sigma_1)}{\det(\Sigma_2)}\right)+\int\langle x,(\Sigma_1^{-1}-\Sigma_2^{-1})x\rangle\mu(dx)\right).$$

PROOF We have

$$\mathrm{Ent}_{\nu_1}(\nu_2) = -\frac{1}{2}\int_{\mathbb{R}^n}\left(\langle(\Sigma_2^{-1}-\Sigma_1^{-1})x,x\rangle+\log\left(\sqrt{\frac{\det(\Sigma_1)}{\det(\Sigma_2)}}\right)\right)\nu_2(dx).$$

So we obtain for any orthonormal basis (e_k) of $\mathbb{R}^n$:

$$\begin{aligned}
&\int_{\mathbb{R}^n} -\frac{1}{2}\langle(\Sigma_2^{-1} - \Sigma_1^{-1})x, x\rangle \nu_2(dx) \\
&= -\frac{1}{2}\mathbb{E}_{\nu_2}\langle(\Sigma_2^{-1} - \Sigma_1^{-1})x, x\rangle \\
&= -\frac{1}{2}\sum_{k=1}^{n} \mathbb{E}_{\nu_2}\langle(x, (\Sigma_2^{-1} - \Sigma_1^{-1})^* e_k\rangle\langle x, e_k\rangle \\
&= -\frac{1}{2}\sum_{k=1}^{n} \langle(\Sigma_2^{-1} - \Sigma_1^{-1})^* e_k, \Sigma_2 e_k\rangle \\
&= -\frac{1}{2}\sum_{k=1}^{n} \langle e_k, \Sigma_1^{-1}\Sigma_2 e_k\rangle - \langle e_k, e_k\rangle \\
&= -\frac{1}{2}(\mathrm{Tr}(\Sigma_1^{-1}\Sigma_2) - n)
\end{aligned}$$

and (i) follows.
For (ii) we have

$$\mathrm{Ent}_{\nu_1}(\mu) - \mathrm{Ent}_{\nu_2}(\mu) = \int_{\mathbb{R}^n} \left[\log\left(\frac{d\mu}{d\nu_1}\right) - \log\left(\frac{d\mu}{d\nu_2}\right)\right] d\mu = \int \log\left(\frac{d\nu_2}{d\nu_1}\right) d\mu$$

and the result follows along the preceding calculations. ■

Lemma 7.5.3 (Radon-Nikodym derivative) *For $i = 1, 2$, let $\mu^i = \mathcal{N}(0, \Sigma^i)$ be Gaussian measures on H with covariance operators Σ^i. We assume that both Σ^1 and Σ^2 have the same eigenbasis $(e_n)_{n\in\mathbb{N}}$ and we denote the eigenvectors by σ_n^i respectively. Set $\mu_n^i := \mathcal{N}(0, \sigma_n^i)$, a normal distribution on $\mathbb{R}$, so that we have formally: $\mu^i = \otimes_{n\in\mathbb{N}}\mu_n^i$. If μ^1 and μ^2 are equivalent, then we can write the Radon-Nikodym derivative as a tensor product:*

$$\frac{d\mu^1}{d\mu^2}(x) = \prod_{n\in\mathbb{N}} \frac{d\mu_n^1}{d\mu_n^2}(\langle x, e_n\rangle) \qquad x \in H$$

Proof Let us denote

$$\mu^0(dx) := \prod_{n\in\mathbb{N}} \frac{d\mu_n^1}{d\mu_n^2}(\langle x, e_n\rangle)\mu^2(dx),$$

then we have to show that

$$\int_H F(x)\mu^1(dx) = \int_H F(x)\mu^0(dx) \tag{7.9}$$

holds for any measurable and bounded function F. By a monotone class argument it is sufficient to show it for functions F of the form

$$F(x) = f(\langle x, e_1\rangle, ..., \langle x, e_N\rangle), \tag{$*$}$$

where $N \in \mathbb{N}$ is arbitrary and $f : \mathbb{R}^N \to \mathbb{R}$ is a smooth function. The system formed by these functions is multiplicative and since its countable subset $(\langle x, e_N\rangle)_{N\in\mathbb{N}}$ separates the points of H the system also generates $\sigma(H)$.

Thus, let us take a function F as in $(*)$ and check (7.9):

$$\begin{aligned}
\int_H F(x)\mu^1(dx) &= \int_H f(\langle x, e_1\rangle, ..., \langle x, e_N\rangle)\mu^1(dx) \\
&= \int_{\mathbb{R}^N} f(x)\,(\mathcal{P}^N_\sharp \mu^1)\,(dx) \\
&= \int_{\mathbb{R}^N} f(x)\,(\otimes_{n=1}^N \mu_n^1)\,(dx) \\
&= \int_{\mathbb{R}^N} f(x_1, ..., x_N)\,\prod_{n=1}^N \left(\frac{d\mu_n^1}{d\mu_n^2}(x_n)\right)\,\mu_1^2(dx_1)...\mu_N^2(dx_N) \\
&= \int_{\mathbb{R}^N} f(x)\,(\mathcal{P}^N_\sharp \mu^0)\,(dx) = \int_H F(x)\mu^0(dx),
\end{aligned}$$

where $\mathcal{P}^N : H \to \mathbb{R}^N, \quad x \mapsto (\langle x, e_1\rangle, ..., \langle x, e_N\rangle)$. ∎

Bibliography

[1] ADAMS, S. ; DIRR, N. ; PELETIER, M. AND ZIMMER, J. : *Large deviations and gradient flows* - arXiv:1201.4601 [math.AP] 2012.

[2] ALBEVERIO, S. AND RÜDIGER, B. : *Stochastic Integrals and the Lévy-Ito Decomposition Theorem on Separable Banach Spaces* - Stochastic Analyis and Applications 23 p. 217-253, 2005.

[3] AMBROSIO, L. GIGLI, N. AND SAVARÉ, G. : *Gradient Flows in Metric Spaces and in the Space of Probability Measures* - Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005.

[4] APPLEBAUM, D. : *Lévy processes and stochastic calculus* - Cambridge University Press, 2004.

[5] APPLEBAUM, D. : *Martingale-valued measures, Ornstein-Uhlenbeck processes with jumps and operator self-decomposability in Hilbert space* - Séminaire de Probabilités 39 p. 171-196, 2006.

[6] ARENDT, W. in: *One-parameter Semigroups of Positive Operators* - Heidelberg : Springer, 1986.

[7] BOGACHEV, V. ; RÖCKNER, M. AND SCHUMLAND, B. : *Generalized Mehler semigroups and applications* - Probab. Theory Relat. Fields 105 p.193-225, 1996.

[8] BREZIS, H. : *Operateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert* - North-Holland Publishing Co., Amsterdam, 1973.

[9] MCCANN, R.J. : *A convexity principle for interacting gases* - Advances in Mathematics 128 p. 153-179, 1997.

[10] CHICONE, C. AND LATUSHKIN, Y. : *Evolution Semigroups in Dynamical Systems and Differential Equations* - Providence : American Mathematical Society , 1999.

[11] CHOJNOWSKA-MICHALIK, A. : *On processes of Ornstein-Uhlenbeck type in Hilbert space* - Stochastics 21 p. 251-286, 1987.

[12] Clement, P. : *Introduction to Gradient Flows in Metric Spaces (II)* - https://igk.nath.uni-bielefeld.de/study-materials/notes-clement-part2.pdf, 2010.

[13] Daleckii, Ju. L. and Krein, M. G. : *Stability of solutions of differential equations in Banach space* - Providence : American Mathematical Society, 1974.

[14] Da Prato, G. and Lunardi A. : *Ornstein-Uhlenbeck operators with time periodic coefficients* - Journal of Evolution Equations 7 p. 587-614, 2007.

[15] Da Prato, G. : *Kolmogorov Equations for Stochastic PDEs* - Birkhäuser, 2004.

[16] Da Prato, G. and Zabczyk, J. : *Ergodicity for infinite dimensional systems* - Cambridge University Press, 2003.

[17] Da Prato, G and Röckner, M : *A Note on Evolution Systems of Measures for Time-Dependent Stochastic Differential Equations* -Seminar on Stochastic Analysis, Random Fields and Applications V Progress in Probability 59, Part 1, p.115-122, 2008.

[18] Da Prato, G; Röckner, M. and Wang, F. :*Singular stochastic equations on Hilbert spaces: Harnack inequalities for their transition semigroups* Journal of Functinal Analysis 257 p. 992–1017, 2009.

[19] Da Prato, G. and Zabczyk, J. : *Stochastic Equations in Infinite Dimensions* - Cambridge University Press, 2008.

[20] Dynkin, E : *Sufficient Statistics and Extreme Points* - Ann. Probab. 5 p.705-730, 1978.

[21] Dudley, R. : *Real Analysis and Probability* - Wadsworth & Brooks/Cole, 1989.

[22] Fang, S. ; Shao, J. and Sturm, K.T. : *Wasserstein space over the Wiener space* - Probab. Theory Relat. Fields 146 p.535-565, 2010.

[23] Feyel, D and Üstünel, A.S. : *Monge-Kantorovitch Measure Transportation and Monge-Ampère Equation on Wiener Space* - Probab. Theory Relat. Fields 128 p.347-385, 2004.

[24] Friedman, A. : *Partial Differential Equations* - Holt, Reinhart and Winston, 1969.

[25] Fuhrmann, M. and Röckner, M. : *Generalized Mehler Semigroups: The Non-Gaussian Case* - Potential Analysis 12 p. 1-47, 2000.

[26] Geissert, M. and Lunardi, A. : *Invariant measures and maximal L^2 regularity for nonautonomous Ornstein–Uhlenbeck equations* - J. London Math. Soc. 77 p. 719-740, 2008.

[27] JORDAN, R. ; KINDERLEHRER, D. AND OTTO, F. : *The Variational Formulation of the Fokker-Planck Equation* - SIAM J. Math. Anal 29 p. 1-17, 1999.

[28] KNÄBLE, F. : *Ornstein-Uhlenbeck Equations with time-dependent coefficients and Lévy Noise in finite and infinite dimensions* - Journal of Evolution Equations 11 p. 959-993, 2011.

[29] LESCOT, P. AND RÖCKNER, M. : *Generators of Mehler-Type Semigroups as Pseudo-Differential Operators* - Infinite Dimensional Analysis, Quantum Probability and Related Topics 5 p. 297-316, 2002.

[30] LESCOT, P. AND RÖCKNER, M. : *Perturbations of Generalized Mehler Semigroups and Applications to Stochastic Heat Equations with Lévy Noise and Singular Drift* - Potential Analysis 20 p. 317-344, 2004.

[31] LINDE, W. : *Infinitely Divisible and Stable Measures on Banach Spaces* - Leipzig : Teubner, 1983.

[32] MAAS, J. : *Analysis of Infinite Dimensional Diffusions* - University of Delft, Ph.D.-Thesis , 2009.

[33] MA, Z. AND RÖCKNER, M. : *Introduction to the Theory of (Non-Symmetric) Dirichlet Forms* - Heidelberg : Springer, 1992.

[34] NEIDHARDT, H. AND ZAGREBNOV, V. : *Linear non-autonomous Cauchy problems and evolution semigroups* - arXiv:0711.0284v1 [math-ph] 2007.

[35] NICKEL, G. : *On evolution semigroups and wellposedness of nonautonomous Cauchy problems* - Ph.D. thesis, Tübingen: Univ. Tübingen, 1996.

[36] ORNSTEIN, L. AND UHLENBECK, G. : *On the theory of Brownian Motion* - Physical Review. 36 p.823-841, 1930.

[37] OUYANG, S. : *Harnack inequalities and applications for stochastic equations* - Ph.D. thesis, Bielefeld University, 2009.

[38] OUYANG, S. AND RÖCKNER, M. : *Non-homogeneous generalized Mehler semigroups and applications* - arXiv:1009.5314v2 [math.PR], 2010.

[39] PAZY, A. : *Semi-groups of linear operators and applications to partial differential equations* - College Park Maryland, 1974

[40] PESZAT, S. AND ZABCZYK, J. : *Stochastic Partial Differential Equations with Lévy Noise* - Cambridge University Press, 2007.

[41] PREVOT, C. AND RÖCKNER, M. : *A Concise Course on Stochastic Partial Differential Equations* - Heidelberg : Springer, 2007.

[42] RÖCKNER, M. AND WANG, F. : *Harnack and functional inequalities for generalized Mehler semigroups* - Journal of Functional Analysis 203 p. 237-261, 2003.

[43] SATO, K. : *Lévy processes and infinitely divisible distributions* - Cambridge University Press , 1999.

[44] SCHWARTZ, L. : *Radon measures on arbitrary topological spaces and cylindrical measures* - Oxford University Press , 1973.

[45] VERAAR, M. : *Non-autonomous stochastic evolution equations and applications to stochastic partial differential equations* Journal of Evolution Equations 10 p. 85–127, 2010.

[46] VILLANI, C. : *Optimal Transport, Old and New* - Springer, 2009.

[47] WOOSTER, R. : *Evolution Systems of Measures for Non-autonomous Ornstein-Uhlenbeck Processes with Lévy noise* - Communications on Stochastic Analysis 5 p.353-370, 2011.