Scuola Normale Superiore

Classe di Scienze Matematiche e Naturali
Corso di Perfezionamento (PhD) in Matematica

Well-posedness of Diffusion Processes in Metric Measure Spaces

by

Dario Trevisan

Prof. Luigi Ambrosio, Thesis Advisor

Prof. Maurizio Pratelli, Thesis Advisor

December 2014

Contents

Introduction

The aim of this thesis is the study of diffusion processes under low regularity and ellipticity assumptions, both on their coefficients and the ambient space. This is accomplished mainly by extending techniques and results from [Ambrosio and Trevisan, 2014], moving from the deterministic to the stochastic case. When specialized to different contexts, such as Euclidean or Gaussian spaces, large parts of previously known results, as well as novel ones, to the author's knowledge, are obtained at once. Moreover, as in [Ambrosio and Trevisan, 2014], our framework fits within that of Γ-calculus, developed e.g. in the monograph [Bakry et al., 2014], and is well-suited for the class of $\mathsf{RCD}(K, \infty)$ metric measure spaces, recently introduced by Ambrosio et al. [2014b] and object of extensive research.

The problem of well-posedness, i.e. existence, uniqueness (and stability) properties, for ordinary or stochastic differential equations, lies at very heart of many investigations in analysis and probability, both in finite and infinite dimensional spaces, aiming at going beyond the usual (Itô-)Cauchy-Lipschitz theories.

In the theory of ordinary differential equations (ODE's) in Euclidean spaces,

$$dX_t = b(t, X_t)dt, \quad t \in (0, T), \tag{0.1}$$

a major breakthrough is DiPerna-Lions theory, initiated in the seminal paper [DiPerna and Lions, 1989], which provides well-posedness, by means of a suitable notion of flow, for the equations associated to large classes of non-smooth vector fields, most notably that of Sobolev vector fields. More recently, Ambrosio [2004] extended the theory to include BV vector fields and, at the same time, he introduced a more probabilistic axiomatization based on the duality between flows and continuity equation, while the approach of DiPerna and Lions relied on characteristics and the transport equation.

In recent years the theory developed in many different directions, including larger classes of vector fields, quantitative convergence estimates, mild regularity properties of the flow, and non-Euclidean spaces, including infinite-dimensional ones. We refer to [Ambrosio and Crippa, 2008] for a more exhaustive, but still incomplete, description of the developments on this topic. In [Ambrosio and Trevisan, 2014], we extend the theory of well posedness for the continuity equation and the theory of flows to metric measure spaces $(E, d, \mathfrak{m})$. Roughly speaking, and obviously under additional structural assumptions, we prove that if $\{b(t, \cdot)\}_{t\in(0,T)}$ is a time-dependent family of Sobolev vector fields then there is a unique flow associated to b, namely a family of absolutely continuous maps $\{\mathbb{X}(\cdot, x)\}_{x\in E}$ from $[0, T]$ to E satisfying:

(i) $\mathbb{X}(\cdot, x)$ solves the ODE (0.1), with $X(0, x) = x$, for $\mathfrak{m}$-a.e. $x \in E$;

(ii) the push-forward measures $\mathbb{X}(t, \cdot)_{\#}\mathfrak{m}$ are absolutely continuous with respect to $\mathfrak{m}$ and have uniformly bounded densities.

Of course the notions of "Sobolev vector field" and even "vector field", as well as the notion of solution to the ODE have to be properly understood in this nonsmooth context.

Concerning the theory of stochastic differential equations (SDE's),

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad t \in (0, T), \tag{0.2}$$

Itô-Cauchy-Lipschitz theory was first outdone by the martingale approach, developed by Stroock and Varadhan [2006] (collecting results appeared in a series of articles from the late '60s). The crucial observation is that any solution to (0.2) induces many martingales: indeed, by choosing any sufficiently smooth function $f : \mathbb{R}^d \to \mathbb{R}$ and applying Itô formula, we obtain that

$$[0, T] \ni t \quad \mapsto \quad f(X_t) - \int_0^t (\mathcal{L}_s f)(X_s) ds$$

is a martingale with respect to the Brownian filtration, where we let

$$\mathcal{L}_t f(x) := \sum_{i=1}^{d} b^i(t, x) \frac{\partial f}{\partial x^i}(x) + \frac{1}{2} \sum_{i,j=1}^{d} a^{i,j}(t, x) \frac{\partial^2 f}{\partial x^i \partial x^j}(x), \quad \text{for } t \in (0, T), \tag{0.3}$$

be the associated diffusion (Kolmogorov) operator, and $a := \sigma\sigma^*$ is regarded as an infinitesimal covariance. The martingale problem consists in choosing precisely this property as a definition of solution to (0.2), and the main achievement by Stroock and Varadhan is that they provide well-posedness (in law) for equations associated to bounded measurable vector fields and uniformly bounded, continuous and elliptic a's.

Since then, the theory has been growing, due to its robustness and strong connections with the theory of semigroups and PDE's, also in abstract (metric) frameworks, see e.g. [Ethier and Kurtz, 1986]. In our thesis, we are concerned uniquely with martingale problems, but let us mention here that very active research lines deal with well-posedness for strong solutions to SDE's, i.e. closer to Itô's original approach to the problem of diffusions: besides the seminal paper [Veretennikov, 1980], we refer e.g. to [Krylov and Röckner, 2005] and [Da Prato et al., 2013] for recent striking results. Rigorous correspondences between the two descriptions are provided by the classical Yamada and Watanabe [1971] theorem and more recent extensions, see e.g. [Kurtz, 2007].

Figalli [2008] was first to develop a precise connection between DiPerna-Lions theory and martingale problems, providing in particular well-posedness for a wide class of diffusion whose associated operators $\mathcal{L}$ in (0.3) have not necessarily continuous or elliptic coefficients, provided that some Sobolev regularity holds. Of course, well-posedness, in particular uniqueness, has to be understood "in average" with respect to $\mathscr{L}^d$-a.e. initial condition. More precisely, a formalization akin to that of Ambrosio-DiPerna-Lions is introduced, the main objects being Stochastic Lagrangian Flows, i.e., Borel families $(\mathbb{P}_x)_{x \in \mathbb{R}^d}$ of probabilities on $C([0, T]; \mathbb{R}^d)$, satisfying:

(i) $\mathbb{P}_x$ solves the martingale problem associated to $\mathcal{L}$, starting from x, for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$;

(ii) the push-forward measures $(e_t)_\sharp \int \mathbb{P}_x \, d\mathscr{L}^d(x)$, where e_t is the evaluation map at $t \in [0, T]$, are absolutely continuous with respect to $\mathscr{L}^d$, with uniformly bounded densities.

Let us stress the fact that, as in the deterministic theory, uniqueness is understood for flows, thus in a selection sense: we are not claiming well-posedness for $\mathfrak{m}$-a.e. initial datum. Notice that these conditions might seem in perfect correspondence with the deterministic case sketched above, but Stochastic Lagrangian Flows are not necessarily (neither expected to be) deterministic maps of the initial point only; this is reflected also when the the operator $\mathcal{L}$ in (0.3) reduces to a derivation, i.e. when $a = 0$, and a solution to the martingale problem is any probability concentrated on possibly non-unique solutions to the ODE. Despite this discrepancy, the theory provides rather efficient tools to study SDE's under low regularity assumptions, in Euclidean spaces, and, together with [Le Bris and Lions, 2008], which deals with analogous issues from a PDE point of view, has become the starting point for further developments.

In this thesis we investigate well-posedness for diffusion processes, from the point of view of martingale problems, in the spirit of [Figalli, 2008] and in the setting of metric measure spaces, as introduced in [Ambrosio and Trevisan, 2014]. Motivations for performing such a study come at least from two sides. From an analytical one, the theory of "Riemannian" metric measure spaces mentioned in the first paragraph, currently under development, requires new calculus tools, and diffusion processes are a natural extension of flows and ODE's, also strictly connected with parabolic partial differential equations. From a probabilistic side, since the framework in [Ambrosio and Trevisan, 2014] is actually that of Dirichlet forms and Γ-calculus, it is almost compulsory to investigate whether the fruitful approach initiated therein provides results for classes of diffusion processes which lie beyond the scope of non-symmetric Dirichlet forms, as developed e.g. in [Ma and Röckner, 1992], [Stannat, 1999], or the results on well-posedness for singular diffusions developed e.g. in [Eberle, 1999]. In particular, our approach is naturally well-suited for degenerate diffusions, since it emerges as a generalization of the deterministic case.

Moreover, as in [Ambrosio and Trevisan, 2014], we are certain that the advantages of dealing with abstract frameworks are made clear by the wide range of settings to which our results can be then specialized: for the sake of brevity, in this thesis we mostly limit ourselves to a rigorous investigation of Euclidean and Gaussian cases, as chief examples of finite and infinite dimensional spaces. Another interesting feature of our approach is that some techniques, initially developed for the abstract framework, may provide a new point of view on the Euclidean case: as a first example we remark that our results non-trivially extend those from [Figalli, 2008]; a second one is that the so-called commutator estimate, which is certainly the technical core of the theory, seems to rely on a quite different identity from that in [DiPerna and Lions, 1989], as described in Section 11.2.2. This might even reveal some new features, and will be object of further investigations.

The thesis is organized in four parts, each one relying on the previous ones, although the fourth part, where we specialize the general theory, may be also read independently from the abstract development. The first part deals uniquely with diffusions in Euclidean spaces and should be considered a prelude to the following developments: we describe in a concise form the equivalences between "Eulerian" and "Lagrangian" descriptions of diffusions processes, i.e. respectively between Fokker-Planck equations and martingale problems, highlighting the role of the so-called superposition principle, similarly to the approach in [Figalli, 2008]. In the second part, we move from Euclidean to metric measure spaces and study the counterparts of these equivalences in this framework, after the right notions are introduced: diffusion operators, martingale problems, etc. In the third part, we address well-posedness for Fokker-Planck equations under suitable (but low) regularity assumptions both on the space and the diffusion operator, obtaining as a consequence of the well-posedness for martingale problems and associated flows. The fourth part is devoted to the specialization of the results thus obtained, so we return to Euclidean spaces and then consider the case of Gaussian, possibly infinite dimensional, spaces.

As stated in the very first paragraph, large parts of the techniques we employ are extensions to the case of diffusions of those originally introduced in [Ambrosio and Trevisan, 2014]: the first one is a general superposition principle for diffusion processes metric measure spaces that allows us to lift solutions to Fokker-Planck equations to measures on spaces of continuous functions, which solve correspondent martingale problems; the second one is an approach to commutator estimates with Γ-calculus tools, which in this case is extended to deal with diffusion operators. Let us also mention that, in the first part, we focus on a superposition principle for multidimensional diffusions processes, whose coefficients are not necessarily locally bounded or continuous: this result may be of independent interest, beyond its direct application to the study of diffusions in metric measure spaces, and we provide a self-contained exposition.

We now give a more accurate description of the four parts.

Part 1. In this part, we study multidimensional diffusion processes associated to an operator $\mathcal{L}$ as in (0.3), establishing abstract equivalences between well-posedness from different points of view, regardless of the fact that well-posedness actually holds, which may depend on various assumptions on the coefficients. We focus on the "Eulerian" description provided by Fokker-Planck (or forward Kolmogorov) equations

$$\partial_t \nu_t = (\mathcal{L}_t)^* \nu_t, \quad \text{in } (0,T) \times \mathbb{R}^d, \tag{0.4}$$

and the "Lagrangian" approach of martingale problems. We also introduce "martingale flows" on $[0,T] \times \mathbb{R}^d$ as selections of solutions to martingale problems, for $(s,x) \in [0,T] \times \mathbb{R}^d$.

In Chapter 1, we prove that well-posedness for Fokker-Planck equations in the class of curves of probability measures is equivalent to that of martingale problems, as well as existence and uniqueness for martingale flows, provided that a superposition principle holds.

In Chapter 2, we investigate the validity of a general superposition principle, showing that solutions to the Fokker-Planck equations (0.4) can be lifted to solutions to martingale problems in spaces of continuous paths, extending the results in [Figalli, 2008] to the case of diffusion operators with integrable coefficients. Although these results could be also investigated using tools e.g. from [Ethier and Kurtz, 1986], here we provide a complete and self-contained derivation, highlighting a general scheme of proof, which we may say to be classical, based on approximation, tightness and limit arguments. By iterations of this scheme, we pass from smooth to bounded coefficients, then to locally bounded and finally to the general case. A crucial tool is played by estimates on the modulus of continuity of solutions to martingale problems, based on fractional Sobolev energies: as in the case of ODE's, the size of the coefficients determines the regularity of paths, and in this case the processes exhibit Hölder continuity, of any order smaller than $1/2$, in case of bounded coefficients, which gets worse as they become larger.

Part 2. From Chapter 3 to Chapter 5 we introduce our abstract setup, which is the typical one of Γ-calculus and of the theory of Dirichlet forms: the distance is absent and we are given only a topology τ on X and a reference measure $\mathfrak{m}$ on X, which is required to be Borel, nonnegative and σ-finite. On $L^2(\mathfrak{m})$ we are given a symmetric, densely defined and strongly local Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ whose semigroup P is assumed to be Markovian. We write $\mathbb{V} := D(\mathcal{E})$ and assume that a carré du champ $\Gamma : \mathbb{V} \times \mathbb{V} \to L^1(\mathfrak{m})$ is defined, and that we are given a "nice" algebra $\mathscr{A}$ which plays the role of the C_c^∞ functions in the theory of distributions. Using $\mathscr{A}$, in Chapter 4, we define "vector fields" as derivations, as in [Ambrosio and Trevisan, 2014] which is in turn influenced by Weaver [2000] (and parallel developments in the theory of metric currents, [Ambrosio and Kirchheim, 2000]). A derivation $\boldsymbol{b}$ is a linear map from $\mathscr{A}$ to the space of real-valued Borel functions on X, $f \mapsto df(\boldsymbol{b})$, satisfying the Leibniz rule $d(fg)(\boldsymbol{b}) = fdg(\boldsymbol{b}) + gdf(\boldsymbol{b})$, and a pontwise $\mathfrak{m}$-a.e. bound in terms of Γ. Besides the basic example of gradient derivations, the carré du champ provides, by duality, a natural pointwise norm on derivations; such duality can be used to define, via integration by parts, a notion of divergence $\operatorname{div} \boldsymbol{b}$ for a derivation (the divergence depends only on $\mathfrak{m}$, not on Γ).

By introducing the bilinear counterparts of derivations, we are led to study 2-tensors, and finally define diffusion operators $\mathcal{L}$ as linear operators on $\mathscr{A}$, $f \mapsto \mathcal{L}f$, such that the associated carré du champ

$$(f, g) \quad \mapsto \quad \boldsymbol{a}(f, g) := \frac{1}{2}\left[\mathcal{L}(fg) - f\mathcal{L}(g) - g\mathcal{L}(f)\right]$$

defines a symmetric, non-negative, 2-tensor, i.e. $\boldsymbol{a}(f,g) = \boldsymbol{a}(g,f)$, $\boldsymbol{a}(f,f) \geq 0$, for f, $g \in \mathscr{A}$. As with derivations, we introduce a natural notion of divergence $\operatorname{div} \mathcal{L}$, whose negative part plays an important role in several quantitative estimates.

In chapters 6 and 7, we introduce Fokker-Planck equations,

$$\partial_t u = \mathcal{L}_t^* u_t, \quad \text{on } (0,T) \times X, \tag{0.5}$$

which specialize (0.4) in the case $\nu_t = u_t \mathfrak{m}$. We also define martingale problems, flows, and study their abstract equivalences, closely following the first part. In this case, however, we restrict the attention to well-posedness "in average" with respect to the measure $\mathfrak{m}$, in the sense of Ambrosio-DiPerna-Lions theory, as extended by Figalli. Let us remark that we prefer to introduce a slightly different notion than that of Stochastic Lagrangian Flows, by considering families of solutions to martingale problems for $\mathfrak{m}$-a.e. $x \in X$ and for every $s \in [0,T]$; we also prefer to refer to them as regular martingale flows, where the term "regular" stands for condition (ii) above regarding absolute continuity and bounds on densities, while the term "martingale" remarks that that the solution is understood in the sense of Stroock-Varadhan's martingale problem. After abstract equivalence is settled, we investigate the validity of a superposition principle for diffusions in metric measure spaces, by extending the approach in [Ambrosio and Trevisan, 2014], which in turn is heavily influenced by a change of variables technique appearing in the recent paper by Kolesnikov and Röckner [2014], although not in a Lagrangian perspective.

Part 3. We address in this part well-posedness results for solutions to Fokker-Planck equations. In Chapter 8, we formally describe the estimates that provide well-posedness, in elliptic or degenerate cases, and even in presence of Sobolev inequalities. As in DiPerna-Lions theory, a crucial role is played by the quantity $\operatorname{div} \mathcal{L}^-$, which governs the exponential rate of accumulation of mass, e.g. the estimate

$$\frac{d}{dt}\int |u_t|^2 \, d\mathfrak{m} \leq \left\|\operatorname{div} \mathcal{L}^-\right\|_{L^\infty(\mathfrak{m})} \int |u_t|^2 \, d\mathfrak{m}, \quad \mathscr{L}^1\text{-a.e. } t \in (0,T). \tag{0.6}$$

From this differential inequality, Gronwall lemma entails bounds in $L^\infty((0,T); L^2(\mathfrak{m}))$, as long as $u_0 \in L^2(\mathfrak{m})$; in particular, if $u_0 = 0$, we would obtain $u = 0$ and so uniqueness, simply arguing by linearity on the difference $u^1 - u^2$ between two solutions with the same initial datum.

Notably, if d-dimensional Sobolev inequalities hold and the diffusion is elliptic, we can drop (usual) L^∞-bounds in favour of L^d-bounds: ellipticity has the effect of spreading mass and prevents collapsing.

In Chapter 9 we prove existence of solutions. The strategy of the proof is classical: first we add a viscosity term and get a $\mathbb{V}$-valued solution by Hilbert space techniques, then we take a vanishing viscosity limit, exploiting the bound provided by (0.6). Together with existence, we recover also higher (or lower, since our measure $\mathfrak{m}$ might be not finite and therefore no inclusion between L^p spaces might hold) integrability estimates on solutions, depending on the initial condition. Also, under a suitable assumption on $\mathscr{A}$, we prove that the L^1 norm is independent of time.

Chapter 10 is devoted to the proof of uniqueness of solutions, which is the most delicate part of the theory; the literature on this subject is already vast and currently growing, with contributions coming both from analytic and probabilistic sides, on finite and infinite dimensional spaces: see e.g. [Jordan et al., 1998], [Lisini, 2009], [Natile et al., 2011], [Peletier et al., 2013], [Bogachev et al., 2002], [Bogachev et al., 2011], and references therein. The classical proof in [DiPerna and Lions, 1989] is based on a smoothing scheme that, in our context, is played by the semigroup P, or slight variants of it: this approach proved to be successful in [Ambrosio and Figalli, 2009] and [Trevisan, 2014a], in Wiener spaces.

Aiming at rigorously establishing (0.6) for any solution to the FPE (0.5), the main problem is that multiplication by u itself is an operation not allowed by the weak formulation of (0.5), which holds only in duality functions in $\mathscr{A}$. Therefore, we study the equation solved by a smooth approximation of u, namely $u^\alpha := \mathsf{P}_\alpha u$, which reads as

$$\partial_t u^\alpha = \mathcal{L}_t^* u_t^\alpha - r_t^\alpha, \quad \text{on } (0,T) \times X,$$

where the "error term" $r_t^\alpha := [\mathcal{L}_t^*, \mathsf{P}_\alpha] u_t$ is the commutator between P and the action of the diffusion operator $\mathcal{L}$. The approximated FPE above allows for establishing an approximated version of (0.6), which in the limit $\alpha \downarrow 0$, entails uniqueness, provided that the commutator is infinitesimal.

For for the sake of clarity, let us sketch here the strategy to deal with derivations, settled in [Ambrosio and Trevisan, 2014], since the general case is a variation on this theme (although non-trivial). By duality, we are reduced to study the commutator

$$[\mathsf{P}_\alpha, \boldsymbol{b}] f = \mathsf{P}_\alpha (df(\boldsymbol{b})) - d(\mathsf{P}_\alpha f)(\boldsymbol{b}), \qquad \text{for } f \in \mathscr{A},\ \alpha > 0,$$

and the main idea is to imitate Bakry-Émery's Γ-calculus (see the already quoted monograph [Bakry et al., 2014]), interpolating and writing, at least formally,

$$\begin{aligned}
[\mathsf{P}_\alpha, \boldsymbol{b}] f &= \int_0^\alpha \frac{d}{ds} \mathsf{P}_s (d(\mathsf{P}_{\alpha-s} f)(\boldsymbol{b})) ds \\
&= \int_0^\alpha \mathsf{P}_s \left[\Delta \left[d(\mathsf{P}_{\alpha-s} f)(\boldsymbol{b}) \right] - d(\Delta \mathsf{P}_{\alpha-s} f)(\boldsymbol{b}) \right] ds \\
&= \int_0^\alpha \mathsf{P}_s \left[\Delta, \boldsymbol{b} \right] \mathsf{P}_{\alpha-s} f ds.
\end{aligned} \tag{0.7}$$

It turns out that an estimate of the commutator involves only the symmetric part of the derivative (this, in the Euclidean case, was already observed by Capuzzo Dolcetta and Perthame

[1996] for regularizations induced by even convolution kernels). This structure can be recovered in our context: inspired by the definition of Hessian in [Bakry, 1994] we define the symmetric part $D^{sym}\boldsymbol{b}$ by

$$\int D^{sym}\boldsymbol{b}(u,f)d\mathfrak{m} := -\frac{1}{2}\int \left[du(\boldsymbol{b})\Delta f + df(\boldsymbol{b})\Delta u - (\operatorname{div}\boldsymbol{b})\Gamma(u,f)\right] d\mathfrak{m}.$$

Using this definition in (0.7) (assuming here for simplicity $\operatorname{div}\boldsymbol{b} = 0$) we establish the identity

$$\int u[\mathsf{P}_\alpha, \boldsymbol{b}]f d\mathfrak{m} = 2\int_0^\alpha \int D^{sym}\boldsymbol{b}(\mathsf{P}_s u, \mathsf{P}_{\alpha-s}f)d\mathfrak{m}\, ds, \qquad \text{for } u, f \in \mathscr{A}.$$

Then, we assume the validity of the estimates which, in a smooth context, amount to an L^q control ($q > 1$) on the symmetric part of derivative and some regularizing properties of the semigroup P, which hold assuming an abstract Ricci curvature lower bound on the underlying space, and in particular for Euclidean and Gaussian spaces, regardless of their dimension. These provide uniform bounds in α, which by standard density arguments lead to convergence towards 0 for the general commutator.

In case of diffusions, similar strategies can be employed, e.g. by interpolating up to a second order Taylor expansion in the possibly degenerate case, or computing the commutator between ∂_t and a time-dependent family of semigroups, in the elliptic case, similarly as in [Figalli, 2008]. Let us remark again, however, that our strategy seems to be well-suited for degenerate cases, while in the elliptic case the novel contribution becomes more modest, in particular if compared to the extensive studies available in the literature of stochastic analysis, e.g. the theory of non-symmetric Dirichlet forms quoted above. Apparently, however, the DiPerna-Lions approach still provides some new insight since, e.g., it allows to go beyond the L^2-framework and consider more general integrability assumptions.

Part 4. In this part, we mainly specialize the general results to Euclidean and Gaussian settings; for brevity, we prefer not to address the full family of spaces considered in [Ambrosio and Trevisan, 2014], although non-trivial results hold also for $\mathsf{RCD}(K,\infty)$ spaces, as we sketch in Chapter 13; we focus instead with greater detail in these two well-studied frameworks. In Chapter 11, we provide explicit computations that parallel those in the previous part, comparing our approach first with that of DiPerna-Lions, in the deterministic case, and then with Figalli's extension, in the general case. In Chapter 12, we specialize to infinite dimensional Gaussian frameworks, both with respect to Malliavin calculus on Wiener spaces and to Gaussian Hilbert spaces: we also include the main result from [Trevisan, 2014a], where well-posedness for continuity equations associated to vector fields with bounded variation in Wiener spaces is settled, marking an ideal connection with our Master's Thesis, centred around functions of bounded variations in Wiener spaces. Finally, in Chapter 13, we briefly report how curvature assumptions on the underlying geometry of the space, both from the "Eulerian" and the "Lagrangian" point of view, provide an abstract but rich enough structure, where non-trivial examples can be studied, at least in the deterministic case.

From the description above, it is clear that this thesis is based on, but significantly expands, the original contents in [Ambrosio and Trevisan, 2014] and, in the last chapter, results from [Trevisan, 2014a] also well fit within. In general, a PhD thesis should report on the whole research activity: for the sake of uniformity of exposition, we have not included other mathematical works, obtained during the candidate's *Perfezionamento* (PhD) course at Scuola

Normale Superiore. Below, we briefly mention them: they provide a wider picture about our interests on interactions between analysis and probability, in particular for the study of differential equations and analysis on metric measure spaces.

Functions of Bounded Variation on the Classical Wiener Space and an Extended Ocone-Karatzas Formula, [Pratelli and Trevisan, 2012]**.** This joint work is largely based on the candidate's Master's Degree thesis. We prove an extension of Ocone-Karatzas integral representation, roughly giving a function as the Itô integral of (the predictable projection of) its Malliavin derivative. We cover the case of BV functions on the classical Wiener space, i.e. when the distributional derivative is a measure; quite surprisingly, it turns out that the integral term in the formula is nevertheless a function, i.e. an absolutely continuous measure. We also establish an elementary chain rule formula and combine the two results to compute explicit integral representations for some classes of BV composite random variables.

BV**-regularity for the Malliavin derivative of the maximum of a Wiener process,** [Trevisan, 2013a]**.** Also this work is partially based on the candidate's Master's Degree thesis, although parts of it were settled during the first year of PhD studies. In Malliavin calculus, a well-known example of non-smooth random variable is the following one: let $(W_t)_{t\in[0,T]}$ be the real-valued Wiener process and let $M := \sup_{t\in[0,T]} W_t$ be its maximum. Then, M is differentiable in Malliavin's sense, and its derivative $D_t M$ is $I_{\{t<\sigma\}}$, where σ is the first time at which the maximum is hit. In this article, we prove that the second-order derivative $D_{s,t}M$ is a measure, with bounded variation, or equivalently, that the DM is a BV vector field on the classical Wiener space. The total variation measure $\left|D^2M\right|$ is a finite measure on the Wiener space which is singular with respect to Wiener's measure and we show that it is supported on paths which hit their maximum exactly twice.

Zero noise limits using local times, [Trevisan, 2013b] **.** We consider a well-known family of SDE's with irregular drifts and the correspondent zero noise limits, namely

$$dX_t^\varepsilon = \operatorname{sign}(X_t^\varepsilon)\,|X_t^\varepsilon|^\gamma\,dt + \varepsilon dW_t, \quad X_0^\varepsilon = 0,$$

for some $\gamma \in [0,1)$. Using mollified versions of local times, we show which trajectories are selected in the narrow limit $\varepsilon \downarrow 0$. Of course, it was already known that the selection concentrates on the extremal ones, but our approach is new and completely "Lagrangian", relying on elementary stochastic calculus only, in particular a careful analysis of the application of Itô-Tanaka formula to with $f(x) = |x|$, to show that noise and vector field interact positively, allowing for the solution to leave the origin.

Uncertainty and isoperimetric inequalities on groups and homogeneous spaces, [Dall'Ara and Trevisan, 2014]**.** In this joint work, we prove a family of uncertainty inequalities on fairly general groups and homogeneous spaces, both in the smooth and in the discrete framework. They hold on L^p, for $p \in [1,\infty)$, and the proof is based on a link between the L^1 endpoint and a general weak isoperimetric inequality.

A short proof of Stein's universal multiplier theorem, [Trevisan, 2014b] **.** In this note, we give a brief proof of Stein's universal multiplier theorem, purely by probabilistic methods, avoiding any use of harmonic analysis techniques (complex interpolation or transference methods). Stein's celebrated result, [Stein, 1970, Corollary IV.6.3], provides continuity bounds for a general family of operators related to a Markovian semigroup $(T^t)_{t\geq 0}$, virtually without any assumption on the underlying measure space (X, m). Our novel proof of this result relies only on Rota's martingale representation and Burkholder-Gundy inequalities.

Acknowledgements

I would like to deeply thank my supervisors L. Ambrosio and M. Pratelli, because of their constant support and suggestions during all of my PhD studies and in particular the writing of this thesis, and more generally because of their inspiring attitude towards mathematics and its challenges.

I express my gratitude towards G. Crippa, F. Flandoli and M. Novaga for discussions on continuity equations, flows and their stochastic counterparts, even at early stages of my interest in the subject; N. Gigli, G. Savaré and M. Erbar for useful and highlighting comments on the geometry of metric measure spaces. I am also grateful towards G. Da Prato and G. Letta for their series of courses in the years before their retirement, respectively at Scuola Normale Superiore and at Università di Pisa: their lectures profoundly influenced my interests towards this research area.

I warmly thank G. Curato, S. Calamai, S. Di Marino, M. Marini, S. Mongodi, B. Ruffini, B. Velichkov, B. Wróbel, G. Zanco and all the other colleagues with whom I shared part of my experience of PhD studies at Scuola Normale Superiore; special thanks go to M. Maurelli and G. Dall'Ara.

Lastly, I am thankful towards all the professors, researchers, post-docs, other workers and all the students at the institution of Scuola Normale Superiore, who crucially contributed to my mathematical education.

Part I

Diffusion processes in $\mathbb{R}^d$

Chapter 1

Equivalent descriptions for diffusion processes

In this chapter, we study the abstract correspondence between “Eulerian” and “Lagrangian” descriptions of multidimensional (i.e., in $\mathbb{R}^d$) diffusion processes, in particular with respect to existence and uniqueness issues. To this aim, in Section 1.1 we introduce diffusion operators in $\mathbb{R}^d$, as well as Fokker-Planck equations, martingale problems and flows and in Section 1.2 we study their equivalences.

The results appearing in this chapter cannot be considered novel, since they rely on arguments well established in the literature, particularly those contained in the seminal paper [Figalli, 2008], where many ideas from the deterministic setting of continuity equations and ODE’s, see e.g. the lecture notes [Ambrosio and Crippa, 2008], meet with others of more stochastic nature, classically developed in the monograph [Stroock and Varadhan, 2006]. However, differently from [Figalli, 2008], here we seek to provide abstract and full equivalences between various notions, emphasizing the crucial role played by the superposition principle for diffusions, whose validity is addressed in Chapter 2. We also remark that the arguments in Section 1.2 adapt almost verbatim to the setting of metric measure spaces, specifically in Chapter 6.

1.1 Definitions and basic facts

In this section we introduce measure-valued solutions to Fokker-Planck equations and martingale problems, associated to a diffusion operator in $\mathbb{R}^d$ ($d \geq 1$). For a brief historical account of these different approaches to describe multidimensional diffusion processes, we refer to the well-written Introduction in [Stroock and Varadhan, 2006].

We let throughout $\mathscr{A} = C^{1,2}_c((0,T)\times\mathbb{R}^d))$, i.e. $f \in \mathscr{A}$ if and only if f has compact support and is continuously differentiable once with respect to $t \in (0,T)$ and twice with respect to $x \in \mathbb{R}^d$: here and below, the superscript $(1,2)$ counts the number of derivatives with respect to (t,x). We endow $\mathscr{A}$ with the norm

$$\|f\|_{C^{1,2}} = \sup_{(0,T)\times\mathbb{R}^d} \left\{ |f| + |\partial_t f| + |\nabla f| + |\nabla^2 f| \right\}.$$

We also let

$$a = (a^{i,j})_{i,j=1}^d : (0,T) \times \mathbb{R}^d \to \mathrm{Sym}_+(\mathbb{R}^d), \quad b = (b^i)_{i=1}^d : (0,T) \times \mathbb{R}^d \to \mathbb{R}^d, \tag{1.1}$$

be respectively a Borel map, taking values in the space of symmetric, non-negative definite $n \times n$ matrices on $\mathbb{R}$ and a Borel time-dependent vector field on $\mathbb{R}^d$.

Definition 1.1 (diffusion operators in $\mathbb{R}^d$). *We say that the linear operator $\mathcal{L} = \mathcal{L}(a, b)$, given by*

$$\mathscr{A} \ni f \quad \mapsto \quad \mathcal{L}f : (t,x) \mapsto \sum_{i,j=1}^d a^{i,j}(t,x) \frac{\partial^2 f}{\partial x^i \partial x^j}(t,x) + \sum_{i=1}^d b^i(t,x) \frac{\partial f}{\partial x^i}(t,x)$$

is the diffusion operator with coefficients a, b.

The vector field b is sometimes referred as the drift (or infinitesimal mean) of $\mathcal{L}$ and the matrix valued map a as the infinitesimal covariance of $\mathcal{L}$. If $a = 0$, thus $\mathcal{L}$ reduces to a derivation, we also say that we are in the deterministic case.

For brevity, we introduce the notation, for vectors v, $w \in \mathbb{R}^d$ and matrices A, $B \in \mathbb{R}^{d\times d}$,

$$v \cdot w = \sum_{i=1}^d v^i w^i, \quad |v|^2 = v \cdot v, \quad (v \otimes w)^{i,j} := v^i w^j \quad \text{for } i, j \in \{1, \ldots d\}$$

$$A : B = \sum_{i,j=1}^d A^{i,j} B^{i,j}, \quad |A|^2 = A : A, \quad A(v, w) = A : (v \otimes w),$$

and the following standard notation for differential calculus in $\mathbb{R}^d$,

$$g(t, \cdot) = g_t(\cdot), \quad (\mathcal{L}f)_t = \mathcal{L}_t f, \quad \partial_t f = \frac{\partial f}{\partial t}, \quad \text{for } t \in (0, T),$$
$$\partial_i f = \frac{\partial f}{\partial x^i}, \quad \partial_{i,j} f = \frac{\partial^2 f}{\partial x^i \partial x^j}, \quad \text{for } i, j \in \{1, \ldots, d\},$$
$$\nabla f = (\partial_i f)_{i=1}^d, \quad \nabla^2 f = (\partial_{i,j} f)_{i,j=1}^d, \quad \text{thus}$$
$$b \cdot \nabla f = \sum_{i=1}^d b^i \partial_i f \quad \text{and} \quad a : \nabla^2 f = \sum_{i,j=1}^d a^{i,j} \partial_{i,j} f.$$

Given a Polish space (X, τ), we write $\mathscr{M}(X)$ for the space of signed (real-valued) Borel measures on X, whose total variation measure is finite, $\mathscr{M}^+(X) \subseteq \mathscr{M}(X)$ for the cone of finite non-negative measures and $\mathscr{P}(X) \subseteq \mathscr{M}^+(X)$ for the convex set of Borel probability measures. We say that a curve $\nu = (\nu_t)_{t\in(0,T)} \subseteq \mathscr{M}(X)$ is Borel if, for every Borel set $A \subseteq \mathbb{R}^d$, the map $t \mapsto \nu_t(A)$ is Borel. We let $|\nu| = (|\nu_t|)_{t\in(0,T)}$ be the curve of total variation measures: if ν is Borel, then $|\nu|$ is Borel as well. Since many of the bounds that appear below are integral with respect to the variable t, with respect to Lebesgue measure $\mathscr{L}^1$ restricted to $(0, T)$, if $\nu = (\nu_t)_{t\in(0,T)} \subseteq \mathscr{M}(\mathbb{R}^d)$ is Borel, with a slight abuse of notation we let $|\nu|$ also be the measure on $\mathbb{R}^d$ mapping $A \mapsto |\nu|(A) = \int_0^T |\nu_t|(A)dt$, for $A \subseteq \mathbb{R}^d$ Borel. When $\nu = (\nu_t)_t \subseteq \mathscr{M}^+(\mathbb{R}^d)$, we also let $|\nu| = \nu$, with the same abuse of notation. We say that a, $b \in L^p(|\nu|)$, for $p \in [1, \infty)$, if there holds

$$\int (|a|^p + |b|^p)\, d|\nu| \left(:= \int_0^T \int (|a_t|^p + |b_t|^p)\, d|\nu_t|\, dt \right) < \infty.$$

Similarly, we say that a, $b \in L^p_{loc}(|\nu|)$ if, for every compact set $B \subseteq \mathbb{R}^d$, it holds

$$\int_{(0,T)\times B} (|a|^p + |b|^p)\, d|\nu| < \infty.$$

Fokker-Planck equations

We consider weak solutions to Fokker-Planck equations, in duality with $\mathscr{A}$, and state some elementary properties, whose simple proof can be found at the beginning of [Ambrosio et al., 2008, §8.1], in the deterministic setting, i.e., when $a = 0$ and the Fokker-Planck equation reduces to the continuity equation.

Definition 1.2 (solutions to FPE's)**.** *Let a, b be Borel maps as in* (1.1)*. A Borel curve $\nu = (\nu_t)_{t\in(0,T)} \subseteq \mathscr{M}(\mathbb{R}^d)$ is a (weak) solution to the Fokker-Planck equation (FPE)*

$$\partial_t \nu_t = \mathcal{L}_t^* \nu_t, \quad \textit{in } (0,T)\times\mathbb{R}^d, \tag{1.2}$$

if a, $b \in L^1_{loc}(|\nu|)$ and it holds

$$\int_0^T \int \left[\partial_t f_t + \mathcal{L}_t f\right] d\nu_t \, dt = 0, \qquad \textit{for every } f \in \mathscr{A}. \tag{1.3}$$

Our main interest lies in solutions to FPE's that are curves of probability measures, but general measure valued solutions are useful, e.g. since linearity allows for adding or subtracting solutions.

Remark 1.3 (extension of the weak formulation)**.** If a, $b \in L^1(|\nu|)$, the validity of (1.3) can be easily extended to any $f \in C_b^{1,2}((0,T)\times\mathbb{R}^d)$ whose support has compact projection on $(0,T)$, arguing as in [Ambrosio et al., 2008, Remark 8.1.1]. Since the proof of this fact requires the introduction of useful cut-off functions, we sketch it here. For $R \geq 1$, let $\chi_R : \mathbb{R}^d \to [0,1]$ be a smooth function with $\chi_R(x) = 1$, for $|x| \leq R$, $\chi_R(x) = 0$, for $|x| \geq 2R$, with the uniform bounds $|\nabla \chi_R| \leq 4R^{-1}$ and $\left|\nabla^2 \chi_R\right| \leq 4R^{-2}$. Given $f \in C_b^{1,2}((0,T)\times\mathbb{R}^d)$ whose support has compact projection on $(0,T)$, set $f^R = f\chi_R \in \mathscr{A}$, so that (1.3) holds. The chain rule for diffusion operators

$$\mathcal{L}_t f^R = (\mathcal{L}_t f)\chi_R + f_t \mathcal{L}_t \chi_R + 2a_t(\nabla f_t, \nabla \chi_R)$$

entails the bound

$$|\mathcal{L}_t f^R| \leq |\mathcal{L}_t f| + |f_t|\,|\mathcal{L}_t \chi_R| + 2\,|a_t|\,|\nabla f_t|\,|\nabla \chi_R| \leq C\, \|f_t\|_{C_b^2} \left[|a_t| + |b_t|\right].$$

Letting $R \to \infty$, by dominated convergence, we extend the validity of (1.3) as required. ■

Remark 1.4 (narrowly continuous representative)**.** Arguing as in [Ambrosio et al., 2008, Lemma 8.1.2], it is not difficult to prove that any solution $\nu = (\nu)_{t\in(0,T)}$ to the FPE associated with $\mathcal{L}(a,b)$, with a, $b \in L^1(|\nu|)$, enjoys a (unique) narrowly continuous representative $\tilde{\nu} = (\tilde{\nu})_{t\in[0,T]}$ (i.e. for every $f \in C_b(\mathbb{R}^d)$, $f \mapsto \int f d\tilde{\nu}_t$ is continuous) such that $\nu_t = \tilde{\nu}_t$, for $\mathscr{L}^1$-a.e. $t \in (0,T)$ and for every $f \in C_b^{1,2}((0,T)\times\mathbb{R}^d)$, it holds

$$\int f_{t_2} d\tilde{\nu}_{t_2} - \int f_{t_1} d\tilde{\nu}_{t_1} = \int_{t_1}^{t_2} \int \left[\partial_t f + \mathcal{L}_t f\right] d\nu_t dt, \quad \text{for every } t_1,\, t_2 \in [0,T], \text{ with } t_1 \leq t_2.$$

Martingale problems

We introduce solutions to the martingale problem associated to $\mathcal{L}$, following [Stroock and Varadhan, 2006, Chapter 6]. We make use of the following notation: let $T > 0$ and, on the space $C([0,T];\mathbb{R}^d)$ (endowed with the sup norm), let $e_t : \gamma \mapsto \gamma_t := \gamma(t) \in \mathbb{R}^d$ be the evaluation map at $t \in [0,T]$. The natural filtration on $C([0,T];\mathbb{R}^d)$ is the increasing family of σ-algebras $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,T]}$ with $\mathcal{F}_t := \sigma(e_s : s \in [0,t])$. Given $\boldsymbol{\eta} \in \mathscr{P}(C([0,T];\mathbb{R}^d))$, we let $\eta_t := (e_t)_\sharp \boldsymbol{\eta}$ be the 1-marginal law at $t \in [0,T]$. Notice that the family $\eta := (\eta_t)_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ is Borel and actually narrowly continuous.

Definition 1.5 (solutions to MP's)**.** *Let a, b be Borel maps as in* (1.1)*. A probability measure $\boldsymbol{\eta} \in \mathscr{P}(C([0,T];\mathbb{R}^d))$ is a solution to the martingale problem (MP) associated to $\mathcal{L}(a,b)$ if a, $b \in L^1_{loc}(\eta)$ and, for every $f \in \mathscr{A}$, the process*

$$[0,T] \ni t \mapsto f_t \circ e_t - \int_0^t [\partial_t f_s + \mathcal{L}_s f] \circ e_s ds \tag{1.4}$$

is a martingale with respect to the natural filtration on $C([0,T];\mathbb{R}^d)$.

The assumption a, $b \in L^1_{loc}(\eta)$ entails that $t \mapsto \int_0^t [\partial_t f_s + \mathcal{L}_s f] \circ e_s ds$ can be defined as a progressively measurable process and moreover it belongs to $L^\infty_{loc}(\boldsymbol{\eta}, (\mathcal{F}_t)_t)$, i.e. there exists an increasing sequence of stopping times τ_n, converging towards T, $\boldsymbol{\eta}$-a.s., such that $\int_0^{\tau_n} [\partial_t f_s + \mathcal{L}_s f] \circ e_s ds \in L^\infty(\boldsymbol{\eta})$ for every $n \geq 1$: indeed we let

$$\tau_n := \inf \left\{ t \in [0,T] : \int_0^t |\partial_t f_s + \mathcal{L}_s f| \circ e_s ds d\boldsymbol{\eta} \geq n \right\}.$$

Remark 1.6 (the deterministic case)**.** When $a = 0$, solutions to the MP reduce to probability measures concentrated on absolutely continuous solutions to the ordinary differential equation (ODE)

$$\frac{d}{dt}\gamma_t = b_t(\gamma_t), \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T).$$

Indeed, arguing as in [Figalli, 2008, Lemma 3.8], every martingale (1.4) is identically zero: from this the thesis easily follows. ■

Remark 1.7 (solutions to MP's induce solutions to FPE's)**.** Integrating with respect to $\boldsymbol{\eta}$, i.e., taking expectation, we deduce that any solution $\boldsymbol{\eta}$ to the MP provides, by means of its 1-marginals $\eta = (\eta_t)_{t\in(0,T)}$ a narrowly continuous solution to the FPE associated to the same diffusion $\mathcal{L}$. Moreover, arguing as in Remark 1.3, if a, $b \in L^1(\eta)$, then the martingale property extends for processes of the form (1.4), with $f \in C^{1,2}_b((0,T)\times\mathbb{R}^d)$. ■

Solutions to martingale problems (as well as solutions to FPE's) enjoy many stability properties with respect to natural operations. The next proposition investigates the behaviour upon restriction of the interval of definition: quite obviously, all the definitions above can be given also with respect to any interval $[t_1,T]$ in place of $[0,T]$.

Proposition 1.8. *Let $t_1 \in [0,T)$, $\boldsymbol{\eta} \in \mathscr{P}(C([t_1,T];\mathbb{R}^d))$ be a solution to the martingale problem associated to a diffusion $\mathcal{L} = \mathcal{L}(a,b)$, with a, b as in* (1.1)*. Let $t_2 \in [t_1,T]$ and let $\rho : C([t_1,T];\mathbb{R}^d) \to [0,\infty)$ be a probability density (with respect to $\boldsymbol{\eta}$), measurable with respect to $\mathcal{F}_{t_2}$. Let π denote the natural restriction map*

$$C([t_1,T];\mathbb{R}^d) \ni \gamma \mapsto (\gamma_t)_{t\in[t_2,T]} \in C([t_2,T];\mathbb{R}^d).$$

Then, $\pi_\sharp(\rho\boldsymbol{\eta}) \in \mathscr{P}(C([t_2,T];\mathbb{R}^d)$ is a solution to the martingale problem associated to $\mathcal{L}$ in $C([t_2,T];\mathbb{R}^d)$.

Proof. It is sufficient to fix any $f \in C_c^{1,2}((t_2,T);\mathbb{R}^d)$, let $t \in [t_2,T]$ and $g : C([t_2,T];\mathbb{R}^d) \to \mathbb{R}$ be any bounded function, measurable with respect to $\sigma(e_r : t_2 \le r \le t)$, and prove

$$\int \left[f_T - \int_{t_2}^T (\partial_t + \mathcal{L}_s) f_s \, ds \right] g \, d\pi_\sharp(\rho \boldsymbol{\eta}) = \int \left[f_t - \int_{t_2}^t (\partial_t + \mathcal{L}_s) f_s \, ds \right] g \, d\pi_\sharp(\rho \boldsymbol{\eta}), \tag{1.5}$$

(here and below, for simplicity of notation, we omit to write e_s). The key point is to consider f as a function belonging to $C_c^{1,2}((t_1,T)\times\mathbb{R}^d)$, letting $f_s = 0$ for $s \in (t_1,t_2]$. The assumption on $\boldsymbol{\eta}$ gives that

$$[t_1,T] \ni t \mapsto f_t - \int_{t_1}^t (\partial_t + \mathcal{L}_s) f_s \, dr$$

is a martingale on the space $C([t_1,T];\mathbb{R}^d)$ endowed with the probability $\boldsymbol{\eta}$ and the natural filtration. Since $f_t = 0$ for $t \in (t_1,t_2]$, it holds, for $t \in (t_2,T)$,

$$f_t - \int_{t_1}^t (\partial_t + \mathcal{L}_r) f_r \, dr = f_t - \int_{t_2}^t (\partial_t + \mathcal{L}_r) f_r \, dr.$$

On the other hand, as $(g \circ \pi)\rho$ is $\mathcal{F}_{t_2}$-measurable, it holds

$$\int \left[f_T - \int_{t_1}^T (\partial_t + \mathcal{L}_s) f_s \, ds \right] (g \circ \pi) \, \rho d\boldsymbol{\eta} = \int \left[f_t - \int_{t_1}^t (\partial_t + \mathcal{L}_s) f_s \, ds \right] (g \circ \pi) \, \rho d\boldsymbol{\eta}.$$

These two identities entail (1.5). □

Next, we prove stability of solutions to MP's with respect to convex combinations. The correspondent statement for weak solutions to Fokker-Planck equations holds as well, see e.g. [Figalli, 2008, Lemma 2.4].

Proposition 1.9. *Let $\mathcal{L} = \mathcal{L}(a,b)$ be a diffusion operator with a, b as in* (1.1). *On a measurable space $(Z,\mathcal{A})$, let $\bar\nu \in \mathscr{P}(Z)$ and $(\boldsymbol{\eta}_z)_{z\in Z} \subseteq \mathscr{P}(C[0,T];\mathbb{R}^d)$ be a Borel family of probability measures, such that $\boldsymbol{\eta}_z$ is a solution to the MP associated to $\mathcal{L}$, for $\bar\nu$-a.e. $z \in Z$. Assume moreover that, for every $f \in \mathscr{A}$, it holds*

$$\int_Z \int |\mathcal{L} f| \, d\eta_z d\bar\nu(z) < \infty.$$

Let $\boldsymbol{\eta} := \int \boldsymbol{\eta}_z d\bar\nu(z) \in \mathscr{P}(C[0,T];\mathbb{R}^d))$, i.e.

$$\boldsymbol{\eta}(A) := \int \boldsymbol{\eta}_z(A) \, d\bar\nu(z), \quad \textit{for every } A \subseteq C[0,T];\mathbb{R}^d) \textit{ Borel.}$$

Then, $\boldsymbol{\eta}$ is a solution to the MP associated to $\mathcal{L}$.

Proof. The integrability assumption entails that, for every $f \in \mathscr{A}$, $t \in [0,T]$, the function $M_t := f_t \circ e_t - \int_0^t (\partial_t + \mathcal{L}_s) f \circ e_s \, ds$ belongs to $L^1(\boldsymbol{\eta})$. To check the martingale property, it is sufficient to apply Fubini's theorem: for $t \in [0,T]$ and any bounded $\mathcal{F}_t$-measurable function g, it holds

$$\int M_T g \, d\boldsymbol{\eta} = \int \left[\int M_T g \, d\boldsymbol{\eta}_z \right] d\bar\nu(z) = \int \left[\int M_t g \, d\boldsymbol{\eta}_z \right] d\bar\nu(z) = \int M_t g \, d\boldsymbol{\eta}.$$

□

Martingale flows

We define martingale flows as selections of solutions to martingale problems.

Definition 1.10 (martingale flows)**.** *Let $\mathcal{L} = \mathcal{L}(a, b)$ be a diffusion operator with a, b as in (1.1). A family of probability measures $(\boldsymbol{\eta}(s,x))_{(s,x)\in[0,T]\times\mathbb{R}^d} \subseteq \mathscr{P}(C([0,T];\mathbb{R}^d))$ is said to be a martingale flow (MF) associated to $\mathcal{L}$ if, for every $s \in [0,T]$, the map $x \mapsto \boldsymbol{\eta}(s,x)$ is Borel and for every $\bar\nu \in \mathscr{P}(\mathbb{R}^d)$, the probability measure*

$$\boldsymbol{\eta} := \int \boldsymbol{\eta}(s,x)d\bar\nu(x) \in \mathscr{P}(C([0,T];\mathbb{R}^d)) \tag{1.6}$$

is a solution to the martingale problem in $C([0,T];\mathbb{R}^d)$, associated to the diffusion $\chi_{[s,T]}\mathcal{L}$, with $\eta_s = \bar\nu$.

If we let $\bar\nu = \delta_x$ we deduce that for any martingale flow, $\boldsymbol{\eta}(s,x)$ is a solution to the MP such that $\eta(s,x)_s = \delta_x$. In view of Proposition 1.9, one could directly introduce martingale flows requiring that $\boldsymbol{\eta}(s,x)$ solves the martingale problem with initial datum δ_x, but we choose the formulation above since it translates directly to the metric measure spaces setting, see Definition 6.11 and Remark 6.12. Notice that one can equivalently require for $\boldsymbol{\eta}$ in (1.6) to be a solution of the martingale problem in $C([s,T];\mathbb{R}^d)$ associated to the diffusion $\mathcal{L}$, but it is technically easier to consider the flow as consisting of probability measures all defined the same space $C([0,T];\mathbb{R}^d)$. On the other hand, it clearly holds $\eta_t = \bar\nu$ also for $t \le s$.

Remark 1.11. In [Ambrosio and Crippa, 2008] as well as in [Figalli, 2008], the definition of flow is given without the parameter $s \in [0,T]$, i.e. by considering a Borel selection of solutions of the MP only at $s = 0$. Here, we prefer to look at forward solutions starting from every $s \in [0,T]$ and $x \in \mathbb{R}^d$, essentially because in the stochastic setting, Lemma 1.14 below does not provide a full counterpart of [Ambrosio and Crippa, 2008, Theorem 9]. Moreover, the collection $(\boldsymbol{\eta}(s,x))_{(s,x)\in[0,T]\times\mathbb{R}^d}$ provides, in some sense, a more complete description of the diffusion process associated to $\mathcal{L}$. ■

Remark 1.12 (Markov property)**.** We are not imposing to martingale flows any Markov or semigroup property, which in this formulation reads as the Chapman-Kolmogorv equations

$$\eta(s,x)_t = \int_{\mathbb{R}^d} \eta(r,y)_t\, \eta(s,x)_r(dy), \quad \text{for every } x \in \mathbb{R}^d,\ r,\ s,\ t \in [0,T] \text{ with } s \le r \le t. \tag{1.7}$$

We obtain them as a consequence of uniqueness, in Lemma 1.16. However, let us remark that it is of independent interest to restrict the study to flows that are Markov, as in [Stroock and Varadhan, 2006, Chapter 12].

The strong Markov property for flows (roughly, (1.7) with stopping times in place of deterministic times) seems to require for the joint map $(s,x) \mapsto \boldsymbol{\eta}(s,x)$ to be Borel. We say that a flow is strong if such a condition holds: this notion is technical and we prefer not do not address strong flows: throughout this thesis, we restrict the attention to existence and uniqueness issues for martingale flows (once uniqueness is proved, one may investigate whether the flow is strong). ■

The superposition principle

We conclude this section with an abstract definition for the validity of a superposition principle for diffusions, providing a "lift" of solutions to Fokker-Planck equations $\nu = (\nu_t)_t$ to solutions $\boldsymbol{\eta}$ to martingale problems, i.e. a converse to Remark 1.7.

Definition 1.13 (superposition principle). *Let a, b be Borel maps as in* (1.1) *and let $\nu = (\nu_t)_{t\in(0,T)} \subseteq \mathscr{P}(\mathbb{R}^d)$ be a weak solution to the FPE* (1.2)*, with $\mathcal{L} = \mathcal{L}(a, b)$. We say that the superposition principle holds for ν if there exists $\boldsymbol{\eta} \in \mathscr{P}(C([0, T]; \mathbb{R}^d))$, called a superposition solution for ν, which solves the MP associated to $\mathcal{L}(a, b)$ and it holds $\eta_t = \nu_t$ for $\mathscr{L}^1$-a.e. $t \in (0, T)$.*

In the deterministic case of vector fields, continuity equations and ODE's, the term "superposition principle" stems from the fact that, in the general case of non-regular coefficients, the superposition solution $\boldsymbol{\eta}$ is non-trivially distributed among the non-unique solutions to the ODE, thus introducing some randomness in an otherwise deterministic setting. Notice that, in the general case of diffusion operators and martingale problems, all the solutions show already some "intrinsic" randomness, so at this level the term is justified only by extension from the deterministic case.

1.2 Equivalence between FPE's, MP's and flows

The superposition principle is a crucial tool for establishing a perfect correspondence between the Eulerian (FPE's) and Lagrangian (MP's and flows) point of views on diffusion processes, transferring well-posedness results both ways. This connection is firmly established in the deterministic case, see e.g. [Ambrosio and Crippa, 2008, §4], in the case of diffusion operators, e.g. as in [Figalli, 2008, §2], but also for more abstract generators, see [Ethier and Kurtz, 1986, §4]. As already remarked at the beginning of the chapter, in this section we provide a complete equivalence, in particular between well-posedness of FPE's and MP's, in Lemma 1.15, provided that the superposition principle holds. We postpone the investigation of its validity, for general diffusions in $\mathbb{R}^d$, in Chapter 2.

Fokker-Planck equations ⇔ martingale problems

Equivalence for existence results is immediate, if the superposition principle holds: from existence of solutions ν to the FPE associated to $\mathcal{L}$, to which the superposition principle applies, we obtain existence of solutions to the correspondent MP, simply because the superposition principle is already a more precise statement. A first result that allows for transferring uniqueness is the following one, see [Figalli, 2008, Theorem 2.3].

Lemma 1.14 (transfer of uniqueness for 1-marginals). *Let a, b be Borel maps as in* (1.1)*, let $\bar{\nu} \in \mathscr{P}(\mathbb{R}^d)$ and assume that the superposition principle holds for every narrowly continuous solution $\nu = (\nu)_{t\in[0,T]} \subseteq \mathscr{P}(\mathbb{R}^d)$ to the FPE*

$$\partial_t \nu_t = \mathcal{L}_t^* \nu_t, \quad \textit{in } (0, T) \times \mathbb{R}^d, \quad \textit{with } \nu_0 = \bar{\nu}. \tag{1.8}$$

Then, the following conditions are equivalent:

i) there exists at most one narrowly continuous solution ν to the FPE (1.8)*;*

ii) any two solutions $\boldsymbol{\eta}^1$, $\boldsymbol{\eta}^2$ to the MP associated to $\mathcal{L}$, with $\eta^1_0 = \eta^2_0 = \bar{\nu}$, have identical 1-marginals, i.e. $\eta^1_t = \eta^2_t$ for $t \in [0, T]$.

Proof. i) $\Rightarrow$ ii). Let $\boldsymbol{\eta}^i$ be solutions to the MP associated to $\mathcal{L}$, with $\eta^i_0 = \bar{\nu}$ for $i \in \{1, 2\}$. Then $\eta^i = (\eta^1_t)_t$ is a narrowly continuous weak solution to the FPE associated to $\mathcal{L}$ with $\nu_0 = \bar{\nu}$, thus $\eta^1 = \eta^2$.

ii) $\Rightarrow$ i). It is sufficient to consider superposition solutions $\boldsymbol{\eta}^i$ for $i \in \{1, 2\}$ to deduce that $\nu^1 = \eta^1 = \eta^2 = \nu^2$. □

A stronger uniqueness result, at the level of processes, can be deduced assuming uniqueness for every initial datum and for all the restricted problems to intervals of the form $[s, T]$ for $s \in [0, T]$, together with the correspondent superposition principles. The argument we employ dates back at least to [Stroock and Varadhan, 2006, Theorem 6.2.3]; see also [Figalli, 2008, Proposition 5.5]. In the deterministic case, however, a different argument shows that it is not necessary to consider uniqueness starting from intermediate times, see [Ambrosio and Crippa, 2008, Theorem 9].

Lemma 1.15 (transfer of uniqueness)**.** *Let a, b be Borel maps as in* (1.1)*. For every $s \in [0, T]$, let the superposition principle hold, for every solution $(\nu)_{t\in(s,T)} \subseteq \mathscr{P}(\mathbb{R}^d)$ to the FPE*

$$\partial_t \nu_t = \mathcal{L}^*_t \nu_t, \quad \text{in } (s, T) \times \mathbb{R}^d,$$

where any superposition solution is required to solve the MP associated to $\mathcal{L}$ on $C([s, T]; \mathbb{R}^d)$.

Then, the following conditions are equivalent:

i) for every $s \in [0, T]$ and $\bar{\nu} \in \mathscr{P}(\mathbb{R}^d)$, there exists at most one narrowly continuous solution $\nu = (\nu_t)_{t\in[s,T]}$ to the FPE

$$\partial_t \nu_t = \mathcal{L}^*_t \nu_t, \quad \text{in } (s, T) \times \mathbb{R}^d, \quad \text{with } \nu_s = \bar{\nu};$$

ii) for every $s \in [0, T]$, if $\boldsymbol{\eta}^1$, $\boldsymbol{\eta}^2 \in \mathscr{P}(C([s, T]; \mathbb{R}^d))$ are solutions to the MP associated to $\mathcal{L}$ on $C([s, T]; \mathbb{R}^d)$, with $\eta^1_s = \eta^2_s$, then $\boldsymbol{\eta}^1 = \boldsymbol{\eta}^2$.

Proof. ii) $\Rightarrow$ i). goes identical as in Lemma 1.14: given ν^i, we consider a superposition solution $\boldsymbol{\eta}^i \in \mathscr{P}(C([s, T]; \mathbb{R}^d))$, for $i \in \{1, 2\}$, and by the assumption we deduce $\boldsymbol{\eta}^1 = \boldsymbol{\eta}^2$, in particular $\nu^1 = \eta^1 = \eta^2 = \nu^2$.

i) $\Rightarrow$ ii). In this case, the proof relies (implicitly) on the Markov property. Let $s \in [0, T]$ and $\boldsymbol{\eta}^1$, $\boldsymbol{\eta}^2 \in \mathscr{P}(C([s, T]; \mathbb{R}^d))$ be solutions to the MP associated to $\mathcal{L}$ on $C([s, T]; \mathbb{R}^d)$, with $\eta^1_s = \eta^2_s$. To deduce that $\boldsymbol{\eta}^1 = \boldsymbol{\eta}^2$, it is enough to show that, for every $n \geq 1$, the n-marginals of $\boldsymbol{\eta}^1$ an $\boldsymbol{\eta}^2$ coincide, i.e., for any $s \leq t_1 < \ldots < t_n \leq T$ and $A_1, \ldots, A_n \subseteq \mathbb{R}^d$ Borel sets, it holds

$$\boldsymbol{\eta}^1(e_{t_1} \in A_1, \ldots, e_{t_n} \in A_n) = \boldsymbol{\eta}^2(e_{t_1} \in A_1, \ldots, e_{t_n} \in A_n). \tag{1.9}$$

We argue by induction on $n \geq 1$. The case $n = 1$ holds true, as a consequence of *i) $\Rightarrow$ ii)* in Lemma 1.14, i.e. we use the fact that $(\eta^i_t)_{t\in[s,T]}$ for $i \in \{1, 2\}$ are narrowly continuous solutions to the same FPE, with $\eta^1_s = \eta^2_s$. To perform the induction step from n to $n + 1$, we argue as follows. For fixed $s \leq t_1 < \ldots < t_n < t_{n+1} \leq T$ and $A_1, \ldots, A_n, A_{n+1} \subseteq \mathbb{R}^d$ Borel sets, let

$$\rho := \frac{\prod_{i=1}^n \chi_{A_i}(e_{t_i})}{\boldsymbol{\eta}^1(e_{t_1} \in A_1, \ldots, e_{t_n} \in A_n)} : C([s, T]; \mathbb{R}^d) \to [0, \infty),$$

i.e., in probabilistic jargon, the density of $\boldsymbol{\eta}$ conditioned upon the event $\bigcap_{i=1}^n \{e_{t_i} \in A_i\}$. We assume that the denominator above is not null: otherwise there is nothing to prove since, by the inductive assumption, (1.9) holds true. Notice also that the inductive assumption similarly gives $(e_{t_n})_\sharp(\rho\boldsymbol{\eta}^1) = (e_{t_n})_\sharp(\rho\boldsymbol{\eta}^2)$, since it amounts to the identity

$$\boldsymbol{\eta}^1(e_{t_1} \in A_1, \ldots, e_{t_n} \in (A_n \cap B)) = \boldsymbol{\eta}^2(e_{t_1} \in A_1, \ldots, e_{t_n} \in (A_n \cap B)), \quad \text{for all } B \subseteq \mathbb{R}^d \text{ Borel.}$$

For $i \in \{1, 2\}$, define $\boldsymbol{\eta}^i_\rho$ as the push-forward of the measure $\rho\boldsymbol{\eta}^i$ with respect to the natural projection

$$\pi : C([s,T];\mathbb{R}^d) \ni \gamma \mapsto (\gamma_t)_{t\in[t_n,T]} \in C([t_n,T];\mathbb{R}^d).$$

By Proposition 1.8, $\boldsymbol{\eta}^i_\rho$ solves the martingale problem associated to $\mathcal{L}$ on $C([t_n,T];\mathbb{R}^d)$, with identical laws at t_n:

$$(\eta^1_\rho)_{t_n} = (e_{t_n})_\sharp(\rho\boldsymbol{\eta}^1) = (e_{t_n})_\sharp(\rho\boldsymbol{\eta}^2) = (\eta^2_\rho)_{t_n}$$

Again by the implication *i)* $\Rightarrow$ *ii)* in Lemma 1.14, we deduce that $(\eta^1_\rho)_t = (\eta^2_\rho)_t$ for $t \in [t_n,T]$, in particular for t_{n+1}, entailing

$$\frac{\boldsymbol{\eta}^1(e_{t_1} \in A_1, \ldots, e_{t_n} \in A_n, e_{t_{n+1}} \in A_{n+1})}{\boldsymbol{\eta}^1(e_{t_1} \in A_1, \ldots, e_{t_n} \in A_n)} = \frac{\boldsymbol{\eta}^2(e_{t_1} \in A_1, \ldots, e_{t_n} \in A_n, e_{t_{n+1}} \in A_{n+1})}{\boldsymbol{\eta}^2(e_{t_1} \in A_1, \ldots, e_{t_n} \in A_n)},$$

and so the correspondent of (1.9) for the case $n+1$ is settled. □

Martingale problems $\Leftrightarrow$ martingale flows

We investigate the link between well-posedness for martingale problems and flows. Since both these notions are "Lagrangian", there is no need of the superposition principle here. Let us introduce the following notation: for $(s,\bar\nu) \in [0,T] \times \mathscr{P}(\mathbb{R}^d)$, we write $C_{s,\bar\nu}(\mathcal{L}) \subseteq \mathscr{P}(C([0,T];\mathbb{R}^d))$ for the set of solutions $\boldsymbol{\eta}$ to the martingale problem associated to $\chi_{[0,s]}\mathcal{L}$, with $\eta_s = \bar\nu$.

Lemma 1.16 (well-posedness for MF's)**.** *Let $\mathcal{L} = \mathcal{L}(a,b)$ be a diffusion with a, b as in* (1.1), *and assume that $C_{s,x} := C_{s,\delta_x}(\mathcal{L})$ is compact for every $(s,x) \in [0,T] \times \mathbb{R}^d$, with*

$$[0,T] \times \mathbb{R}^d \ni (s,x) \mapsto C_{s,x} \in \mathscr{K}(\mathscr{P}(C([0,T];\mathbb{R}^d)))$$

Borel, where the target space is that of compact sets of $\mathscr{P}(C([0,T]\times\mathbb{R}^d))$ endowed with the Hausdorff distance, see e.g. [Stroock and Varadhan, 2006, Chapter 12, §1]. Assume also that

$$\sup_{x\in\mathbb{R}^d, \eta\in C_{s,x}} \int_s^T \int |\mathcal{L}_t f| \, d\eta_t \, dt < \infty,$$

for every $s \in [0,T]$, $f \in \mathscr{A} = C^{1,2}_c((0,T)\times\mathbb{R}^d)$.

Then, the following conditions are equivalent:

i) for every $s \in [0,T]$, $\bar\nu \in \mathscr{P}(\mathbb{R}^d)$, there exists a unique $\boldsymbol{\eta}(s,x) \in C_{s,\bar\nu}(\mathcal{L})$;

ii) there exists a unique martingale flow associated to $\mathcal{L}$.

In such a case, the (unique) martingale flow satisfies (1.7).

Proof. i) $\Rightarrow$ ii). Uniqueness of a martingale flow assuming uniqueness of martingale problems is trivial. By the assumption, the map $x \mapsto \boldsymbol{\eta}(s,x)$ is Borel being composition of Borel maps (projecting the singleton into its point is Borel), for every $s \in [0,T]$. To check that it provides a martingale flow, it is sufficient to apply Proposition 1.9.

ii) $\Rightarrow$ i). Existence of solutions to the martingale problem, assuming existence of a martingale flow follows trivially from (1.6). To prove uniqueness we argue as follows: assume that there exists $\bar{s} \in [0,T]$, $\bar{\mu}$ such that $\boldsymbol{\eta}^1$, $\boldsymbol{\eta}^2 \in C_{\bar{s},\bar{\mu}}$ with $\boldsymbol{\eta}^1 \neq \boldsymbol{\eta}^2$. By disintegrating with respect to $e_{\bar{s}}$, we deduce with no loss of generality that $\bar{\mu} = \delta_{\bar{x}}$, for some $\bar{x} \in \mathbb{R}^d$. Let $(\boldsymbol{\eta}(s,x))_{s,x}$ be a martingale flow (here we use existence) and then modify it only at the point $(\bar{s},\bar{x})$, letting $\boldsymbol{\eta}(\bar{s},\bar{x}) = \boldsymbol{\eta}^i(\bar{s},\bar{x})$, for $i \in \{1,2\}$. Clearly, the two maps obtained are Borel and provide two different martingale flows, in contrast with the uniqueness assumption. To check that they are martingale flows, it is sufficient to argue for $s = \bar{s}$ and rely on the identity

$$\int \boldsymbol{\eta}^i(\bar{s},x) d\bar{\nu}(x) = \int_{\{x \neq \bar{x}\}} \boldsymbol{\eta}(\bar{s},x) d\bar{\nu}(x) + \boldsymbol{\eta}^i(s,x)\bar{\nu}(\bar{x}),$$

and Proposition 1.9.

To prove that (1.7) holds, it is enough to notice that, if we let $\pi : C([s,T];\mathbb{R}^d) \mapsto C([r,T];\mathbb{R}^d)$ be the natural projection, by Proposition 1.8, then $\pi_\sharp(\boldsymbol{\eta}(s,x)) \in C_{r,\bar{\nu}}$ (to be rigorous, we have to extend it trivially on $[0,r]$), where $\bar{\nu} = \eta(s,x)_r$, by definition. By uniqueness, we deduce

$$\pi_\sharp(\boldsymbol{\eta}(s,x)) = \int_{\mathbb{R}^d} \boldsymbol{\eta}(r,y)\, \eta(s,x)_r(dy), \quad \text{as measures on } C([r,T];\mathbb{R}^d),$$

which entails (1.7). □

An identical proof shows that if

$$[0,T] \times \mathbb{R}^d \ni (s,x) \mapsto C_{s,x} \in \mathscr{K}(\mathscr{P}(C([0,T] \times \mathbb{R}^d)))$$

is Borel, then the unique martingale flow, if it exists, is strong in the sense introduced in Remark 1.12.

Remark 1.17 (flows in the deterministic case)**.** When $a = 0$, we noticed in Remark 1.6 that solutions to the martingale problem associated to $\mathcal{L}(0,b)$ correspond to probability measures concentrated solutions to the ODE driven by b. Under the assumptions of the result above, for every $(s,x) \in [0,T] \times \mathbb{R}^d$ there exists a unique solution $X(s,x) \in C([0,T];\mathbb{R}^d)$ and $\boldsymbol{\eta}(s,x) = \delta_{X(s,x)}$, thus the martingale flow reduces to the unique flow in the usual sense. In particular, Chapman-Kolmogorov (1.7) equations reads as the semigroup law $\mathbb{X}(s,x)(t) = \mathbb{X}(r,\mathbb{X}(s,x))(t)$, for $x \in \mathbb{R}^d$, r, s, $t \in [0,T]$, with $s \leq r \leq t$. ■

The results developed in this section apply in the case of diffusions with bounded and continuous coefficients a, b, as investigated by Stroock and Varadhan [2006]. In particular, the superposition principle holds, e.g. by Theorem 2.12, and Corollary 2.11 shows that the sets $C_{s,\bar{\nu}}$ are pre-compact. To prove that they are closed, one relies and the continuity of the coefficients to show that narrow limits of solutions to martingale problems solve the limit problem. Notice that continuity for $(s,\bar{\nu}) \mapsto C_{s,\bar{\nu}}$ is equivalent to the statement that for every sequence $(s_n,\bar{\nu}_n)$ with $s_n \to s$ in $[0,T]$, $\bar{\nu}_n \to \bar{\nu}$ narrowly in $\mathscr{P}(\mathbb{R}^d)$, up to extracting a subsequence, there exists $\boldsymbol{\eta}_n \in C_{s_n,\bar{\nu}_n}$ narrowly convergent to some $\boldsymbol{\eta} \in C_{s,\bar{\nu}}$. We deduce that the following statements are equivalent:

i) for every $s \in [0,T]$, $\bar{\nu} \in \mathscr{P}(\mathbb{R}^d)$, there exists a unique narrowly continuous solution to the FPE;
$$\partial_t \nu_t = \mathcal{L}_t^* \nu_t, \quad \text{in } (s,T), \quad \text{with } \nu_s = \bar{\nu};$$

ii) for every $s \in [0,T]$, $\bar{\nu} \in \mathscr{P}(\mathbb{R}^d)$, there exists a unique solution to the martingale problem associated to $\mathcal{L}$ in $C([s,T];\mathbb{R}^d)$, with law of e_s given by $\bar{\nu}$;

iii) there exists a unique martingale flow associated to $\mathcal{L}$.

Theorem 7.2.1 in Stroock and Varadhan [2006] entails that these conditions hold whenever a is uniformly bounded, continuous and elliptic, i.e. there exists $\lambda > 0$ such that $a \geq \lambda Id$ in $\mathrm{Sym}_+ \mathbb{R}^d$.

Chapter 2

The superposition principle

In this chapter, we establish the superposition principle for rather general diffusions in $\mathbb{R}^d$, in particular under minimal regularity and no ellipticity assumptions on coefficients: this both settles the equivalence results in the previous chapter and provides a rigorous foundation on which we build our deductions in the metric measure space setting, in Chapter 7.

The main result obtained in this chapter, i.e., Theorem 2.14, provides a far-reaching extension of [Figalli, 2008, Theorem 2.6], which gives superposition solutions in the case of diffusions whose coefficients are uniformly bounded in $(0,T)\times\mathbb{R}^d$. Let us mention that results in a similar spirit appear quite often in the literature, e.g. [Smirnov, 1993] in the framework of (deterministic) currents, or Echeverria's theorem [Ethier and Kurtz, 1986, Theorem 4.9.17] and extensions [Kurtz and Stockbridge, 1998] in the framework of martingale problems in spaces of càdlàg paths. Large parts of our deductions are to be regarded as counterparts of [Ambrosio et al., 2008, §8.1 and §8.2] in the setting of diffusions. From this premises, the result itself is rather natural, but let us remark that its derivation is not immediate from the available literature, due to non-trivial technical points (see Remark 2.15).

We also aim at providing a complete and self-contained exposition, thus in Section 2.1, we begin by establishing the superposition principle for diffusions having sufficiently smooth coefficients, which is our base case for subsequent deductions. In Section 2.2, we focus on some general features shared by many proofs of superposition principles in the literature, as well as in our case. In Section 2.3.1, we extend the superposition to diffusions with uniformly bounded coefficients, that is roughly the case covered in [Figalli, 2008, Theorem 2.6]. In Section 2.3.2 and Section 2.3.3 we move forward in generality, first to the case of locally bounded coefficients and then to our main result.

2.1 Case of smooth and bounded diffusions

We address the superposition principle for diffusions in $\mathbb{R}^n$ with sufficiently smooth and bounded coefficients. Precisely, we prove the following theorem.

Theorem 2.1 (superposition principle, smooth case)**.** *Let a, b be Borel maps as in* (1.1)*, with*

$$\int_0^T \left[\|a_t\|_{C^2_b(\mathbb{R}^d)} + \|b_t\|_{C^2_b(\mathbb{R}^d)}\right] dt < \infty. \tag{2.1}$$

Then, the superposition principle holds for every solution $\nu = (\nu_t)_{t\in(0,T)} \subseteq \mathscr{P}(\mathbb{R}^d)$ to the FPE (1.2) *associated to $\mathcal{L} = \mathcal{L}(a,b)$.*

The argument is standard and based on two well-known facts: the first one, of "Lagrangian" nature, is well-posedness for Itô stochastic differential equations with bounded Lipschitz coefficients, Theorem 2.2; the second one is "Eulerian", being uniqueness for narrowly continuous solutions to FPE's (1.2) for $\mathcal{L} = \mathcal{L}(a, b)$ with a b satisfying (2.1), Theorem 2.4. Here, the only novelty with respect to the classical approach with these problems consists in relaxing uniform bounds with respect to $t \in (0, T)$ to mere integral bounds.

Let us remark that, for the only purpose of establishing a case base for the validity of a general superposition principle, thanks to the scheme that we introduce in the next section (see also Remark 2.6), we may also argue under much stronger assumptions on the coefficients, e.g. $a, b \in C_b^\infty((0,T) \times \mathbb{R}^d)$, but it seems that there is no real gain in simplicity of our deductions.

Theorem 2.2 (existence for smooth MP's)**.** *Let a, b be Borel maps as in* (1.1)*, satisfying* (2.1)*. Then, for every $\bar{\nu} \in \mathscr{P}(\mathbb{R}^d)$, there exists a solution $\boldsymbol{\eta}$ to the MP associated to $\mathcal{L}(a, b)$, with $\eta_0 = \bar{\nu}$.*

Proof. The assumption $a \in L^1_t(C^2_b(\mathbb{R}^d))$ entails that the symmetric non-negative square-root of a, i.e. the essentially unique map

$$\sigma : (0,T) \times \mathbb{R}^d \to \mathrm{Sym}_+(\mathbb{R}^d) \quad \text{such that } \sigma_t^2 = a_t, \quad \mathscr{L}^1\text{-a.e. } t \in (0,T),$$

is bounded and Lipschitz with respect to $x \in \mathbb{R}^d$, with bounds integrable with respect to $t \in (0,T)$, see e.g. [Stroock and Varadhan, 2006, Lemma 3.2.3].

Let $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, be a filtered probability space (it is not necessary that the so-called usual assumptions on the filtration hold, see [Stroock and Varadhan, 2006, Lemma 4.3.3] and the discussion above it) endowed with a continuous $\mathcal{F}$-Wiener process $W = (W_t)_{t\in[0,T]}$ with values in $\mathbb{R}^d$. Let also $\overline{X}$ be a real-valued random variable, defined on the same probability space, and $\mathcal{F}_0$-measurable, thus independent of W.

Assume first that $\overline{X} \in L^2(\mathbb{P})$. Then, by Picard fixed point we solve the Itô stochastic differential equation (SDE)

$$dX_t = b_t(X_t)dt + \sqrt{2}\sigma_t(X_t)dW_t, \quad X_0 = \overline{X}, \tag{2.2}$$

obtaining a $\mathbb{P}$-a.s. continuous progressively measurable process $X = (X_t)_{t\in[0,T]}$ [Stroock and Varadhan, 2006, Theorem 5.11] such that it holds

$$X_t = \overline{X} + \int_0^t b_s(X_s)ds + \sqrt{2}\int_0^t \sigma_s(X_s)dW_s, \quad \text{ for } t \in [0,T], \ \mathbb{P}\text{-a.e. in } \Omega.$$

Itô's formula entails that the law of X, i.e. $\mathbb{P}_\sharp X$, is a solution to the martingale problem associated to $\mathcal{L}(a, b)$.

Moreover, by pathwise uniqueness, hence uniqueness in law, for the SDE (2.2), the function mapping the law of $\overline{X}$ into $\mathbb{P}_\sharp X$ is injective, one is able to provide a Borel family of probabilities $(\boldsymbol{\eta}(x))_{x\in\mathbb{R}^d} \subseteq \mathscr{P}(C([0,T];\mathbb{R}^d))$, each solving the martingale problem associated to $\mathcal{L}(a, b)$, with $\eta(x)_0 = \delta_x$, see [Stroock and Varadhan, 2006, Theorem 5.11]. Therefore, even if the law of $\bar{\nu}$ has no finite second moments, we show existence of a solution to the MP letting $\boldsymbol{\eta} = \int \boldsymbol{\eta}(x)d\bar{\nu}(x)$. □

Arguing as above, for $s \in [0,T]$, one is able to define a Borel family $(\boldsymbol{\eta}(s,x))_{(s,x)\in[0,T]\times\mathbb{R}^d}$ with $\boldsymbol{\eta}(s,x)$ solving the MP on $C([s,T];\mathbb{R}^d)$, i.e. a martingale flow after Definition 1.10. One can also prove that $(s,x) \mapsto \boldsymbol{\eta}(s,x)$ is narrowly continuous, arguing as in [Stroock and Varadhan, 2006, Theorem 5.1.4].

Remark 2.3 (the smooth, locally bounded, case)**.** Assuming that the coefficients a, b satisfy the local analogue of (2.1), i.e.

$$\int_0^T \left[\|a_t\|_{C^2(B)} + \|b_t\|_{C^1(B)} \right] dt < \infty, \quad \text{for every bounded } B \subseteq \mathbb{R}^d,$$

one is able to provide existence as well as unique for a maximal solution $X = (X_t)_{t \in [0,\tau)}$ to the SDE (2.2), up to an explosion time τ. By taking the one-point compactification of $\mathbb{R}^d$, i.e. adding the point ∞ and letting $X_t = \infty$ for $t \geq \tau$, one can prove that X thus defined is continuous and that its law $\boldsymbol{\eta}$ is a solution to the martingale problem associated to $\mathcal{L}(a,b)$, in duality with $C^2_c((0,T) \times \mathbb{R}^d)$ (with a slight abuse, since we view $\boldsymbol{\eta}$ as a probability measure on continuous curves with values in the compactification of $\mathbb{R}^d$ or as a sub-probability measure in $\mathbb{R}^d$). ■

Before we address the uniqueness result for FPE's, we recall some known results on (backward) Kolmogorov equations, referring e.g. to the expository notes by Krylov [1999] for more details. For simplicity, we let a, b be $C^\infty_b((0,T) \times \mathbb{R}^d)$ and $g \in C^\infty_c((0,T) \times \mathbb{R}^d)$. Then, Kolmogorov equations of the form

$$\partial_t f = -\mathcal{L}_t f + g, \quad \text{in } (0,T) \times \mathbb{R}^d, \quad f_T = \bar{f}, \tag{2.3}$$

provide a dual point of view to that of FPE's. Notice that in the deterministic case they reduce to (bakward) transport equations, with a source term g. A solution to the equation (2.3) is by definition a function $f \in C^{1,2}_b((0,T) \times \mathbb{R}^d)$ such that

$$\partial_t f(s,x) = -\mathcal{L}_s f(s,x) + g(s,x), \text{ for every } (s,x) \in (0,T) \times \mathbb{R}^d \text{ and } \lim_{s \uparrow T} f(s,x) = \bar{f}(x).$$

Kolmogorov equations can be investigated with martingale flows $\boldsymbol{\eta} = (\boldsymbol{\eta}(s,x))_{[0,T] \times \mathbb{R}^d}$, at least if the coefficients are sufficiently smooth, as we assume here. Indeed, one can prove by stochastic methods that, if $\bar{f} \in C^2_b(\mathbb{R}^d)$, then

$$f(s,x) := \int \bar{f} \, d\eta(s,x)_T - \int_s^T \int g_r \, d\eta(r,x)_T dr$$

provides a solution to (2.3). Notice that $\|f\|_\infty \leq \|\bar{f}\|_\infty + T \|g\|_\infty$ and if $g \geq 0$ then $f \leq \left\|\bar{f}\right\|_\infty$, which can be seen as consequence either of the fact that $\boldsymbol{\eta}(s,x)$ are probability measures or by the maximum principle and ultimately because a is non-negative, see [Stroock and Varadhan, 2006, Theorem 3.1.1]. Indeed, the maximum principle entails also uniqueness for solutions to (2.3), while existence can be proved by PDE techniques, e.g. by vanishing viscosity approximation as in [Stroock and Varadhan, 2006, Theorem 3.2.5 and Theorem 3.2.6]. We sketch the proof the following bound for solutions f, in terms of (2.1), which is useful for our purpose:

$$\sup_{t \in [0,T]} \|f_t\|_{C^2(\mathbb{R}^d)} \leq C \left\{ \int_0^T \left[\|a_t\|_{C^2_b} + \|b_t\|_{C^1_b} \right] dt + T \|g\|_{C^2} \right\} \left(\|\bar{f}\|_{C^2} + T \|g\|_{C^2} \right), \tag{2.4}$$

where $z \mapsto C\{z\}$ denotes some function depending on the dimension d only. As already noticed, the maximum principle entails a bound for $\|f_t\|_\infty$, uniform in $t \in [0,T]$, so that we are reduced to investigate the bounds for the derivatives. Let us highlight the formal argument in the proof, referring to [Stroock and Varadhan, 2006, Theorem 3.2.4] for a more general result. Assuming that f is sufficiently smooth, we write the equations solved by $|\nabla f_t|^2 = \nabla f_t \cdot \nabla f_t$ and $\left|\nabla^2 f\right|^2 = \nabla^2 f : \nabla^2 f$, which read as

$$\partial_t |\nabla f|^2 = 2\nabla f \cdot \nabla \partial_t f = -\mathcal{L} |\nabla f|^2 - 2R(f), \quad |\nabla f_T|^2 = \left|\nabla \bar{f}\right|^2,$$

and

$$\partial_t \left|\nabla^2 f\right|^2 = 2\nabla^2 f : \nabla^2 \partial_t f = -\mathcal{L} \left|\nabla^2 f\right|^2 - 2R'(f), \quad \left|\nabla^2 f_T\right|^2 = \left|\nabla^2 \bar{f}\right|^2,$$

where we also omit to denote the dependence upon $t \in (0,T)$ and let

$$R(f) := \frac{1}{2}\left[\mathcal{L} |\nabla f|^2 - 2\nabla f \cdot \nabla \mathcal{L} f\right] - \nabla g \cdot \nabla f,$$

$$R'(f) := \frac{1}{2}\left[\mathcal{L} \left|\nabla^2 f\right|^2 - 2\nabla^2 f : \nabla^2 \mathcal{L} f\right] - \nabla^2 g : \nabla^2 f.$$

The crucial point is to provide suitable bounds from below for R, R' in terms of a, b and their derivatives. By linearity, we deal separately with the cases $\mathcal{L}(a,0)$ and $\mathcal{L}(0,b)$ and let $g = 0$, the general case being a simple variant. In the latter case,

$$2R(f) = b \cdot \nabla |\nabla f|^2 - 2\nabla f \cdot \nabla(b \nabla f) = -2\nabla b : (\nabla f \otimes \nabla f) \geq - \|\nabla b\|_\infty |\nabla f|^2,$$

for $R(f)$, while for $R'(f)$, we estimate

$$\begin{aligned}
2R'(f) &= b \cdot \nabla \left|\nabla^2 f\right|^2 - 2\nabla^2 f : \nabla^2 (b \cdot \nabla f) \\
&= b \cdot \nabla \left|\nabla^2 f\right|^2 - 2\nabla^2 f : \left[(\nabla^2 b)\nabla f + 2(\nabla b)\nabla f + b\nabla^3 f\right] \\
&= -2\nabla^2 f : \left[(\nabla^2 b)\nabla f\right] - 4\nabla^2 f : \left[(\nabla b)\nabla^2 f\right] \\
&\geq -2\left[\left\|\nabla^2 b\right\|_\infty^2 + 2\|\nabla b\|_\infty\right] \left|\nabla^2 f\right|^2 - \frac{1}{2}\left\|\nabla^2 b\right\|_\infty^2 \|\nabla f\|_\infty^2 .
\end{aligned}$$

Computing the commutator for $\mathcal{L}(a,0)$, we obtain

$$\begin{aligned}
2R(f) &= a : \nabla^2 |\nabla f|^2 - 2\nabla f \cdot \nabla(a \nabla^2 f) \\
&= 2a : (\nabla^3 f \, \nabla f) + 2a : (\nabla^2 f)^2 - 2\nabla f \cdot \left[(\nabla a)(\nabla^2 f)\right] - 2\nabla f \cdot a \nabla^3 f \\
&= 2a : (\nabla^2 f)^2 - 2\nabla f \cdot \left[(\nabla a)(\nabla^2 f)\right] \\
&= 2\left|\sigma \nabla^2 f\right|^2 - 4\nabla f \cdot \left[(\nabla \sigma)\sigma \, \nabla^2 f)\right] \pm 2\left|(\nabla \sigma)\nabla f\right|^2 \\
&= 2\left|\sigma \nabla^2 f - (\nabla \sigma)\nabla f\right|^2 - 2\left|(\nabla \sigma)\nabla f\right|^2 \\
&\geq -2\|\nabla \sigma\|_\infty^2 |\nabla f|^2,
\end{aligned}$$

using the identities $A : (Bv) = AB \cdot v$ and $A^2 : B^2 = |AB|^2$. To handle the second order commutator, we argue similarly and obtain in conclusion the bound

$$R'(f) \geq -\left|(\nabla \sigma)\nabla^2 f\right|^2 - \nabla^2 f : \left[(\nabla^2 a)\nabla^2 f\right] \geq -\left[\|\nabla \sigma\|_\infty^2 + \left\|\nabla^2 a\right\|_\infty\right] \left|\nabla^2 f\right|^2 .$$

Thanks to these bounds, we obtain that $|\nabla f|^2$ satisfies

$$\partial_t |\nabla f|^2 \leq -\mathcal{L} |\nabla f|^2 + \left[\|\nabla b\|_\infty^2 + \|\nabla \sigma\|_\infty^2\right] |\nabla f|^2, \quad |\nabla f_T|^2 = \left|\nabla \bar{f}\right|^2 .$$

By Gronwall lemma and the maximum principle for $\mathcal{L}$ we deduce

$$\sup_{t\in[0,T]} \|\nabla f\|_\infty \le \exp\left\{\int_0^T \left[\|\nabla b_t\|_\infty^2 + \|\nabla\sigma_t\|_\infty^2\right] dt\right\} \|\nabla \bar{f}\|_\infty .$$

Now that this uniform bound is established, from the fact that $\left|\nabla^2 f\right|^2$ solves

$$\partial_t |\nabla f|^2 \le -\mathcal{L}\,|\nabla f|^2 + 2\left[\left\|\nabla^2 b\right\|_\infty^2 + 2\,\|\nabla b\|_\infty\right] \left|\nabla^2 f\right|^2 + \frac{1}{2}\left\|\nabla^2 b\right\|_\infty^2 \|\nabla f\|_\infty^2 , \ \left|\nabla^2 f_T\right|^2 = \left|\nabla^2 \bar{f}\right|^2$$

and again Gronwall lemma and the maximum principle for $\mathcal{L}$, we obtain a uniform upper bound for $\left|\nabla^2 f\right|^2$ in terms of

$$\exp\left\{\int_0^T \left[\|\nabla b_t\|_\infty^2 + \|\nabla\sigma_t\|_\infty^2\right] dt\right\} \left[\left\|\nabla^2 \bar{f}\right\|_\infty + \sup_{t\in[0,T]} \|\nabla f_t\|_\infty^2 \int_0^T \left\|\nabla^2 b_t\right\|_\infty^2 dt\right],$$

which leads to the inequality (2.4). These computations are formal, but can be made rigorous by considering the uniformly elliptic case, which provides sufficient regularity and then argue in the vanishing viscosity limit.

We are in a position to prove the following uniqueness and comparison theorem for weak solutions to Fokker-Planck equations, akin to [Ambrosio et al., 2008, Proposition 8.1.7].

Theorem 2.4 (uniqueness and comparison for smooth FPE's)**.** *Let a, b be Borel maps as in* (1.1) *and satisfy the following local version of* (2.1)*:*

$$\int_0^T \|a_t\|_{C^2(B)} + \|b_t\|_{C^2(B)}\, dt < \infty, \quad \textit{for every bounded open } B \subseteq \mathbb{R}^d.$$

Let $\nu = (\nu_t)_{t\in[0,T]} \subseteq \mathscr{M}(\mathbb{R}^d)$ be a narrowly continuous solution to the FPE associated to $\mathcal{L}(a,b)$, with a, $b \in L^1(\nu)$. Then, the condition $\nu_0 \le 0$ entails $\nu_t \le 0$, for every $t \in [0,T]$, and in particular there exists at most one narrowly continuous solution ν with a, $b \in L^1(\nu)$.

Proof. Let $g \in C_c^\infty((0,T)\times\mathbb{R}^d)$, with $g \ge 0$. Our aim is to show that $\int g\, d\nu \le 0$. Fix $R \ge 1$ large enough so that the support of g is contained in $(0,T)\times B_R(0)$ and let χ_R be as in Remark 1.3. Notice that letting $a_R = a\chi_R$ and $b_R = b\chi_R$ in place of a, b, condition (2.1) holds and $\mathcal{L}_R f = \mathcal{L} f$ on $(0,T)\times B_R(0)$, for every $f \in C_b^2((0,T)\times\mathbb{R}^d)$.

For $\varepsilon > 0$, let a_R^ε, b_R^ε be a double mollification with respect to the space and time variables, and define $\mathcal{L}_R^\varepsilon = \mathcal{L}(a_R^\varepsilon, b_R^\varepsilon)$, which is a diffusion operator with smooth and bounded coefficients, satisfying (2.1) uniformly in $\varepsilon > 0$. Let f^ε be a solution to the backward Kolmogorov equation

$$\partial_t f^\varepsilon = -\mathcal{L}_R^\varepsilon f^\varepsilon + g, \quad f_T^\varepsilon = 0,$$

and choose $f^\varepsilon \chi_R$ in the weak formulation (1.3), which is admissile because $f^\varepsilon \in C_b^{1,2}((0,T)\times\mathbb{R}^d)$ (see also Remark 1.4). Since $f^\varepsilon \le 0$ and $\nu_0 \le 0$, we have

$$\begin{aligned}
0 &\ge -\int f^\varepsilon \chi_R\, d\nu_0 = \int \left[\chi_R\, \partial_t f^\varepsilon + \mathcal{L}(f^\varepsilon \chi_R)\right] d\nu \\
&= \int \left[-\chi_R\, \mathcal{L}_R^\varepsilon f + \mathcal{L}(f^\varepsilon\chi_R)\right] d\nu \\
&= \int \left\{\chi_R\left[g + \mathcal{L}_R^\varepsilon f^\varepsilon - \mathcal{L} f^\varepsilon\right] + f^\varepsilon \mathcal{L}\chi_R + 2a(\nabla f^\varepsilon, \nabla\chi_R)\right\} d\nu \\
&\ge \int g\, d\nu - \sup_{t\in[0,T]} \|f_t^\varepsilon\|_{C_b^2(\mathbb{R}^d)} \int \left[\chi_R\,|a_R^\varepsilon - a| + |b_R^\varepsilon - b| + |\mathcal{L}\chi_R| + 2\,|a|\,|\nabla\chi_R|\right] d\,|\nu| .
\end{aligned}$$

As $\varepsilon \downarrow 0$, since $a_R = a$ and $b_R = b$ on $(0,T) \times B(0,R)$, the second integral converges to $\int [|\mathcal{L}\chi_R| + 2\,|a|\,|\nabla\chi_R|]\, d\,|\nu|$, and $\sup_{t\in[0,T]} \|f^\varepsilon_t\|_{C^2_b}$ is uniformly bounded in $\varepsilon > 0$, by (2.4). Finally, we let $R \to \infty$ and conclude, since $|\nabla\chi_R| + |\nabla\chi_R| \to 0$, pointwise and uniformly bounded. □

Proof of Theorem 2.1. By Remark 1.4, any weak solution $\nu = (\nu_t)_{t\in(0,T)}$ admits a narrowly continuous representative $\tilde{\nu}$. Let $\boldsymbol{\eta}$ be a solution to the martingale problem associated to $\mathcal{L}$, as provided by Theorem 2.2, with $\bar{\nu} = \tilde{\nu}_0$. By Remark 1.7, $\eta = (\eta_t)_{t\in[0,T]}$ provides a narrowly continuous solution to the Fokker-Planck equation associated to $\mathcal{L}$, with $\eta_0 = \tilde{\nu}_0$. By Theorem 2.4, we conclude that $\eta_t = \tilde{\nu}_t$, for $t \in [0,T]$. □

2.2 The approximation-tightness-limit scheme

There is a common structure in many proofs of superposition principles available in the literature, e.g. [Ambrosio et al., 2008, Theorem 8.2.1], [Ambrosio and Crippa, 2008, Theorem 12], [Ambrosio and Figalli, 2009, Theorem 4.5], [Figalli, 2008, Theorem 2.6], [Ambrosio and Trevisan, 2014, Theorem 7.1]. In this section, we highlight its main steps, providing at the same time useful auxiliary results.

Let $\nu = (\nu_t)_{t\in(0,T)}$ be a solution to the FPE associated to a diffusion operator $\mathcal{L}(a,b)$. To deduce existence of a superposition solution for ν, we argue in three steps.

Step 1 (approximation)**.** We build from ν a sequence of solutions $(\nu^n)_n$ to FPE's associated respectively to diffusions $(\mathcal{L}^n)_n$, for which the superposition principle is already known to hold, thus providing a sequence of solutions $(\boldsymbol{\eta}^n)_n$ to correspondent MP's. Here the difficulty is to provide a good approximation, so that ν^n converge towards ν, e.g., narrowly, and $\mathcal{L}$ towards $\mathcal{L}$, in a sense to be made precise, as $n \to \infty$.

Step 2 (tightness)**.** We prove that the sequence $(\boldsymbol{\eta}^n)_n \subseteq \mathscr{P}(C([0,T];\mathbb{R}^d))$ is tight, thus obtaining existence of a narrow limit $\boldsymbol{\eta}$, up to extracting a subsequence. By Ascoli-Arzelà criterion, this step reduces to prove bounds on the modulus of continuity.

Step 3 (limit)**.** From convergence $\nu^n \to \nu$, $\mathcal{L}^n \to \mathcal{L}$, as $n \to \infty$, we argue that $\boldsymbol{\eta}$ is a superposition solution for ν.

2.2.1 Approximation

The approximation step consists in mollifying by convolutions or considering suitable push-forwards of solutions via smooth maps. In this section we remark some general features of these transformations.

Push forward via smooth maps

Let $\nu = (\nu_t)_{t\in(0,T)} \subseteq \mathscr{M}^+(\mathbb{R}^d)$ be a solution to the FPE associated to $\mathcal{L}(a,b)$, where a, b are Borel maps as in (1.1) in $L^1(\nu)$. Let $k \geq 1$, and let π be a map

$$\pi : \mathbb{R}^d \to \mathbb{R}^k, \quad \text{with } \pi^i \in C^2_b(\mathbb{R}^d), \quad \text{for } i \in \{1,\dots k\}.$$

Then, it is possible to define a diffusion operator $\pi(\mathcal{L})$ on $\mathbb{R}^k$ such that $\pi(\nu) := (\pi_\sharp\nu_t)_{t\in(0,T)}$ is a solution to the associated FPE, in duality with $\mathscr{A}_k = C^{1,2}_c((0,T) \times \mathbb{R}^k)$. Indeed, the composition $f \circ \pi(t,x) := f(t,\pi(x))$ belongs to $C_b^{1,2}((0,T)\times\mathbb{R}^d)$, thus by Remark 1.3 we can take $f\circ\pi$ as a test function in the weak formulation (1.3) and the chain rule for diffusion operators entails that

$$\mathcal{L}(f\circ\pi) = \sum_{i=1}^k \mathcal{L}(\pi^i)\left[(\partial_i f)\circ\pi\right] + \sum_{i,j=1}^k a(\nabla\pi^i,\nabla\pi^j)\left[(\partial_{i,j}f)\circ\pi\right].$$

We define, for $(t,x)\in(0,T)\times\mathbb{R}^d$,

$$\pi(a)_t^{i,j}(x) := \mathbb{E}_{\nu_t}\left[a(\nabla\pi^i,\nabla\pi^j)\,|\,\pi = x\right] = \frac{d\pi_\sharp\left[a(\nabla\pi^i,\nabla\pi^j)\nu_t\right]}{d\pi_\sharp\nu_t}(x), \quad \text{for } i,j\in\{1,\dots k\},$$

and similarly

$$\pi(b)_t^i(x) := \mathbb{E}_{\nu_t}\left[\mathcal{L}(\pi^i)\,|\,\pi = x\right] = \frac{d\pi_\sharp\left[\mathcal{L}(\pi^i)\nu_t\right]}{d\pi_\sharp\nu_t}(x), \quad \text{for } i\in\{1,\dots k\},$$

we obtain that $\pi(\mathcal{L}) := \mathcal{L}(\pi(a),\pi(b))$ is a well-defined diffusion operator on $\mathbb{R}^k$ and $\pi(\nu)$ is a weak solution to the FPE

$$\partial_t\pi(\nu) = \pi(\mathcal{L})^*\pi(\nu), \qquad \text{in } (0,T)\times\mathbb{R}^d.$$

Notice that, if a, $b\in L^p(\nu)$, then $\pi(\mathcal{L})$ has coefficients in $L^p(\pi(\nu))$, since conditional expectations reduce norms and the derivatives of π^i are uniformly bounded. Moreover, uniform bounds on the coefficients are preserved by the operation $(\nu,\mathcal{L})\mapsto(\pi(\nu),\pi(\mathcal{L}))$, but in general local bounds are not. Moreover, since π is continuous, narrowly continuous curves are preserved.

Convolutions

Let $\nu = (\nu_t)_{t\in(0,T)} \subseteq \mathscr{M}^+(\mathbb{R}^d)$ be a solution to the FPE associated to $\mathcal{L}(a,b)$, with a, b Borel maps as in (1.1) and let $\rho\geq 0$ be a Borel probability density on $\mathbb{R}^d$ (with respect to $\mathscr{L}^d$). For our discussion, if a, b belong only to $L^1_{loc}(\nu)$, we have to assume that ρ has compact support, but as our interest lies on the case when ρ has full support, thus we let a, $b\in L^1(\nu)$. Then, it is possible to prove that the family of measures $\nu*\rho := (\nu_t*\rho)_{t\in(0,T)}$, obtained by convolution, consists of solution to FPE's associated to suitably defined diffusion operators. Indeed, for $f\in\mathscr{A}$, it holds $f*\rho\in C_b^{1,2}((0,T)\times\mathbb{R}^d)$ and

$$\begin{aligned}\mathcal{L}(f*\rho) &= \sum_{i=1}^d b^i\partial_i(f*\rho) + \sum_{i,j=1}^d a^{i,j}\partial_{i,j}(f*\rho)\\ &= \sum_{i=1}^d b^i(\partial_i f)*\rho + \sum_{i,j=1}^d a^{i,j}(\partial_{i,j}f)*\rho.\end{aligned}$$

We define

$$(a^\rho)^{i,j} := \frac{d(a^{i,j}\nu_t)*\rho}{d(\nu_t*\rho)}, \quad (b^\rho)^i := \frac{d(b^i\nu_t)*\rho}{d(\nu_t*\rho)}, \quad \forall i,j\in\{1,\dots d\}.$$

Then, $(\nu_t * \rho)_{t\in(0,T)}$ is a weak solution to the FPE associated to $\mathcal{L}^\rho := \mathcal{L}(a^\rho, b^\rho)$, since for $f \in \mathscr{A}$ it holds

$$\begin{aligned}\int \partial_t f \, d(\nu * \rho) &= \int (\partial_t f) * \rho \, d\nu = \int \partial_t (f * \rho) \, d\nu \\ &= -\int \mathcal{L}(f * \rho) \, d\nu = -\int \mathcal{L}^\rho f \, d(\nu * \rho).\end{aligned}$$

In the next lemma, we study integrability and regularity properties for (a^ρ, b^ρ), referring to [Ambrosio et al., 2008, Lemma 8.1.10] for more details.

Lemma 2.5. *Let ρ be a smooth probability kernel on $\mathbb{R}^d$ with $\rho > 0$ and $\left|\nabla^i \rho\right| \le C\rho$, for $i \in \{1, \dots k\}$, where $C \ge 0$ is some constant. Let μ, $\nu \in \mathscr{M}^+(\mathbb{R}^d)$, with $\mu \ll \nu$.*

Then, it holds $\mu * \rho \ll \nu * \rho$, and the following version of its density,

$$\frac{d(\mu * \rho)}{d(\nu * \rho)}(x) = \frac{\int \rho(x-y)\, d\mu(y)}{\int \rho(x-y)\, d\nu(y)}, \quad \textit{for every } x \in \mathbb{R}^d \tag{2.5}$$

is $C^k(\mathbb{R}^d)$. Moreover, for every convex, lower semicontinuous function $\beta : \mathbb{R} \mapsto [0, \infty]$, it holds

$$\int \beta\left(\frac{d(\mu * \rho)}{d(\nu * \rho)}\right) d(\nu * \rho) \le \int \beta\left(\frac{d\mu}{d\nu}\right) d\nu. \tag{2.6}$$

Similar conclusions hold when $\mu = (\mu_t)_{t\in[0,T]} \subseteq \mathscr{M}(\mathbb{R}^d)$ is Borel and $\nu = (\nu_t)_{t\in[0,T]} \subseteq \mathscr{M}^+(\mathbb{R}^d)$ is narrowly continuous, with $\mu_t \ll \nu_t$ for every $t \in [0,T]$. We obtain moreover the bound

$$\sup_{t\in[0,T]} \left\| \frac{d(\mu_t * \rho)}{d(\nu_t * \rho)} \right\|_{C_b^k(B)} < \infty,$$

for every open bounded set $B \subseteq \mathbb{R}^d$.

When we apply the lemma above to a narrowly continuous solution $\nu = (\nu_t)_{t\in[0,T]}$ to the FPE associated to $\mathcal{L}$, we deduce that, for $p \in [1, \infty]$, if a, $b \in L^p(\nu)$, then a^ρ, $b^\rho \in L^p(\nu * \rho)$ and moreover, if ν is narrowly continuous and a, $b \in L^1(\nu)$, then a^ρ_t, b^ρ_t are $C^k(\mathbb{R}^d)$, uniformly in $t \in [0,T]$ and in particular, locally bounded, uniformly in $t \in [0,T]$.

Proof. The proof of $\mu * \rho \ll \nu * \rho$ is trivial and we omit it. With a slight abuse of notation, we denote let $\mu * \rho(x)$, (respectively $\nu * \rho(x)$) the numerator (respectively the denominator) in the right hand side of (2.5). These functions are clearly $C^k(\mathbb{R}^d)$, with

$$\left|\nabla^i (\mu * \rho)\right|(x) \le C(\mu * \rho)(x), \quad \text{for } x \in \mathbb{R}^d, \quad \forall i \in \{1, \dots, k\}.$$

and similar bounds hold for $\nabla^i(\nu * \rho)$. The assumption $\rho > 0$ entails $\nu * \rho(x) > 0$ for every $x \in \mathbb{R}^d$, thus the quotient in (2.5) is $C^k(\mathbb{R}^d)$.

The second statement follows from Jensen's inequality at fixed $x \in \mathbb{R}^d$, applied to the 1-homogeneous, convex and lower semicontinuous function $(0,T) \times \mathbb{R}^d (t, z) \mapsto t\beta(|z|/t)$ and the measure $\rho(x - \cdot)\nu$ (notice that it does not need to be a probability measure, thanks to 1-homogeneity). The conclusion follows then by integration over $x \in \mathbb{R}^d$.

Finally, the t-dependent case is handled similarly, simply noticing that $(t, x) \mapsto \nu_t * \rho(x)$ is continuous and always strictly positive. □

Remark 2.6. Notice that we may let the kernel ρ depend also on t, by extending any narrowly continuous solution ν, letting $\nu_t = \nu_0$ for $t < 0$ and $\nu_t = \nu_T$ for $t > T$. The measures obtained by convolution solve suitable FPE's, whose coefficients are smooth also with respect to t. ■

Let us point out that more general approximations can be devised, e.g. by replacing π and ρ by Borel probability kernels $K = (K_x)_{x \in \mathbb{R}^d} \subseteq \mathscr{P}(\mathbb{R}^k)$, for some $k \geq 1$. The commutator between K and $\mathcal{L}$ plays an important role, and it seems difficult (perhaps worthless) at this stage to look for variants, as the two strategies discussed above are sufficient for our purposes.

2.2.2 Tightness

In this section we provide a compactness criterion for solutions to martingale problems in $\mathbb{R}^d$. By Ascoli-Arzelà theorem, tightness is achieved by estimating the modulus of continuity of solutions to martingale problems. In the deterministic case, solutions to ODE's are absolutely continuous curves; here we rely on analogous results for stochastic processes, leading to Hölder regularity for their paths from bounds on the quadratic variation of associated martingales. The technical tools that we employ are fractional Sobolev spaces of curves, clearly related to Kolmogorov criterion. Let us provide general definitions, for curves with values in metric spaces, as they become useful also in Part II.

Let (Y, d) a metric space. For $\delta \in (0,1)$, $p \in [1, \infty)$, we introduce the energy functional $\|\cdot\|_{\delta,p}$ on Borel curves $\gamma : (0,T) \mapsto Y$, given by

$$\|\gamma\|_{\delta,p}^p := \int_0^T \int_0^T \frac{d(\gamma_t, \gamma_s)^p}{|t-s|^{1+\delta p}} ds dt \in [0, \infty],$$

For $\delta \in (0,1)$, recall that a curve $\gamma : (0,T) \mapsto \mathbb{R}$ is said to be δ-Hölder continuous if

$$\|\gamma\|_{\delta,\infty} := \sup_{s \neq t \in (0,T)} \frac{d(\gamma_t, \gamma_s)}{|t-s|^\delta} < \infty.$$

The following embedding theorem holds, see e.g. [Di Nezza et al., 2012, Theorem 8.2].

Theorem 2.7 (fractional Sobolev embedding)**.** *Let (Y, d) be a complete metric space, let $\delta \in (0,1)$, and $p \in [1, \infty)$, satisfy $\delta p > 1$. Then, every Borel curve $\gamma : (0,T) \mapsto Y$ with $\|\gamma\|_{\delta,p} < \infty$ admits a (unique) $(\delta - 1/p)$-Hölder continuous representative $\tilde{\gamma}$, i.e. $\gamma_t = \tilde{\gamma}_t$, $\mathscr{L}^1$-a.e. $t \in (0,T)$. Moreover, it holds*

$$\|\tilde{\gamma}\|_{(\delta - 1/p)} \leq C \|\gamma\|_{\delta,p},$$

for some constant C depending on δ, p and T only.

Before we address the specific situation of solutions to martingale problem, we give a general result relying on the previous theorem and Burkholder-Gundy inequalities.

Lemma 2.8. *Let $(\Omega, (\mathcal{F}_t)_{t \in [0,T]}, \mathbb{P})$ be a filtered probability space and let*

$$\varphi = (\varphi_t)_t, \quad \ell = (\ell_t)_t, \quad \alpha = (\alpha_t)_t$$

be progressively measurable processes, with $\alpha \geq 0$. Assume moreover that

$$[0,T] \ni t \mapsto M_t := \varphi_t - \int_0^t \ell_s \, ds, \quad \textit{and} \quad [0,T] \ni t \mapsto M_t^2 - \int_0^t \alpha_s \, ds$$

are $\mathbb{P}$-a.s. continuous local martingales. For $p \in (1, \infty)$, $\delta \in (1/p, 1)$, let

$$C([0,T];\mathbb{R}) \ni \mathcal{A}(\gamma) := \inf_{\gamma=\gamma^1+\gamma^2} \left\{ \left\|\gamma^1\right\|_{\delta-1/p} + \left\|\gamma^2\right\|_{(\delta-1/p)/2} \right\}. \tag{2.7}$$

Then, for some constant C depending on p, δ and T only, it holds

$$\mathbb{E}\left[\mathcal{A}(\varphi)\right] \leq C \left\{ \left[\int_0^T \mathbb{E}\left[|\ell_t|^p\right] dt \right]^{1/p} + \left[\int_0^T \mathbb{E}\left[\alpha_t^p\right] dt \right]^{1/2p} \right\}. \tag{2.8}$$

Proof. The assumptions give that $(M_t)_t$ is a local martingale, whose quadratic variation process is $t \mapsto \int_0^t \alpha_s ds$. We let

$$\gamma_t^1 := \int_0^t \ell_s\, ds, \quad \gamma_t^2 := M_t, \quad \text{for } t \in [0,T],$$

thus the left hand side in (2.8) is smaller than

$$\mathbb{E}\left[\left\|\gamma^1\right\|_{\delta-1/p}\right] + \mathbb{E}\left[\left\|\gamma^2\right\|_{(\delta-1/p)/2}\right], \tag{2.9}$$

for which we provide separate bounds. For the first term, we use Hölder inequality

$$\mathbb{E}\left[\left\|\gamma^1\right\|_{\delta-1/p}\right] \leq \mathbb{E}\left[\left\|\gamma^1\right\|_{\delta-1/p}^p\right]^{1/p},$$

Theorem 2.7 and Fubini's theorem, reducing the problem to bound from above

$$\begin{aligned} \mathbb{E}\left[\left\|\gamma^1\right\|_{\delta,p}^p\right] &= \int_{(0,T)^2} \frac{\mathbb{E}\left[\int_s^t |\ell_r|\, dr\right]^p}{|s-t|^{1+\delta p}} dsdt \\ &\leq \int_{(0,T)^2} \frac{\int_s^t \mathbb{E}\left[|\ell_r|^p\right] dr}{|s-t|^{2-p(1-\delta)}} dsdt \\ &= \int_0^T \mathbb{E}\left[|\ell_r|^p\right] g_{p(1-\delta)}, \end{aligned}$$

where we let, for $\sigma > 0$,

$$g_\sigma(r) := \int_r^T \int_0^r |s-t|^{\sigma-2}\, dsdt.$$

Notice that g_σ is uniformly bounded, since

$$g_\sigma(r) = \frac{1}{\sigma-1} \int_r^T \left[(t-r)^{\sigma-1} - t^{\sigma-1}\right] dt = \frac{(T-r)^\sigma - (T^\sigma - r^\sigma)}{\sigma(\sigma-1)},$$

thus we obtain, for some constant C depending on δ, p and T only,

$$\mathbb{E}\left[\left\|\gamma^1\right\|_\delta\right] \leq C \left[\int_0^T \mathbb{E}\left[|\ell_t|^p\right]\right]^{1/p}.$$

Next, we provide a bound for the second term in (2.9). We use again Hölder inequality, with exponent $2p$,

$$\mathbb{E}\left[\|M\|_{(\delta-1/p)/2}\right] \leq \mathbb{E}\left[\|M\|_{(\delta-1/p)/2}^{2p}\right]^{1/2p},$$

Theorem 2.7 and Fubini's theorem, reducing the problem to bound from above the quantity

$$\mathbb{E}\left[\|M\|_{\delta/2,2p}^{2p}\right] = \int_{(0,T)^2} \frac{\mathbb{E}\left[|M_t - M_s|^{2p}\right]}{|s-t|^{1+\delta p}} dsdt.$$

By Burkholder-Gundy inequalities with exponent $2p$, there exists some constant depending on p only such that

$$\mathbb{E}\left[|M_t - M_s|^{2p}\right] \le C\mathbb{E}\left[\left|\int_s^t \alpha_r dr\right|^p\right] \le C\,|t-s|^{p-1}\int_s^t \mathbb{E}\left[\alpha_r^p\right] dr,$$

thus, arguing as in the previous case, we obtain

$$\mathbb{E}\left[\|M\|_{\delta/2,2p}^{2p}\right] \le C\int_0^T \mathbb{E}\left[\alpha_r^p\right] g_{p(1-\delta)}(r)dr.$$

and (2.8) is settled. □

Notice that, given any coercive function $\theta : \mathbb{R} \mapsto [0,\infty]$, the functional $\gamma \mapsto \theta(\gamma_0) + \mathcal{A}(\gamma)$ is coercive, since if $\mathcal{A}(\gamma) \le K$, then it belongs to the image of a compact rectangle in $C([0,T];\mathbb{R})^2$ by means of the mapping $(\gamma^1,\gamma^2) \mapsto \gamma^1 + \gamma^2$.

Remark 2.9. In the deterministic case, i.e. when $a = 0$, we obtain that φ is $\mathbb{P}$-a.s. δ-Hölder continuous on $(0,T)$, for any $\delta \in (0, 1-1/p)$. This result is almost optimal, since it does not entail $(1-1/p)$-Hölder continuity. On the other side, it is well-known that the real valued Wiener process, obtained by letting $\ell = 0$ and $\alpha = 1$, is concentrated on paths that are $\mathbb{P}$-a.s. δ-Hölder continuous for every $\delta \in (0,1/2)$, but not $1/2$-Hölder continuous: from this point of view the result is optimal (as we may let $p \to \infty$). We may also provide a variant allowing for different integrability on ℓ and α. It is reasonable to assume that in the case $p = 1$ one can provide a different coercive functional, for the absolutely continuous part, as in the proof of [Ambrosio and Crippa, 2008, Theorem 12]. On the other side, for the martingale part, it is not clear how to deal with the case $p = 1$. ■

We now discuss how Lemma 2.8 becomes useful in the tightness step. Given a solution $\boldsymbol{\eta}$ to the martingale problem associated to $\mathcal{L}(a,b)$ with a, b as in (1.1), for any $f \in \mathscr{A}$, we claim that letting

$$\varphi_t = f_t \circ e_t, \quad \ell_t = [\partial_t f + \mathcal{L}_t f] \circ e_t, \quad \alpha_t = 2a_t(\nabla f_t, \nabla f_t) \circ e_t, \quad \text{for } t \in [0,T].$$

we are in the situation of Lemma 2.8, with $\Omega = C([0,T];\mathbb{R}^d)$, $\mathbb{P} = \boldsymbol{\eta}$ and $\mathcal{F}$ being the canonical filtration. Indeed, $M_t := f_t - \int_0^t \ell_s\, ds$ is a martingale by the very definition of solution to the martingale problem, so we only have to prove that $M_t^2 - \int_0^t \alpha_s\, ds$ is a local martingale.

To this aim, we notice first that, as $f^2 \in \mathscr{A}$, the process

$$[0,T] \ni t \mapsto f_t^2 - \int_0^t \left[\partial_t + \mathcal{L}_s\right] f^2 \circ e_s\, ds \tag{2.10}$$

is a martingale as well. The key point is to use the identity

$$a(\nabla f, \nabla f) = \frac{1}{2}\left[(\partial_t + \mathcal{L})f^2 - 2f(\partial_t + \mathcal{L})f\right].$$

To keep notation simple, we omit to write any appearance of e_t, e_s or e_r in the following identities. Developing the square of M_t, it holds

$$M_t^2 = (f_t - \int_0^t \ell_s \, ds)^2 = f_t^2 + 2\int_0^t (\partial_t + \mathcal{L}) f_s \left[\int_s^t (\partial_t + \mathcal{L}) f_r \, dr - f_t \right] ds.$$

Then, we replace f_t^2 by using (2.10), to deduce that

$$t \mapsto M_t^2 - \int_0^t (\partial_t + \mathcal{L}) f_s^2 \, ds - 2\int (\partial_t + \mathcal{L}) f_s \left[\int_s^t (\partial_t + \mathcal{L}) f_r dr - f_t \right] ds$$

is a martingale. We add and subtract $2\int_0^t f_s (\partial_t + \mathcal{L}) f_s ds$, obtaining that

$$t \mapsto M_t^2 - \int_0^t \alpha_s \, ds + 2\int_0^t (\partial_t + \mathcal{L}) f_s \left[f_t - f_s - \int_s^t (\partial_t + \mathcal{L}) f_r dr \right] ds$$

is a martingale. We conclude by noticing that

$$t \mapsto \int_0^t (\partial_t + \mathcal{L}) f_s \left[f_t - f_s - \int_s^t (\partial_t + \mathcal{L}) f_r \, dr \right] ds = \int_0^t \ell_s \left(M_t - M_s \right) ds$$

is a local martingale, as a consequence of the following lemma (see also [Stroock and Varadhan, 2006, Theorem 1.2.8]).

Lemma 2.10. *Let $(\Omega, (\mathcal{F}_t)_{t \in [0,T]}, \mathbb{P})$ be a filtered probability space and let $\ell = (\ell_t)_t$ be progressively measurable, $M = (M_t)_t \in L^\infty_{loc}(\mathbb{P}, (\mathcal{F}_t)_t)$ be a local martingale, and*

$$\int_0^T \mathbb{E}\left[|\ell_t| \right] dt < \infty.$$

Then, the process $\int_0^t \ell_s \left(M_t - M_s \right) ds$ is a local martingale.

Proof. After localization, we are reduced to the case $M \in L^\infty(\mathbb{P})$, so that $\int_0^t \ell_s \left(M_t - M_s \right) ds \in L^1(\mathbb{P})$, for every $t \in [0,T]$, so we focus on orthogonality of increments. It is sufficient to fix $t \in [0,T]$ and show that

$$\mathbb{E}\left[\int_0^T \ell_s \left(M_T - M_s \right) ds \,|\, \mathcal{F}_t \right] = \int_0^t \ell_s \left(M_t - M_s \right) ds.$$

The integrability assumptions provide a justification for exchanging the order between conditional expectation and integration with respect to s. We consider two cases: if $s \in [0,t]$, then ℓ_s in $\mathcal{F}_s$ measurable so

$$\mathbb{E}\left[\ell_s \left(M_T - M_s \right) \,|\, \mathcal{F}_t \right] = \ell_s \, \mathbb{E}\left[M_T - M_s \,|\, \mathcal{F}_t \right] = \ell_s \left(M_t - M_s \right).$$

If $s \in [t,T]$, then

$$\mathbb{E}\left[\ell_s \left(M_T - M_s \right) \,|\, \mathcal{F}_t \right] = \mathbb{E}\left[\mathbb{E}\left[\ell_s \left(M_T - M_s \right) \,|\, \mathcal{F}_s \right] \,|\, \mathcal{F}_t \right] = \mathbb{E}\left[\ell_s \, \mathbb{E}\left[M_T - M_s \,|\, \mathcal{F}_s \right] \,|\, \mathcal{F}_t \right] = 0,$$

and the thesis follows. □

As a conclusion, we can specialize Lemma 2.8 as follows.

Corollary 2.11. *Let a, b be Borel maps as in (1.1), let $\boldsymbol{\eta} \in \mathscr{P}(C([0,T];\mathbb{R}^d))$ be a solution to the martingale problem associated to $\mathcal{L}(a,b)$. For any $p \in (1,\infty)$, $\delta \in (0, 1-1/p)$ define $\mathcal{A}$ as in (2.7) let and $f \in C_b^{1,2}((0,T)\times\mathbb{R}^d)$. Then, letting $\varphi_t := f_t \circ e_t$, it holds*

$$\mathbb{E}\left[\mathcal{A}(\varphi)\right] \le C\left\{\left[\int_0^T\int |\partial_t f_t + \mathcal{L}_t f|^p \, d\eta_t \, dt\right]^{1/p} + \left[\int_0^T \int |a_t(\nabla f_t, \nabla f_t)|^p \, d\eta_t \, dt\right]^{1/2p}\right\},$$

where C is some constant depending on p, δ and T only.

2.2.3 Limit

In the third step we assume that the probability measures $(\boldsymbol{\eta}^n)_n$, obtained as superposition solutions for an approximating sequence $(\nu^n)_n$, narrowly converge in $\mathscr{P}(C([0,T];\mathbb{R}^d))$ towards some limit $\boldsymbol{\eta}$. To deduce that $\boldsymbol{\eta}$ provides a superposition solution for ν is not obvious, due to the fact that, although test functions $f \in \mathscr{A}$ are continuous, we must deal with a limit in the weak formulation, with terms involving the coefficients a, b, in general not continuous. The strategy is to rely on density arguments and exploit the approximation procedure that provides the sequence $(\nu^n)_n$.

More precisely, to show that $\boldsymbol{\eta}$ is a solution for the martingale problem associated to a limit diffusion $\mathcal{L}$, it is enough to establish following property: for every s, $t \in [0,T]$ with $s \le t$, for every $f \in \mathscr{A}$ (with $\|f\|_{C^{1,2}} \le 1$) and for every g bounded (with $\|g\|_\infty \le 1$), continuous and $\mathcal{F}_s$-measurable on $C([0,T];\mathbb{R}^d)$, it holds

$$\int g\left[f\circ e_t - f\circ e_s - \int_s^t \left[(\partial_t + \mathcal{L}_r)f\right]\circ e_r \, dr\right] d\boldsymbol{\eta} = 0.$$

Assuming that the correspondent identity holds for $\boldsymbol{\eta}^n$ and $\mathcal{L}^n$, i.e.

$$\int g\left[f\circ e_t - f\circ e_s - \int_s^t \left[(\partial_t + \mathcal{L}_r^n)f\right]\circ e_r \, dr\right] d\boldsymbol{\eta}^n = 0,$$

to deduce that $\boldsymbol{\eta}$ is a solution to the martingale problem associated to $\mathcal{L}$, since f and $\partial_t f$ are bounded and continuous, the crucial limit is

$$\int g\left[\int_s^t (\mathcal{L}_r^n f)\circ e_r \, dr\right] d\boldsymbol{\eta}^n - \int g\left[\int_s^t (\mathcal{L}_r f)\circ e_r \, dr\right] d\boldsymbol{\eta} \to 0. \tag{2.11}$$

We argue accordingly to the two approximating strategies introduced in Section 2.2.1.

Push forward via smooth maps

Let a, $b \in L^1(\nu)$ and, for $n \ge 1$, let $\pi^n \in C_b^2(\mathbb{R}^d, \mathbb{R}^k)$ converge towards a $\pi \in C_b^2(\mathbb{R}^d, \mathbb{R}^k)$, in the following sense: $\pi^n \to \pi$ pointwise, uniformly on compact sets and $\nabla\pi^n \to \nabla\pi$ and $\nabla^2\pi^n \to \nabla^2\pi$, pointwise and uniformly bounded. Our interests lie in the case $k = d$ and $\pi = Id$, but the general argument might gain in clarity.

Let $\nu^n = \pi^n(\nu)$, $\mathcal{L}^n = \pi^n(\mathcal{L})$, and $\boldsymbol{\eta}^n$ be superposition solutions for ν^n. We argue that $\boldsymbol{\eta}$ to be a superposition solution for $\pi(\nu)$, with respect to the diffusion operator $\pi(\mathcal{L})$. We begin by adding and subtracting the term

$$\int g\left[\int_s^t (\overline{\mathcal{L}}_r f)\circ e_r \, dr\right] d\boldsymbol{\eta}^n - \int g\left[\int_s^t (\overline{\mathcal{L}}_r f)\circ e_r \, dr\right] d\boldsymbol{\eta}$$

in (2.11), where $\overline{\mathcal{L}} = \mathcal{L}(\overline{a}, \overline{b})$ is a diffusion operator on $\mathbb{R}^k$, with continuous and compactly supported coefficients $\overline{a}$, $\overline{b}$. The difference terms above are infinitesimal as $n \to \infty$, by narrow convergence of $\boldsymbol{\eta}^n$, obtaining therefore a bound from above for the absolute value of (2.11), as $n \to \infty$, in terms of

$$\limsup_{n\to\infty} \int \left|\mathcal{L}^n f - \overline{\mathcal{L}} f\right| d\nu^n + \int \left|\pi(\mathcal{L}) f - \overline{\mathcal{L}} f\right| d\pi(\nu). \tag{2.12}$$

Let us focus on the last term in the right hand side above. By definition of $\pi(\mathcal{L})$, it holds

$$\pi(\mathcal{L}) f(y) = \mathbb{E}_\nu \left[\mathcal{L}(f \circ \pi) \,|\, \pi = y\right], \qquad \nu\text{-a.e. } y \in \mathbb{R}^k,$$

so by the abstract change of variables with respect to π,

$$\int \left|\pi(\mathcal{L}) f - \overline{\mathcal{L}} f\right| d\pi(\nu) = \int \left|\mathbb{E}_\nu \left[\mathcal{L}(f \circ \pi) \,|\, \pi\right] - (\overline{\mathcal{L}} f) \circ \pi\right| d\nu.$$

Being $(\overline{\mathcal{L}} f) \circ \pi$ a function of π, the conditional expectation leaves it unchanged, up to ν-negligible sets, so

$$\begin{aligned} \int \left|\mathbb{E}_\nu \left[\mathcal{L}(f \circ \pi) \,|\, \pi\right] - (\overline{\mathcal{L}} f) \circ \pi\right| d\nu &= \int \left|\mathbb{E}_\nu \left[\mathcal{L}(f \circ \pi) - (\overline{\mathcal{L}} f) \circ \pi \,|\, \pi\right]\right| d\nu \\ &\leq \int \left|\mathcal{L}(f \circ \pi) - (\overline{\mathcal{L}} f) \circ \pi\right| d\nu, \end{aligned}$$

where the last inequality holds because conditional expectation reduces L^1-norms. Writing explicitly the difference

$$\mathcal{L}(f \circ \pi) - (\overline{\mathcal{L}} f) \circ \pi = \sum_{i,j=1}^k \left[a(\nabla \pi^i, \nabla \pi^j) - \overline{a}^{i,j} \circ \pi\right] (\partial_{i,j} f) \circ \pi + \sum_{i=1}^k \left[\mathcal{L}(\pi^i) - \overline{b}^i \circ \pi\right] (\partial_i f) \circ \pi,$$

and recalling that $\|f\|_{C^{1,2}} \leq 1$, we conclude that

$$\int \left|\mathcal{L} f - \overline{\mathcal{L}} f\right| d\nu \leq \int \sum_{i,j=1}^k \left|a(\nabla \pi^i, \nabla \pi^j) - \overline{a}^{i,j} \circ \pi\right| \, d\nu + \int \sum_{i=1}^k \left|\mathcal{L}(\pi^i) - \overline{b}^i \circ \pi\right| \, d\nu. \tag{2.13}$$

A similar bound can be proved for the first term in (2.12):

$$\limsup_{n\to\infty} \int \left|\mathcal{L}^n f - \overline{\mathcal{L}} f\right| d\nu^n \leq \int \sum_{i,j=1}^k \left|a(\nabla \pi^i, \nabla \pi^j) - \overline{a}^{i,j} \circ \pi\right| \, d\nu + \int \sum_{i=1}^k \left|\mathcal{L}(\pi^i) - \overline{b}^i \circ \pi\right| \, d\nu.$$

Indeed, arguing similarly but for fixed $n \geq 1$, with π^n in place of π, we obtain

$$\int \left|\mathcal{L}^n f - \overline{\mathcal{L}} f\right| d\nu^n \leq \int \sum_{i,j=1}^k \left|a(\nabla (\pi^n)^i, \nabla (\pi^n)^j) - \overline{a}^{i,j} \circ \pi^n\right| \, d\nu + \int \sum_{i=1}^k \left|\mathcal{L}((\pi^n)^i) - \overline{b}^i \circ \pi^n\right| \, d\nu,$$

By the assumptions on $\pi^n \to \pi$, as $n \to \infty$, Lebesgue dominated convergence applies.

To conclude, we have to choose $\overline{a}$, $\overline{b}$, minimizing the right hand side in (2.13), which can be made arbitrary small if $a(\nabla \pi^i, \nabla \pi^j)$ and $\mathcal{L}(\pi^i)$ are measurable with respect to π, for $i, j \in \{1, \ldots, k\}$: indeed, we perform again a change of variables back to the measure $\pi(\nu)$ and rely on density of continuous, compactly supported functions in $L^1(\pi(\nu))$.

Convolution

In this case, the argument follows closely that in the proof of [Ambrosio et al., 2008, Theorem 8.2.1]. We assume that a sequence ρ^n of probability densities on $\mathbb{R}^d$ is given, let $\nu^n = \nu * \rho_n$ and $\mathcal{L}^n$ be the diffusion operator with coefficients

$$a^n := \frac{d(a\nu * \rho_n)}{d(\nu * \rho_n)} \quad \text{and} \quad b^n := \frac{d(b\nu * \rho_n)}{d(\nu * \rho_n)}.$$

Moreover, let $\rho^n \mathscr{L}^d \to \delta_0$ narrowly as $n \to \infty$, so that $\nu^n \to \nu$ narrowly. As in the previous case, the argument begins with adding and subtracting in (2.11) the term

$$\int g \left[\int_s^t \overline{\mathcal{L}}_r f \circ e_r \, dr \right] d\boldsymbol{\eta}^n - \int g \left[\int_s^t \overline{\mathcal{L}}_r f \circ e_r \, dr \right] d\boldsymbol{\eta},$$

where $\overline{\mathcal{L}} = \mathcal{L}(\overline{a}, \overline{b})$ is a diffusion operator on $\mathbb{R}^d$, with continuous and compactly supported coefficients $\overline{a}$, $\overline{b}$. Write $\overline{\omega}$ for a common (bounded and continuous) modulus of continuity for $\overline{a}$, $\overline{b}$.

As in the previous case, the difference terms above are infinitesimal as $n \to \infty$, by narrow convergence of $\boldsymbol{\eta}^n$, entailing therefore a bound from above for the absolute value of (2.11), as $n \to \infty$, in terms of

$$\limsup_{n\to\infty} \int \left| \mathcal{L}^n f - \overline{\mathcal{L}} f \right| d\nu^n + \int \left| \mathcal{L} f - \overline{\mathcal{L}} f \right| d\nu.$$

We claim that

$$\lim_{n\to\infty} \int \left| \overline{\mathcal{L}}^n f - \overline{\mathcal{L}} f \right| d\nu^n = 0,$$

where $\overline{\mathcal{L}}^n$ is the diffusion on $\mathbb{R}^d$ with coefficients

$$\overline{a}^n := \frac{d(\overline{a}\nu * \rho_n)}{d(\nu * \rho_n)}, \qquad \overline{b}^n := \frac{d(\overline{b}\nu * \rho_n)}{d(\nu * \rho_n)}.$$

Indeed, recalling that $\|f\|_{C^{1,2}} \le 1$, we estimate

$$\begin{aligned}
\int \left| \overline{\mathcal{L}}^n f - \overline{\mathcal{L}} f \right| d\nu^n &\le \int |\overline{a}^n(x) - \overline{a}(x)| \, d\nu^n + \int |\overline{b}^n - \overline{b}| \, d\nu^n \\
&= \int |(\overline{a}\nu * \rho_n)(x) - \overline{a}(x)(\nu * \rho_n)(x)| \, dx + \int \left| (\overline{b}\nu * \rho_n)(x) - \overline{b}(x)(\nu * \rho_n)(x) \right| dx \\
&\le 2 \int \left[\int \overline{\omega}(y - x) \rho_n(y - x) dx \right] \nu(dy) = 2 \int \overline{\omega}(z) \rho_n(z) dz \to 0.
\end{aligned}$$

Thanks to this fact, we write

$$\begin{aligned}
\limsup_{n\to\infty} \int \left| \mathcal{L}^n f - \overline{\mathcal{L}} f \right| d\nu^n &= \limsup_{n\to\infty} \int \left| \mathcal{L}^n f - \overline{\mathcal{L}}^n f \right| d\nu^n \\
&= \limsup_{n\to\infty} \int_B |a^n - \overline{a}^n| + \int |b^n - \overline{b}^n| d\nu^n \\
&\le \int_B |a - \overline{a}| + |b - \overline{b}| d\nu
\end{aligned}$$

where in the last step we apply (2.6) (whose validity does not rely on smoothness assumption on ρ) and B is any bounded set containing the support of f.

A similar and actually easier argument gives that

$$\int \left|\mathcal{L}f - \overline{\mathcal{L}}f\right| d\nu \le \int_B |a - \overline{a}| + |b - \overline{b}|\, d\nu$$

as well. To conclude that (2.11) is infinitesimal, it is sufficient to rely on the density of continuous, compactly supported functions in the space $L^1(\chi_B \nu)$.

2.3 Proof of the superposition principle

In this section, we address the proof of our general superposition superposition principle for diffusions in $\mathbb{R}^d$, i.e. Theorem 2.14 below, by using Theorem 2.1 as a base case and iterating the approximation-tightness-limit scheme.

2.3.1 Case of bounded diffusions

We extend the validity of the superposition principle to case of solutions to Fokker-Planck equations associated to diffusion operators $\mathcal{L}(a, b)$ with bounded coefficients. This result already provides a slight extension of [Figalli, 2008, Theorem 2.6], as uniform bounds are only imposed on the variable $x \in \mathbb{R}^d$.

Theorem 2.12 (superposition for bounded diffusions)**.** *Let a, b be Borel maps as in* (1.1)*, with*

$$\int_0^T \sup_{x \in \mathbb{R}^d} \left[|a_t(x)| + |b_t(x)|\right] dt < \infty. \tag{2.14}$$

Then, the superposition principle holds for every weak solution $\nu = (\nu_t)_{t \in (0,T)} \subseteq \mathscr{P}(\mathbb{R}^d)$ to the FPE (1.2)*.*

Proof. We follow the approximation-tightness-limit scheme discussed in the previous section.

Step 1 (approximation)**.** We argue by convolution with a Gaussian kernel ρ. For $\varepsilon \in (0,1)$, let $\rho^\varepsilon(x) = \varepsilon^n \rho(x/\varepsilon)$: notice that $\left|\nabla^i \rho^\varepsilon\right| \le C\varepsilon^{-2}$, for $i \in \{1, 2\}$, where C is some absolute constant.

Let $\nu^\varepsilon = \nu * \rho^\varepsilon$, which solves a FPE with respect to a diffusion operator with coefficients a^ε, b^ε satisfying (the correspondent of) (2.1), as a consequence of the last statement in Lemma 2.5. By Theorem 2.1, existence of superposition solutions $\boldsymbol{\eta}^\varepsilon \in \mathscr{P}(C([0,T]; \mathbb{R}^d))$ for the associated martingale problems follows.

Step 2 (tightness)**.** For $R \ge 1$, let $\chi_R : \mathbb{R}^d \to [0,1]$ be the usual cut-off function (as in Remark 1.3) and, for $i \in \{1, \ldots, d\}$, let $x^i_R(x) := x_i \chi_R \in \mathscr{A}$. For any $p \in (1, \infty)$ and $\delta \in (1/p, 1)$, Corollary 2.11 with $\varphi^i_R := x^i_R \circ e_t$ provides a bound for the energy $\mathcal{A}(\varphi^i_R)$ in terms of

$$\ell = \left[\partial_t + \mathcal{L}^\varepsilon\right] x^i_R, \quad \alpha = 2a^\varepsilon(\nabla x^i_R, \nabla x^i_R).$$

Since $\partial_t x^i_R = 0$, $\left\|\nabla x^i_R\right\|_\infty$ is bounded and $\left\|\nabla^2 x^i_R\right\|_\infty$ is infinitesimal as $R \to \infty$, we deduce a bound for the energy $\mathcal{A}(\varphi^i)$, where $\varphi^i_t = x^i \circ e_t$, in terms of the quantities

$$\ell = (b^\varepsilon)^i, \quad \alpha = 2(a^\varepsilon)^{i,i}.$$

We notice that (2.8) combined with (2.6) entails that these bound are uniform for $\varepsilon \in (0,1)$. Since the measures ν_0^ε are tight, as they narrowly converge towards ν_0, there exists a coercive functional $\theta : \mathbb{R}^d \to [0,\infty]$ such that $\sup_{\varepsilon\in(0,1)} \int \theta d\nu_0^\varepsilon < \infty$. Introducing the coercive functional

$$C([0,T];\mathbb{R}^d) \ni \gamma \mapsto \Theta(\gamma) := \theta(\gamma_0) + \sum_{i=1}^d \mathcal{A}(\gamma^i),$$

we conclude that $\int \Theta d\boldsymbol{\eta}^\varepsilon$ is uniformly bounded for $\varepsilon \in (0,1)$, thus $\boldsymbol{\eta}^\varepsilon$ is tight.

Step 3 (limit)**.** This step is fully covered by the discussion in the previous section. □

2.3.2 Case of locally bounded diffusions

We extend our result from uniform bounds to local bounds on the coefficients.

Theorem 2.13 (superposition for locally bounded diffusions)**.** *Let a, b be Borel maps as in* (1.1) *such that*

$$\int_0^T \sup_{x\in B} \left[|a_t(x)| + |b_t(x)|\right] dt < \infty, \quad \textit{for every bounded borel } B \subseteq \mathbb{R}^d, \tag{2.15}$$

Then, the superposition principle holds for every weak solution $\nu = (\nu_t)_{t\in(0,T)} \subseteq \mathscr{P}(\mathbb{R}^d)$ to the FPE (1.2) *such that, for some $p \in (1,\infty)$, it holds*

$$\int_0^T \int \left[|b|^p + |a|^p\right] d\nu_t dt < \infty. \tag{2.16}$$

Proof. **Step 1** (approximation)**.** We argue here by push-forward via smooth maps. For $R \geq 1$, let χ_R be a cut-off function as in Remark 1.3 and let $\pi_R : \mathbb{R}^d \mapsto \mathbb{R}^d$ be the map

$$\pi_R(x) = x\chi_R(x), \quad \text{so that } \pi_R^i(x) = x^i\chi_R(x) \in C_c^2(\mathbb{R}^d).$$

By (2.15), it holds

$$\left|\mathcal{L}(\pi_R^i)\right| \leq \left\|\pi_R^i\right\|_{C^2} \sup_{|x|\leq 2R} \left[|a(x)| + |b(x)|\right], \quad \text{for } x \in \mathbb{R}^d,\ i \in \{1,\dots d\},$$

and similarly

$$\left|a(\nabla\pi_R^i, \nabla\pi_R^j)\right| \leq \left\|\pi_R^i\right\|_{C^1} \sup_{|x|\leq 2R} |a(x)| \quad \text{for } x \in \mathbb{R}^d,\ i,j \in \{1,\dots d\}.$$

Since conditional expectations reduce norms, we deduce that $\nu^R := \pi^R(\nu)$ solves a FPE associated to a diffusion on $\mathbb{R}^d$, whose coefficients a^R, b^R satisfy (2.14): Theorem 2.12 provides correspondent superposition solutions $\boldsymbol{\eta}^R$.

Step 2 (tightness)**.** We argue similarly as in the proof of Theorem 2.12 (with the same notation), but at fixed p and $\delta \in (1/p, 1)$. This leads to a bound for the energy $\mathcal{A}(\gamma^i)$, for any $i \in \{1,\dots,d\}$, in terms of

$$\ell^R := (b^R)^i, \quad \text{and} \quad \alpha^R := 2(a^R)^{i,i}.$$

Jensen's inequality and (2.16) entail that these bounds are uniform as $R \to \infty$, so that tightness follows at once.

Step 3 (limit)**.** Here, we rely on the deductions in Section 2.2.3, for the case of push-forward of measures, because that all the assumptions therein are fulfilled in this situation. The only fact to notice is that, since the limit map π is the identity map, the optimization for $\overline{a}$, $\overline{b}$ can be performed. □

2.3.3 General case

We finally prove the superposition principle for diffusions in $\mathbb{R}^n$, assuming only the bound (2.16), for some $p \in (1, \infty)$.

Theorem 2.14 (superposition for diffusions in $\mathbb{R}^d$)**.** *Let a, b be Borel maps as in* (1.1)*. Then, the superposition principle holds for every weak solution $\nu = (\nu_t)_{t\in(0,T)} \subseteq \mathscr{P}(\mathbb{R}^d)$ to the FPE* (1.2) *such that, for some $p \in (1, \infty)$, it holds*

$$\int [|a|^p + |b|^p] \, d\nu < \infty.$$

Proof. **Step 1** (approximation)**.** We perform the approximation procedure by convolution, exactly as in the proof of Theorem 2.12. This provides measures $(\nu^\varepsilon)_\varepsilon$ that solve FPE's with respect to diffusions satisfying the assumptions of Theorem 2.13, as a consequence of Lemma 2.5. We consider superposition solutions $(\boldsymbol{\eta}^\varepsilon)_\varepsilon$ to the correspondent martingale problem.

Step 2 (tightness)**.** As a consequence of Corollary 2.11 and arguing as in the final part of the correspondent step in the proof of Theorem 2.12, we deduce again that the family $(\boldsymbol{\eta}^\varepsilon)_\varepsilon$ is tight.

Step 3 (limit)**.** We argue as in the correspondent step in the proof of Theorem 2.12. □

Remark 2.15. One could combine all the arguments in the three sections above and deduce Theorem 2.14 at once from Theorem 2.1. We choose to argue by establishing first Theorem 2.12 and then Theorem 2.13 in order to clarify the different approximation procedures involved. Indeed, the crucial improvement with respect to [Figalli, 2008, Theorem 2.6] is to pass from bounded to locally bounded coefficients, which is a rather delicate step if one only knows that the superposition principle holds for diffusions with smooth and bounded coefficients. In Section 2.1 and particularly in Theorem 2.1, uniform bounds play an important role, but in the deterministic case, one is able to deal directly with locally smooth coefficients (compare with [Ambrosio et al., 2008, Proposition 8.1.8]), essentially because deterministic paths either go to infinity, i.e., the solution explodes in a finite time, or stay in a compact set. Roughly speaking, the source of difficulties in the stochastic case is that we have to deal with averages of such behaviours: indeed the solution to a truly stochastic martingale problem is expected to instantaneously "diffuse" over every compact set, of course with small probability as the sets become large. ■

Part II

Diffusions processes in metric measure spaces

Chapter 3

The metric measure space setting

In this chapter, we introduce the general framework where we study diffusion processes: we consider spaces endowed with a sufficiently rich structure, allowing for basic calculus operations, in particular where gradients and Laplacians of functions, vector fields and diffusions operators can be suitably defined.

We are led therefore to consider, as in [Ambrosio and Trevisan, 2014], Polish spaces and symmetric Dirichlet forms, enjoying carré du champ operators, as abstract, possibly infinite dimensional, Riemannian manifold-like spaces. Let us remark that our framework provides a rigorous foundation to the so-called Γ-calculus, very close to that developed e.g. in the recent monograph [Bakry et al., 2014]. The formulation introduced is both rigorous and flexible, as it allows for dealing with finite and infinite dimensional space at the same time: this is further clarified in Part IV. Moreover, it is strongly linked with the theory of Markov processes, via Dirichlet forms, for which we refer mainly to [Bouleau and Hirsch, 1991], but see also [Fukushima et al., 2011] or [Ma and Röckner, 1992]; and also with the growing field of analysis on metric measure spaces, at least when the infinitesimal structure is Riemannian, see [Ambrosio et al., 2014b].

We can rigorously summarize our framework as follows: we let (X, τ) be a Polish topological space, endowed with a σ-finite Borel measure $\mathfrak{m}$ with full support, i.e., $\operatorname{supp} \mathfrak{m} = X$, and

$$\begin{gathered}\text{a strongly local, densely defined and symmetric Dirichlet form } \mathcal{E} \text{ on } L^2(X, \mathscr{B}(X), \mathfrak{m}) \\ \text{enjoying a carré du champ } \Gamma : D(\mathcal{E}) \times D(\mathcal{E}) \to L^1(X, \mathscr{B}(X), \mathfrak{m}) \text{ and} \\ \text{generating a Markov semigroup } (\mathsf{P}_t)_{t\geq 0} \text{ on } L^2(X, \mathscr{B}(X), \mathfrak{m}).\end{gathered} \tag{3.1}$$

The precise meaning of (3.1) is recalled in Section 3.1, while in Section 3.2, we introduce and study basic properties of Sobolev spaces of functions. Most of the results are well-known and for their proof we frequently refer to the first chapter of [Bouleau and Hirsch, 1991].

3.1 Notation and abstract setup

To keep notation simple, we write $L^p(\mathfrak{m})$ (or even L^p_x) instead of $L^p(X, \mathscr{B}(X), \mathfrak{m})$ and denote $L^p(\mathfrak{m})$ norms by $\|\cdot\|_p$. We also write $L^0(\mathfrak{m})$ for the space of $\mathfrak{m}$-a.e. equivalence classes of Borel functions $f : X \mapsto [-\infty, +\infty]$ that take finite values $\mathfrak{m}$-a.e. in X. Sums and intersections of Lebesgue spaces will be constantly used, i.e. $L^p(\mathfrak{m})+L^q(\mathfrak{m})(= L^p_x+L^q_x)$, $L^p\cap L^q(\mathfrak{m})(= L^p_x\cap L^q_x)$, for p, $q \in [1,\infty]$: these spaces are endowed with natural Banach norms, denoted respectively with $\|\cdot\|_{L^p_x+L^q_x}$ and $\|\cdot\|_{L^p_x\cap L^q_x}$. Since $\mathfrak{m}$ is σ-finite, these spaces are separable for p, $q \in [1,\infty)$, and dual spaces are obtained exchanging intersections with sums (and exponents with their duals), in a natural way.

3.1.1 Dirichlet form and carré du champ

A symmetric Dirichlet form $\mathcal{E}$ is a $L^2(\mathfrak{m})$-lower semicontinuous quadratic form satisfying the Markov property

$$\mathcal{E}(\eta\circ f)\le \mathcal{E}(f) \qquad \text{for every normal contraction } \eta:\mathbb{R}\to\mathbb{R}, \tag{3.2}$$

i.e., a 1-Lipschitz map satisfying $\eta(0) = 0$. We refer to the already quoted monographs [Bouleau and Hirsch, 1991, Fukushima et al., 2011] for equivalent formulations of (3.2). Recall that

$$\mathbb{V} := \mathrm{D}(\mathcal{E}) \subset L^2(\mathfrak{m}), \quad \text{endowed with } \|f\|^2_{\mathbb{V}} := \|f\|_2^2 + \mathcal{E}(f)$$

is a Hilbert space. Furthermore, $\mathbb{V}$ is separable because $L^2(\mathfrak{m})$ is separable, see [Ambrosio et al., 2014b, Lemma 4.9] for the simple proof.

We still denote by $\mathcal{E}(\cdot,\cdot):\mathbb{V}\times\mathbb{V}\to\mathbb{R}$ the associated continuous and symmetric bilinear form

$$\mathcal{E}(f,g) := \frac{1}{4}\Big(\mathcal{E}(f+g)-\mathcal{E}(f-g)\Big).$$

We will assume strong locality of $\mathcal{E}$, namely

$$\forall\, f,\, g\in\mathbb{V}, (f+a)g = 0, \mathfrak{m}\text{-a.e. in } X, \text{ for some } a\in\mathbb{R}, \quad \Rightarrow \quad \mathcal{E}(f,g)=0.$$

It is possible to prove [Bouleau and Hirsch, 1991, Proposition I.2.3.2] that $\mathbb{V}\cap L^\infty(\mathfrak{m})$ is an algebra with respect to pointwise multiplication, so that for every $f\in\mathbb{V}\cap L^\infty(\mathfrak{m})$ the linear form on $\mathbb{V}\cap L^\infty(\mathfrak{m})$

$$\Gamma[f;\varphi] := 2\mathcal{E}(f,f\varphi)-\mathcal{E}(f^2,\varphi), \quad \varphi\in\mathbb{V}\cap L^\infty(\mathfrak{m}), \tag{3.3}$$

is well defined and, for every normal contraction $\eta:\mathbb{R}\to\mathbb{R}$, it satisfies [Bouleau and Hirsch, 1991, Proposition I.2.3.3]

$$0\le\Gamma[\eta\circ f;\varphi]\le\Gamma[f;g]\le\|\varphi\|_\infty\,\mathcal{E}(f) \quad \text{for all } f,\,\varphi\in\mathbb{V}\cap L^\infty(\mathfrak{m}),\,\varphi\ge 0. \tag{3.4}$$

The inequality (3.4) shows that for every nonnegative $\varphi\in\mathbb{V}\cap L^\infty(\mathfrak{m})$ the function $f\mapsto\Gamma[f;\varphi]$ is a quadratic form in $\mathbb{V}\cap L^\infty(\mathfrak{m})$ which satisfies the Markov property and can be extended by continuity to $\mathbb{V}$.

We assume that for all $f\in\mathbb{V}$ the linear form $\varphi\mapsto\Gamma[f;\varphi]$ can be represented by an absolutely continuous measure, with respect to $\mathfrak{m}$, with density $\Gamma(f)\in L^1_+(\mathfrak{m})$, the so-called carré du champ. Since $\mathcal{E}$ is strongly local, [Bouleau and Hirsch, 1991, Theorem I.6.1.1] yields the representation formula

$$\mathcal{E}(f,f) = \int_X \Gamma(f)\,d\mathfrak{m}, \quad \text{for all } f\in\mathbb{V}. \tag{3.5}$$

It is not difficult to check that Γ as defined by (3.5) (see e.g. [Bouleau and Hirsch, 1991, Definition I.4.1.2]) is a quadratic continuous map defined in $\mathbb{V}$ with values in $L^1_+(\mathfrak{m})$, and that $\Gamma[f-g;\varphi] \geq 0$ for all $\varphi \in \mathbb{V} \cap L^\infty(\mathfrak{m})$ with $\varphi \geq 0$, yields

$$|\Gamma(f,g)| \leq \sqrt{\Gamma(f)}\sqrt{\Gamma(g)}, \qquad \mathfrak{m}\text{-a.e. in } X. \tag{3.6}$$

We use the Γ notation also for the symmetric, bilinear and continuous map

$$\Gamma(f,g) := \frac{1}{4}\Big(\Gamma(f+g) - \Gamma(f-g)\Big) \in L^1(\mathfrak{m}) \qquad f,\, g \in \mathbb{V},$$

which, thanks to (3.5), represents the bilinear form $\mathcal{E}$ by the formula

$$\mathcal{E}(f,g) = \frac{1}{2}\int_X \Gamma(f,g)d\mathfrak{m}, \quad \text{for all } f,\, g \in \mathbb{V}.$$

Because of the Markov property and locality, Γ satisfies the chain rule [Bouleau and Hirsch, 1991, Corollary I.7.1.2]

$$\Gamma(\eta(f),g) = \eta'(f)\Gamma(f,g) \quad \text{for all } f,\, g \in \mathbb{V},\, \eta : \mathbb{R} \to \mathbb{R} \text{ Lipschitz with } \eta(0) = 0, \tag{3.7}$$

and the Leibniz rule:

$$\Gamma(fg,h) = f\Gamma(g,h) + g\Gamma(f,h) \qquad \text{for all } f,\, g,\, h \in \mathbb{V} \cap L^\infty(\mathfrak{m}).$$

Notice that by [Bouleau and Hirsch, 1991, Theorem I.7.1.1] (3.7) is well defined, since for every Borel set $N \subset \mathbb{R}$ (as the set where η is not differentiable) one has

$$\mathscr{L}^1(N) = 0 \quad \Rightarrow \quad \Gamma(f) = 0 \quad \mathfrak{m}\text{-a.e. on } f^{-1}(N).$$

3.1.2 Laplace operator and Markov semigroup

The Dirichlet form $\mathcal{E}$ induces a densely defined, negative and selfadjoint operator $\Delta : D(\Delta) \subset \mathbb{V} \to L^2(\mathfrak{m})$, via the integration by parts formula $\mathcal{E}(f,g) = -\int_X g\,\Delta f d\mathfrak{m}$ for all $g \in \mathbb{V}$. The operator Δ is of "diffusion" type, since it satisfies the following chain rule for every $\eta \in C^2(\mathbb{R})$ with $\eta(0) = 0$ and bounded first and second derivatives, see [Bouleau and Hirsch, 1991, Corollary I.6.1.4]: whenever $f \in D(\Delta)$ with $\Gamma(f) \in L^2(\mathfrak{m})$, then $\eta(f) \in D(\Delta)$ and

$$\Delta\eta(f) = \eta'(f)\Delta f + \eta''(f)\Gamma(f)\,. \tag{3.8}$$

The "heat flow" P_t associated to $\mathcal{E}$ is well defined starting from any initial condition $f \in L^2(\mathfrak{m})$. Recall that in this framework the heat flow $(\mathsf{P}_t)_{t\geq 0}$ is an analytic Markov semigroup and that $f^t = \mathsf{P}_t f$ can be characterized as the unique C^1 map $f : (0,\infty) \to L^2(\mathfrak{m})$, with values in $D(\Delta)$, satisfying

$$\begin{cases} \dfrac{d}{dt} f^t = \Delta f^t & \text{for } t \in (0,\infty), \\ \lim\limits_{t\downarrow 0} f^t = f & \text{in } L^2(\mathfrak{m}). \end{cases}$$

Because of this, Δ can equivalently be characterized in terms of the strong convergence $(\mathsf{P}_t f - f)/t \to \Delta f$ in $L^2(\mathfrak{m})$ as $t \downarrow 0$.

We have the regularization estimates, in the more general context of gradient flows of convex functionals, see for instance [Ambrosio et al., 2008, Theorem 4.0.4(ii)])

$$\mathcal{E}(\mathsf{P}_t f) \leq \inf_{v\in\mathbb{V}}\big\{\mathcal{E}(v,v) + \frac{\|v-f\|_2^2}{2t}\big\} < \infty, \qquad \forall t>0,\ f\in L^2(\mathfrak{m}), \tag{3.9}$$

$$\|\Delta \mathsf{P}_t f\|_2^2 \leq \inf_{v\in D(\Delta)}\big\{\|\Delta v\|_2^2 + \frac{\|v-f\|_2^2}{t^2}\big\} < \infty, \qquad \forall t>0,\ u\in L^2(\mathfrak{m}). \tag{3.10}$$

One useful consequence of the Markov property is the L^p contraction of $(\mathsf{P}_t)_{t\geq 0}$ from $L^2\cap L^p(\mathfrak{m})$ to $L^2\cap L^p(\mathfrak{m})$. By density in L^p, for $p\in[1,\infty)$, this allows to extend uniquely P_t to a strongly continuous semigroup of linear contractions in $L^p(\mathfrak{m})$, $p\in[1,\infty)$, for which we retain the same notation. Furthermore, $(\mathsf{P}_t)_{t\geq 0}$ is sub-Markovian (cf. [Bouleau and Hirsch, 1991, Proposition I.3.2.1]), since it preserves one-sided essential bounds, namely $f\leq C$ (resp. $f\geq C$) $\mathfrak{m}$-a.e. in X for some $C\geq 0$ (resp. $C\leq 0$) implies $\mathsf{P}_t f\leq C$ (resp. $\mathsf{P}_t f\geq C$) $\mathfrak{m}$-a.e. in X for all $t\geq 0$.

It is easy to check, using L^1-contractivity of P, that the dual semigroup $\mathsf{P}_t^\infty : L^\infty(\mathfrak{m}) \to L^\infty(\mathfrak{m})$ given by

$$\int g\mathsf{P}_t^\infty f d\mathfrak{m} = \int f\mathsf{P}_t g d\mathfrak{m}, \qquad f\in L^\infty(\mathfrak{m}),\ g\in L^1(\mathfrak{m})$$

is well defined. It is a contraction semigroup in $L^\infty(\mathfrak{m})$, sequentially weak-* continuous, and it coincides with P on $L^2\cap L^\infty(\mathfrak{m})$.

Similar continuity and contraction properties can be established for intersections $L^p\cap L^q(\mathfrak{m})$ and sums $L^p(\mathfrak{m})+L^q(\mathfrak{m})$ of Lebesgue spaces, for p, $q\in[1,\infty]$.

3.2 Spaces $\mathbb{V}^p$ and $D^p(\Delta)$

In this section, we study subspaces of $\mathbb{V}$ and $D(\Delta)$, where stronger integrability conditions are imposed.

3.2.1 Spaces $\mathbb{V}^p$

We define

$$\mathbb{V}^p := \left\{u\in\mathbb{V}\cap L^p(\mathfrak{m}) : \int(\Gamma(u))^{p/2}d\mathfrak{m}<\infty\right\}, \qquad \text{for } p\in[1,\infty),$$

with the obvious extension to $p=\infty$. As in [Bouleau and Hirsch, 1991, §I.6.2], one endow $\mathbb{V}^p$ with the norm

$$\|f\|_{\mathbb{V}^p} = \|f\|_{\mathbb{V}} + \|f\|_p + \|\sqrt{\Gamma(u)}\|_p,$$

obtaining a Banach space, akin to the intersection of classical Sobolev spaces $W^{1,2}\cap W^{1,p}$. We notice that $\mathbb{V}^2=\mathbb{V}$, with an equivalent norm, and the inclusion $\mathbb{V}^p\subseteq\mathbb{V}^q$ holds whenever $2\leq p\leq q\leq\infty$ or $1\leq q\leq p\leq 2$.

The following result is a useful criterion to deduce convergence in $\mathbb{V}^p$.

Proposition 3.1. *Let $p\in[1,\infty)$, $f\in\mathbb{V}$, $(f_n)_n\subseteq\mathbb{V}^p$ satisfy $f_n\to f$ in $L^2\cap L^p(\mathfrak{m})$,*

$$\Gamma(f-f_n)\to 0,\ \mathfrak{m}\text{-a.e. in } X \quad \text{and} \quad \|\sqrt{\Gamma(f_n)}\|_{L^2\cap L^p}\to\|\sqrt{\Gamma(f)}\|_{L^2\cap L^p},\ \text{as } n\to\infty.$$

Then, $f\in\mathbb{V}^p$ and $f_n\to f$ in $\mathbb{V}^p$.

Proof. The proof is a straightforward application of Fatou's lemma and the triangle inequality

$$|\sqrt{\Gamma(f_n)} - \sqrt{\Gamma(f)}| \le \sqrt{\Gamma(f_n - f)} \le \sqrt{\Gamma(f_n)} + \sqrt{\Gamma(f)}, \quad \mathfrak{m}\text{-a.e. in } X.$$

Indeed this inequality implies $\Gamma(f_n) \to \Gamma(f)$, $\mathfrak{m}$-a.e. in X, as $n \to \infty$ and

$$(\Gamma(f_n - f))^{p/2} \le 2^{p-1} \left[(\Gamma(f_n))^{p/2} + (\Gamma(f))^{p/2} \right].$$

Fatou's lemma gives

$$\begin{aligned} 2^p \int (\Gamma(f))^{p/2} d\mathfrak{m} &\le \liminf_{n\to\infty} \int \left\{ 2^{p-1} \left[(\Gamma(f_n))^{p/2} + (\Gamma(f))^{p/2} \right] - (\Gamma(f_n - f))^{p/2} \right\} d\mathfrak{m} \\ &\le 2^p \int (\Gamma(f))^{p/2} d\mathfrak{m} - \limsup_{n\to\infty} \int (\Gamma(f_n - f))^{p/2} d\mathfrak{m}, \end{aligned}$$

from which the thesis follows, arguing also for $p = 2$. □

Remark 3.2. The argument above is similar to a classical lemma by F. Riesz, entailing convergence in $L^p(\mathfrak{m})$ for any sequence of functions $(f_n)_n \subseteq L^p(\mathfrak{m})$ such that $f_n \to f$, $\mathfrak{m}$-a.e. in X and $\|f_n\|_p \to \|f_n\|_p$, as $n \to \infty$, for $p \in [1, \infty)$. ■

We introduce here a family of inequalities which play a key role in Part III, dealing with uniqueness for Fokker-Planck equations. They provide a smoothing effect for P in the spaces $\mathbb{V}^p$ and their validity corresponds, in the smooth setting, to integral bounds on the gradient of the kernel of P, see Chapter 11. Here, we provide the definition and some basic consequences of their validity, which however we do not assume in all what follows, remarking explicitly when it is the case.

Definition 3.3 (L^p-Γ inequalities)**.** *Let $p \in [1, \infty]$. We say that the L^p-Γ inequality holds if there exists $c_p^\Gamma \ge 0$ satisfying*

$$\|\sqrt{\Gamma(\mathsf{P}_t f)}\|_p \le \frac{c_p^\Gamma}{\sqrt{t}} \|f\|_p, \quad \text{for every } f \in L^2 \cap L^p(\mathfrak{m}),\ t \in (0,1). \tag{3.11}$$

Although the L^p-Γ inequality is expressed for $t \in (0,1)$, from its validity and L^p contractivity of P, we easily deduce that

$$\|\sqrt{\Gamma(\mathsf{P}_t f)}\|_p \le c_p^\Gamma (t \wedge 1)^{-1/2} \|f\|_p, \quad \text{for every } f \in L^2 \cap L^p(\mathfrak{m}),\ t \in (0, \infty). \tag{3.12}$$

Notice also that (3.9) shows that the the L^2-Γ inequality always holds, with $c_2^\Gamma = 1/\sqrt{2}$. By Marcinkiewicz interpolation, if the L^p-Γ inequality holds, then the L^q-Γ inequality holds as well, for every q between 2 and p.

Other straightforward consequences of the validity of L^p-Γ are that, for every $t > 0$, the operator

$$L^2 \cap L^p(\mathfrak{m}) \ni f \mapsto \mathsf{P}_t f \in \mathbb{V}^p$$

is well defined and continuous and that the space $\mathbb{V}^p$ is dense in $L^2 \cap L^p(\mathfrak{m})$. Moreover, for every $f \in L^2 \cap L^p(\mathfrak{m})$ the curve $(0, \infty) \ni t \mapsto \mathsf{P}_t f$ is continuous with values in V_p and $t \mapsto \sqrt{t}\mathsf{P}_t f \in \mathbb{V}^p$ is uniformly bounded, for $t \in (0,1)$. Since it converges to 0 in $L^2 \cap L^p(\mathfrak{m})$ as $t \downarrow 0$, it is natural to ask whether a similar convergence, with respect to the stronger norm on $\mathbb{V}^p$, holds. The following proposition reduces the problem to continuity at 0 for the map $t \mapsto \mathsf{P}_t f \in \mathbb{V}^p$, when $f \in \mathbb{V}^p$.

Proposition 3.4. *Let $p \in [1,\infty)$, let the L^p-Γ inequality hold and assume that $t \mapsto \mathsf{P}_t f \in \mathbb{V}^p$ is continuous at 0, for every $f \in \mathbb{V}^p$. Then,*

$$\lim_{t\downarrow 0} \sqrt{t}\, \|\mathsf{P}_t f\|_{\mathbb{V}^p} = 0, \quad \textit{for every } f \in L^2 \cap L^p(\mathfrak{m}).$$

Proof. This is a standard density and uniform boundedness argument. Clearly, if $f \in \mathbb{V}^p$ the limit holds since $\|\mathsf{P}_t f\|_{\mathbb{V}^p} \to \|f\|_{\mathbb{V}^p}$, as $t \downarrow 0$. Given $f \in L^2 \cap L^p(\mathfrak{m})$ and $g \in \mathbb{V}^p$, we have, for every $t \in (0,1)$,

$$\sqrt{t}\, \|\mathsf{P}_t f\|_{\mathbb{V}^p} \le \sup_{s\in(0,1)} \sqrt{s}\, \|\mathsf{P}_s(f-g)\|_{\mathbb{V}^p} + \sqrt{t}\, \|\mathsf{P}_t g\|_{\mathbb{V}^p}\,,$$

thus, for some constant c depending on c_p^Γ only, it holds

$$\limsup_{t\downarrow 0} \sqrt{t}\, \|\mathsf{P}_t f\|_{\mathbb{V}^p} \le c\, \|f-g\|_{L^2\cap L^p}\,.$$

By density of $\mathbb{V}^p$ in $L^2 \cap L^p(\mathfrak{m})$, the thesis follows. □

To study continuity at 0, let us first consider the case $p = 2$, for which it always holds, since the energy decreases, i.e. $\mathcal{E}(\mathsf{P}_t f) \le \mathcal{E}(f)$, for $t \ge 0$, e.g. by (3.10) with $v = f$, which also gives $\mathcal{E}(\mathsf{P}_t f) \to \mathcal{E}(f)$, as $t \downarrow 0$. In particular, the curve $(\mathsf{P}_t f)_t$ is bounded in the Hilbert space $\mathbb{V}$, thus it weakly converges to f as $t \downarrow 0$, arguing by density of $D(\Delta)$ in $\mathbb{V}$. Moreover, $\|\mathsf{P}_t f\|_{\mathbb{V}} \to \|f\|_{\mathbb{V}}$, thus $\mathsf{P}_t f \to f$ strongly in $\mathbb{V}$. The general case $p \in [1,\infty)$ is handled similarly, using Fatou's lemma and Proposition 3.1 in place of Hilbert space arguments, and we obtain the following necessary and sufficient condition for continuity at 0 of the curve $t \mapsto \mathsf{P}_t f$, for $f \in \mathbb{V}^p$:

$$\limsup_{t\downarrow 0} \|\sqrt{\Gamma\left(\mathsf{P}_t f\right)}\|_p \le \|\sqrt{\Gamma\left(f\right)}\|_p. \tag{3.13}$$

3.2.2 Spaces $D^p(\Delta)$

We define

$$D^p(\Delta) := \{f \in D(\Delta) \cap L^p(\mathfrak{m}) :\ \Delta f \in L^p(\mathfrak{m})\}\,, \qquad \text{for } p \in [1,\infty),$$

with the obvious extension for $p = \infty$. We endow $D^p(\Delta)$ with the norm

$$\|f\|_{D^p(\Delta)} = \|f\|_{L^2\cap L^p} + \|\Delta f\|_{L^2\cap L^p}\,,$$

obtaining a Banach space. Notice that, for $p \in [1,\infty)$, $D^p(\Delta)$ is only contained in the domain of the generator of P in $L^p(\mathfrak{m})$, characterized in terms of strong convergence for $(\mathsf{P}_t f - f)/t$ in $L^p(\mathfrak{m})$, as $t \downarrow 0$. Indeed, for every $f \in D(\Delta)$ it holds

$$\mathsf{P}_t f - f = \int_0^t \Delta \mathsf{P}_s f ds = \int_0^t \mathsf{P}_s(\Delta f) ds,$$

where the integral above is the sense of Bochner. By the L^p-contraction property for P, it holds $\|\mathsf{P}_s(\Delta f)\|_p \le \|\Delta f\|_p$, for every $s \ge 0$, thus the right hand side divided by t is bounded and it actually converges in $L^p(\mathfrak{m})$ to Δf. For $p = \infty$ one obtains only weak-* convergence.

For $p \in [1, \infty]$, we consider the following analogue of the L^p-Γ inequality, that reads an a L^p version of (3.10), that we refer as the L^p-Δ inequality:

$$\|\Delta \mathsf{P}_t f\|_p \le \frac{c_p^\Delta}{t} \|f\|_p\,, \quad \text{for every } f \in L^2 \cap L^p(\mathfrak{m}) \text{ and every } t \in (0,1). \tag{3.14}$$

Quite differently from the L^p-Γ inequality, it turns out that the L^p-Δ inequality always holds for $p \in (1, \infty)$ and can be obtained as a consequence of the fact that P is analytic [Stein, 1970, Theorem III.1]: it is actually equivalent to it, see [Yosida, 1995, §X.10]. Let us remark that, in some settings, e.g. for the standard heat semigroup in Euclidean spaces, see Chapter 11, the L^p-Δ inequality for $p = \infty$ holds and further consequences could be drawn.

A direct consequence of (3.14) is the following estimate.

Corollary 3.5. *Let $p \in [1, \infty]$ and let c_p^Δ denote the constant in (3.14). Then*

$$\|\mathsf{P}_t f - \mathsf{P}_{t-t'} f\|_p \le \min\left\{ c_p^\Delta \log\left(1 + \frac{t'}{t-t'}\right), 2\right\} \|f\|_p\,, \quad \textit{for every } f \in L^2 \cap L^p(\mathfrak{m}).$$

for every t, $t' \in (0,1)$, with $t' \le t$.

Proof. The estimate with the constant 2 follows from L^p-contractivity. For the other one, we apply (3.14) as follows:

$$\|\mathsf{P}_t f - \mathsf{P}_{t-t'} f\|_p \le \int_0^{t'} \|\Delta \mathsf{P}_{t-t'+r} f\|_p \, dr \le \int_0^{t'} \frac{c_p^\Delta}{t-t'+r} dr\, \|f\|_p = c_p^\Delta \log\left(1 + \frac{t'}{t-t'}\right) \|f\|_p\,.$$

□

We conclude this section by providing an analogue of Proposition 3.4, whose proof goes along the same lines, but uses the fact that continuity at 0 for the map $t \mapsto \mathsf{P}_t f \in D^p(\Delta)$ is straightforward, for $f \in D^p(\Delta)$, because of strong continuity of P and the commutation $\Delta \mathsf{P}_t f = \mathsf{P}_t(\Delta f)$.

Proposition 3.6. *For every $p \in (1, \infty)$, it holds*

$$\lim_{t \downarrow 0} t \|\mathsf{P}_t f\|_{D^p(\Delta)} = 0, \quad \textit{for every } f \in L^2 \cap L^p(\mathfrak{m}).$$

Chapter 4

Derivations and diffusion operators

In this chapter, we introduce and study suitable notions of vector fields, by means of derivations (Section 4.2), of maps with values in symmetric, non-negative, matrices (Section 4.3) and then of diffusions operators (Section 4.4).

Our main assumption is the existence of some algebra of functions $\mathscr{A}$, that we regard as test functions on X, enjoying suitable stability and density properties, for which we perform a detailed study in Section 4.1. Diffusion operators that are then defined on $\mathscr{A}$ and subsequently extended to larger domains, provided that a priori bounds and density results for $\mathscr{A}$ hold: as the reader may expect, density is crucial especially dealing with uniqueness. Besides linearity, another remarkable feature is locality, so the usual rules of calculus for diffusion operators hold.

4.1 The algebra $\mathscr{A}$

In all what follows, we assume that an algebra $\mathscr{A} \subseteq L^1 \cap L^\infty(\mathfrak{m})$ is prescribed, with

$$\Phi(f_1, \ldots, f_n) \in \mathscr{A} \quad \text{whenever } \Phi \in C_b^2(\mathbb{R}^n) \text{ with } \Phi(0) = 0 \text{ and } f_1, \ldots, f_n \in \mathscr{A}, \tag{4.1}$$

for every $n \geq 1$.

Let us remark that further assumptions on $\mathscr{A}$ are to be imposed below, in particular dealing with the time-dependent setting, see Chapter 5.

Notice that $\mathscr{A}$ is an algebra also directly from (4.1): this stability property is rather useful to extend the validity of suitable density assumptions, as we investigate below.

Proposition 4.1. *If $\mathscr{A}$ is dense in $L^2(\mathfrak{m})$, then it is also dense in $L^p \cap L^q(\mathfrak{m})$, for $1 \leq p \leq q < \infty$, and weakly-* dense in $L^p \cap L^\infty(\mathfrak{m})$, for $p \in (1, \infty]$.*

Proof. The argument relies on the following two facts. First, we consider a sequence of functions $\Phi_k \in C_b^2(\mathbb{R})$ with $\Phi_k(0) = 0$, $|\Phi_k(z)| \leq |z|$, for $k \geq 1$ and $\Phi_k(z) \to z$ for every $z \in \mathbb{R}$, as $k \to \infty$. For $p \in [1, \infty)$ and $f \in L^p$, $\Phi_k(f)$ converges to f in L^p, by Lebesgue theorem (and weakly-* in $L^\infty(\mathfrak{m})$, because of pointwise and uniformly bounded convergence).

As a second fact, we notice that if $(f_n)_n \subseteq L^2(\mathfrak{m})$ converges to f in $L^2(\mathfrak{m})$ and $\Phi \in C_b^2(\mathbb{R})$ satisfies $\Phi(0) = \Phi'(0) = 0$, then $\Phi(f_n) \to f_n$ in $L^p(\mathfrak{m})$ for every $p \in [1, \infty)$ and weakly-* in $L^\infty(\mathfrak{m})$. For $p \in [1, \infty)$, it is sufficient to notice that the assumptions entail, up to a subsequence, $\Phi(f_n) \to \Phi(f)$ $\mathfrak{m}$-a.e. in X and $|\Phi(x)| \leq c\,|x|^{2/p}$ (for some constant $c \geq 0$

depending on Φ only) so that by Lebesgue theorem, $\|\Phi(f_n)\|_p \to \|\Phi(f)\|_p$ and by Remark 3.2 we conclude. Again, weak-* convergence in L^∞ follows by pointwise and uniformly bounded convergence.

To conclude, since $L^2 \cap L^p \cap L^q(\mathfrak{m})$ is dense in $L^p \cap L^q(\mathfrak{m})$, it is sufficient to approximate any f in the former space with a sequence in $\mathscr{A}$. Furthermore, by the first fact and a diagonal argument, it is sufficient to approximate $\Phi_k(f)$, for any fixed $k \geq 1$. By density of $\mathscr{A}$ in $L^2(\mathfrak{m})$, we consider a sequence $(f_n)_n \subseteq \mathscr{A}$ which converges to f in $L^2(\mathfrak{m})$ and set $g_n := \Phi_k(g_n) \in \mathscr{A}$, which converges to $\Phi_k(f)$ in $L^p \cap L^q(\mathfrak{m})$ by the second fact. To show weak-* convergence in $L^p \cap L^\infty(\mathfrak{m})$, we argue similarly. $\square$

A similar result can be proved starting from density of $\mathscr{A}$ in $\mathbb{V}$, with the notable difference that density in the spaces $\mathbb{V}^p$ is obtained only for $1 \leq p \leq 2$: as it is intuitively clear, it is not possible to improve regularity simply by composition with smooth functions.

Proposition 4.2. *If $\mathscr{A}$ is contained and dense in $\mathbb{V}$, it is also dense in $\mathbb{V}^p$, for $p \in [1,2]$.*

Proof. The argument is a variant of the previous proof. Regarding the first fact, we notice that, for $f \in \mathbb{V}^p$, with $1 \leq p \leq 2$, the functions $\Phi_k(f)$ converge to f in $\mathbb{V}^p$, by the chain rule $\sqrt{\Gamma(\Phi_k(f))} = |\Phi_k'(f)| \sqrt{\Gamma(f)}$ and Proposition 3.1.

In the second step, we notice that if $f_n \in \mathbb{V}$ is a sequence converging to f in $\mathbb{V}$ then, for any $\Phi \in C^2_b(\mathbb{R})$, with $\Phi(0) = \Phi'(0)$, then $\Phi(f_n) \to f_n$ in $\mathbb{V}^p$ for every $p \in [1,2]$ again by the chain rule and Proposition 3.1.

The conclusion is then identical, being sufficient to approximate any function of the form $\Phi(f)$ in $\mathbb{V}^p$, where $f \in \mathbb{V}^p$, with a sequence of functions in $\mathscr{A}$. $\square$

Intersection spaces $D^p(\Delta) \cap \mathbb{V}^{2p}$ can be considered as well.

Proposition 4.3. *If $\mathscr{A}$ is dense in $D(\Delta) \cap \mathbb{V}^4$, then it is contained and dense in $D^p(\Delta) \cap \mathbb{V}^{2p}$, for $p \in [1,2]$.*

Proof. Again, regarding the first fact, given $f \in D^p(\Delta) \cap \mathbb{V}^{2p}$, with $1 \leq p \leq 2$, the functions $\Phi_k(f)$ converge to f in $\mathbb{V}^p$, by the chain rule for Γ, $\sqrt{\Gamma(\Phi_k(f))} = |\Phi_k'(f)| \sqrt{\Gamma(f)}$, the chain rule for Δ stated in (3.8), Proposition 3.1 and Remark 3.2. We argue similarly for the second fact and then conclude. $\square$

Clearly, analogues can be considered, deducing from density in $\mathbb{V}^q$, for any $q \in [1,\infty)$ and in intersection spaces $D^q(\Delta) \cap \mathbb{V}^{2q}$, entailing densities in spaces with smaller exponents.

Remark 4.4. Under the additional condition that

$$\mathscr{A} \text{ is invariant under the action of } \mathsf{P}, \text{ i.e. } \mathsf{P}_t\mathscr{A} \subseteq \mathscr{A} \text{ for every } t > 0, \tag{4.2}$$

the density assumption of $\mathscr{A}$ in $\mathbb{V}$ can be weakened to density of $\mathscr{A}$ in $L^2(\mathfrak{m})$; indeed, standard semigroup theory shows that an invariant subspace is dense in $\mathbb{V}$ if and only if it is dense in $L^2(\mathfrak{m})$, see for instance [Ambrosio et al., 2014b, Lemma 4.9].

For $p \in [1,\infty)$, density of $\mathscr{A}$ in $L^2 \cap L^p(\mathfrak{m})$, (4.2) and the L^p-Γ inequality entail density in $\mathbb{V}^p$, provided that $\mathsf{P}_t f \to f$ in $\mathbb{V}^p$ as $t \downarrow 0$, for every $f \in \mathbb{V}^p$, or equivalently if (3.13) holds for every $f \in \mathbb{V}^p$. Indeed, this continuity assumption and the L^p-Γ inequality entail that $\bigcup_{t>0} \mathsf{P}_t(\mathcal{A})$ is dense in $\mathbb{V}^p$, for every set $\mathcal{A}$, dense in $L^2 \cap L^p(\mathfrak{m})$, as one can approximate $f \in \mathbb{V}^p$ by $\mathsf{P}_t f$ for $t > 0$ small enough, and then approximate $\mathsf{P}_t f$ with $\mathsf{P}_t g$ in $L^2 \cap L^p(\mathfrak{m})$, for $g \in \mathcal{A}$. ■

4.2 Derivations

This section follows closely [Ambrosio and Trevisan, 2014, §2], where vector fields are introduced as derivations, i.e. linear operators on $\mathscr{A}$ satisfying a pointwise upper bound in terms of Γ.

Throughout this section, we assume for simplicity of exposition that

$$\mathscr{A} \text{ is dense in } \mathbb{V}, \tag{4.3}$$

which, by the results in the previous section, entails various density properties for $\mathscr{A}$.

Definition 4.5 (derivations)**.** *A derivation is a linear operator $\boldsymbol{b} : \mathscr{A} \to L^0(\mathfrak{m})$, $f \mapsto df(\boldsymbol{b})$, satisfying*

$$|df(\boldsymbol{b})| \leq h\sqrt{\Gamma(f)}, \quad \mathfrak{m}\text{-a.e. in } X, \text{ for every } f \in \mathscr{A},$$

for some $h \in L^0(\mathfrak{m})$. The smallest function h (in the $\mathfrak{m}$-a.e. sense) with this property is denoted by $|\boldsymbol{b}|$.

For $q \in [1, \infty]$, we write $\boldsymbol{b} \in L^q$ if $|\boldsymbol{b}| \in L^q(\mathfrak{m})$, and similarly for intersections and sum of Lebesgue spaces, $L^p \cap L^q(\mathfrak{m})$, $L^p(\mathfrak{m}) + L^q(\mathfrak{m})$, for p, $q \in [1, \infty]$. Clearly, $|\boldsymbol{b}|$ is the $\mathfrak{m}$-essential supremum among all functions $f \in \mathscr{A}$ of the expression $|df(\boldsymbol{b})| / \sqrt{\Gamma(f)}$ (set equal to 0 on $\{\Gamma(f) = 0\}$).

Linearity and the $\mathfrak{m}$-a.e. upper bound are sufficient to entail "locality" and thus Leibniz and chain rules, with proof akin to that of [Ambrosio and Kirchheim, 2000, Theorem 3.5]. We point out also the recent work [Gigli, 2014], where derivations are introduced in a similar setting, and their structure is deeply investigated. For our purpose, which is to study diffusion processes by means of Fokker-Planck equations and martingale problems, few basic properties of derivations are sufficient, and for completeness we prove them, in this section.

Proposition 4.6 (Leibniz and chain rules)**.** *Let $\boldsymbol{b}$ be a derivation and let $\Phi \in C^2_b(\mathbb{R}^n, \mathbb{R})$, with $\Phi(0) = 0$. Then, for any $\boldsymbol{f} = (f_1, \ldots, f_n) \in \mathscr{A}^n$, one has*

$$d(\Phi \circ \boldsymbol{f})(\boldsymbol{b}) = \sum_{i=1}^n \partial_i \Phi(\boldsymbol{f}) df_i(\boldsymbol{b}) = \nabla\Phi(\boldsymbol{f}) \cdot d\boldsymbol{f}(\boldsymbol{b}), \quad \mathfrak{m}\text{-a.e. in } X, \tag{4.4}$$

where we let $d\boldsymbol{f}(\boldsymbol{b}) := (df_1(\boldsymbol{b}), \ldots, df_n(\boldsymbol{b}))$. In particular, for every f, $g \in \mathscr{A}$, it holds

$$d(fg)(\boldsymbol{b}) = df(\boldsymbol{b})g + f dg(\boldsymbol{b}), \quad \mathfrak{m}\text{-a.e. in } X.$$

Proof. Since $\Phi(\boldsymbol{f}) \in \mathscr{A}$, the terms in (4.4) are well-defined as elements in $L^0(\mathfrak{m})$, thus we only have to prove that they coincide. To this aim, let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \mathbb{Q}^n$, apply the chain rule in $\mathbb{V}$ and use bilinearity of Γ to obtain the identity

$$\begin{aligned} \Gamma(\Phi(\boldsymbol{f}) - \boldsymbol{\lambda} \cdot \boldsymbol{f}) &= \sum_{i,j=1}^n (\partial_i \Phi(\boldsymbol{f}) - \lambda_i) \Gamma(f_i, f_j) (\partial_j \Phi(\boldsymbol{f}) - \lambda_j) = \\ &= [\Gamma(\boldsymbol{f})] \left(\nabla\Phi(\boldsymbol{f}) - \boldsymbol{\lambda}, \nabla\Phi(\boldsymbol{f}) - \boldsymbol{\lambda}\right) \end{aligned}$$

where we let $[\Gamma(\boldsymbol{f})]^{i,j} := \Gamma(f_i, f_j)$, for i, $j \in \{1, \ldots, n\}$. Notice that $[\Gamma(\boldsymbol{f})]$ is $\mathfrak{m}$-a.e. valued in the space of $n \times n$, non-negative symmetric matrices, $\mathrm{Sym}_+(\mathbb{R}^n)$.

From (4.4), we deduce the inequality

$$|d\Phi(\boldsymbol{f})(\boldsymbol{b}) - \boldsymbol{\lambda}\cdot \boldsymbol{f}(\boldsymbol{b})| \le |\boldsymbol{b}| \sqrt{[\Gamma(\boldsymbol{f})]\,(\nabla\Phi(\boldsymbol{f})-\boldsymbol{\lambda}, \nabla\Phi(\boldsymbol{f})-\boldsymbol{\lambda})}, \quad \mathfrak{m}\text{-a.e. in } X.$$

By sub-additivity, the same inequality holds for every $\boldsymbol{\lambda} \in \mathbb{Q}^n$, for every $x \in A$, where A^c is $\mathfrak{m}$-negligible. Given $x \in A$, we choose a sequence $\boldsymbol{\lambda}_k \to \nabla\Phi(\boldsymbol{f}(x))$, as $k \to \infty$, and we conclude that

$$|d\Phi(\boldsymbol{f})(\boldsymbol{b})(x) - \nabla\Phi(\boldsymbol{f}(x))\cdot [d\boldsymbol{f}(\boldsymbol{b})]\,(x)| = 0.$$

Leibniz rule follows letting $n = 2$ and $\Phi(x_1, x_2) = x_1 x_2$. □

Remark 4.7 (derivations $u\boldsymbol{b}$)**.** Let $\boldsymbol{b}$ be a derivation and let $u \in L^0(\mathfrak{m})$. Then, the operator $u\boldsymbol{b}$, defined on $\mathscr{A}$ by $f \mapsto u df(\boldsymbol{b})$ is a derivation, with $|u\boldsymbol{b}| \le |u||\boldsymbol{b}|$. In particular, if $\boldsymbol{b} \in L^q(\mathfrak{m})$ and $u \in L^r(\mathfrak{m})$, with $q^{-1} + r^{-1} \le 1$, then $u\boldsymbol{b} \in L^{s'}(\mathfrak{m})$, where $(s')^{-1} = q^{-1} + r^{-1}$, i.e. $q^{-1} + r^{-1} + s^{-1} = 1$. and let $u \in L^r(\mathfrak{m})$, with $q^{-1} + r^{-1} \le 1$. Then, $\varphi \mapsto u df(\boldsymbol{b})$ defines a derivation $u\boldsymbol{b}$ in $L^{s'}$, where $(s')^{-1} = q^{-1} + r^{-1}$, i.e. $q^{-1} + r^{-1} + s^{-1} = 1$. By linearity, similar remarks apply when $\boldsymbol{b}$ is a derivation in $L^p \cap L^q(\mathfrak{m})$ or $L^p(\mathfrak{m}) + L^q(\mathfrak{m})$. ■

Example 4.8 (gradient derivations)**.** Every $g \in \mathbb{V}$ induces a derivation $\boldsymbol{b}_g$,

$$\mathscr{A} \ni f \mapsto df(\boldsymbol{b}_g) := \Gamma(f, g). \tag{4.5}$$

These derivations belong to $L^2(\mathfrak{m})$, as (3.6) yields $|\boldsymbol{b}_g| \le \sqrt{\Gamma(g)}$ (with equality assuming $\mathscr{A}$ to be dense in $\mathbb{V}$). If $g \in \mathbb{V}^p$, then $\boldsymbol{b}_g \in L^2 \cap L^p(\mathfrak{m})$.

By linearity, finite sums $\sum_i \chi_i \boldsymbol{b}_{g_i}$ with $\chi_i \in L^\infty(\mathfrak{m})$ and $g_i \in \mathbb{V}$, define derivations in $L^2(\mathfrak{m})$. ■

It is very useful to extend the action of a derivation from $\mathscr{A}$ to larger spaces.

Remark 4.9 (extension of derivations)**.** Let $q \in (1, \infty]$, r, $s \in (1, \infty)$ satisfy $q^{-1}+r^{-1}+s^{-1} = 1$. If $\mathscr{A}$ is dense in $\mathbb{V}^s$, then any derivation $\boldsymbol{b} \in L^q(\mathfrak{m})$ extends uniquely to a continuous linear operator $\boldsymbol{b} : \mathscr{A} \subseteq \mathbb{V}^s \to L^{r'}(\mathfrak{m})$, still denoted by $f \mapsto df(\boldsymbol{b})$, defined on $\mathbb{V}^s$, with values in the space $L^{r'}(\mathfrak{m})$, which still satisfies, for every $f \in \mathbb{V}^s$,

$$|df(\boldsymbol{b})| \le |\boldsymbol{b}| \sqrt{\Gamma(f)}, \qquad \mathfrak{m}\text{-a.e. in } X,$$

as well as Leibniz and chain rules, with respect to composition with functions $\Phi \in C^1_b(\mathbb{R}^n)$ such that $\Phi(0) = 0$. ■

Assuming only density of $\mathscr{A}$ in $\mathbb{V}^r$ and no integral bounds on $|\boldsymbol{b}|$, any derivation $\boldsymbol{b}$ may still be extended uniquely to a linear operator defined on $\mathbb{V}^r$, with values in $L^0(\mathfrak{m})$, continuous when the latter is endowed with convergence in measure. However, this extension is not useful for our purposes, as we often deal with integral functionals involving $df(\boldsymbol{b})$, defined initially on $\mathscr{A}$, which are not continuous with respect to this topology. An important example is $\operatorname{div}\boldsymbol{b}$, whose definition is given below.

Definition 4.10 (divergence of a derivation)**.** *Let $q \in [1, \infty]$, $\mathscr{A} \subset \mathbb{V}^{q'}$ and let $\boldsymbol{b} \in L^q(\mathfrak{m})$ be a derivation. The distributional divergence* $\operatorname{div}\boldsymbol{b}$ *is the linear operator on $\mathscr{A}$ defined by*

$$\mathscr{A} \ni f \mapsto [\operatorname{div}\boldsymbol{b}](f) := -\int df(\boldsymbol{b})d\mathfrak{m}.$$

We say that $\operatorname{div}\boldsymbol{b} \in L^q(\mathfrak{m})$ *if the distribution* $\operatorname{div}\boldsymbol{b}$ *is induced by* $g \in L^q(\mathfrak{m})$*, i.e.*

$$\int df(\boldsymbol{b})d\mathfrak{m} = -\int fgd\mathfrak{m}, \quad \textit{for all } f \in \mathscr{A}.$$

Similarly, we say that $\operatorname{div}\boldsymbol{b}^- \in L^q(\mathfrak{m})$ *if there exists a non-negative* $g \in L^q(\mathfrak{m})$ *such that*

$$\int df(\boldsymbol{b})d\mathfrak{m} \leq \int fgd\mathfrak{m}, \quad \textit{for all } f \in \mathscr{A},\ f \geq 0.$$

The condition $\mathscr{A} \subset \mathbb{V}^{q'}$ ensures integrability for $df(\boldsymbol{b})$. As for $|\boldsymbol{b}|$, $\operatorname{div}\boldsymbol{b}^-$ can be defined as the $\mathfrak{m}$-a.e. smallest non-negative function $g \in L^q(\mathfrak{m})$ for which the inequality above holds: existence follows by a simple convexity argument, because the class of admissible g's is convex and closed in $L^q(\mathfrak{m})$ (for $q = \infty$, it is weakly-* closed).

Example 4.11 (divergence of gradients)**.** The divergence of the gradient derivation induced by $g \in \mathbb{V}$ as in (4.5) coincides with the Laplacian Δg, still understood in distributional terms.

When the divergence of a derivation $\boldsymbol{b}$ is represented by some function, the problem of extension of $\boldsymbol{b}$ can be addressed more precisely, introducing Sobolev spaces of functions.

Definition 4.12 (spaces $W^p(\boldsymbol{b})$)**.** *Let* $q \in [1,\infty]$ *let* $\mathscr{A} \subseteq \mathbb{V}^{q'}$ *and* $\boldsymbol{b} \in L^q(\mathfrak{m})$ *be a derivation with* $\operatorname{div}\boldsymbol{b} \in L^q(\mathfrak{m})$*. Let* $p,\ s \in [q',\infty]$*, satisfy* $p^{-1} + q^{-1} + s^{-1} = 1$*, and let* $\mathscr{A} \subset \mathbb{V}^s$ *.*

We say that $f \in W^p(\boldsymbol{b})$ *if* $f \in L^p(\mathfrak{m})$ *and, for some* $h \in L^p(\mathfrak{m})$*, it holds*

$$\int hg = -\int f\,[dg(\boldsymbol{b}) + g\operatorname{div}\boldsymbol{b}]\,d\mathfrak{m}, \quad \textit{for every } g \in \mathscr{A}.$$

We write $df(\boldsymbol{b}) := h$*.*

When endowed with the norm $\|f\|_p + \|df(\boldsymbol{b})\|_p$, the space $W^p(\boldsymbol{b})$ is Banach and, if $(f_n)_n \in W^p(\boldsymbol{b})$ converges to f, $df_n(\boldsymbol{b}) \to h \in L^p(\mathfrak{m})$, in duality with $L^{p'}(\mathfrak{m})$, then $f \in W^p(\boldsymbol{b})$ and $df(\boldsymbol{b}) = h$.

Remark 4.13 (spaces $H^p(\boldsymbol{b})$)**.** As it happens with Sobolev spaces in Euclidean spaces, density of test functions (in this case, $\mathscr{A}$) is not always fulfilled. For $p \in [q',\infty)$, we let $H^p(\boldsymbol{b})$ be the closure of $\mathscr{A}$ in $W^p(\boldsymbol{b})$. By continuity of the terms in (4.4) with respect to convergence in $W^p(\boldsymbol{b})$, Leibniz and chain rules extend from $\mathscr{A}$ to $H^p(\boldsymbol{b})$. Moreover, when $p = q'$ and $f \in H^{q'}(\boldsymbol{b})$, it holds

$$\int df(\boldsymbol{b}) = -\int (\operatorname{div}\boldsymbol{b}) f d\mathfrak{m}, \quad \text{for every } f \in \mathbb{V}^{q'}.$$

Notice that Remark 4.9 gives the inclusion $\mathbb{V}^s \cap L^{r'}(\mathfrak{m}) \subseteq H^{r'}(\boldsymbol{b})$. ■

Finally, not only Definition 4.5, but also the notions above easily generalize replacing $L^q(\mathfrak{m})$ with sums or intersections of Lebesgue spaces.

4.3 2-tensors

In this section, we briefly introduce the bilinear generalizations of derivations, instrumental to define diffusion operators. The approach is still inspired by that of differential geometry: a matrix valued map $a : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ is described by the functional $(v, w) \mapsto a(v, w)$, for v, w vector fields. A basic example is the operator Γ, thus we try to adopt a consistent notation. Again, we assume that (4.3) holds.

Definition 4.14 (2-tensors)**.** *A 2-tensor is a bilinear operator*

$$\boldsymbol{a} : \mathscr{A} \times \mathscr{A} \to L^0(\mathfrak{m}), \quad (f, g) \mapsto \boldsymbol{a}(f, g),$$

such that

$$|\boldsymbol{a}(f, g)| \le h\sqrt{\Gamma(f)}\sqrt{\Gamma(g)}, \quad \mathfrak{m}\text{-a.e. in } X, \text{ for every } f,\, g \in \mathscr{A},$$

for some $h \in L^0(\mathfrak{m})$. The smallest function h with this property is denoted by $|\boldsymbol{a}|$.

For $q \in [1, \infty]$, we write $\boldsymbol{a} \in L^q(\mathfrak{m})$ if $|\boldsymbol{a}| \in L^q(\mathfrak{m})$, and similarly for sums and intersections of Lebesgue spaces.

A 2-tensor $\boldsymbol{a}$ is said to be symmetric if

$$\boldsymbol{a}(f, g) = \boldsymbol{a}(g, f) \quad \mathfrak{m}\text{-a.e. in } X\text{, for every } f,\, g \in \mathscr{A},$$

non-negative if

$$\boldsymbol{a}(f) := \boldsymbol{a}(f, f) \ge 0 \quad \mathfrak{m}\text{-a.e. in } X\text{, for every } f \in \mathscr{A}$$

λ-elliptic, for some $\lambda > 0$, if it is symmetric and it holds

$$\boldsymbol{a}(f) \ge \lambda\Gamma(f) \quad \mathfrak{m}\text{-a.e. in } X\text{, for every } f \in \mathscr{A},$$

and finally elliptic if it λ-elliptic for some $\lambda > 0$.

Example 4.15 (sum of squares)**.** *Besides the example $\boldsymbol{a}(f, g) = \Gamma(f, g)$, non-negative symmetric 2-tensors can be built as follows. Let $(\boldsymbol{b}_i)_{i\ge1}$ be a sequence of derivations with $\sum_i |\boldsymbol{b}_i|^2 < \infty$ $\mathfrak{m}$-a.e. in X, and define the 2-tensor $\boldsymbol{a} := \sum_{i=1}^{\infty} \boldsymbol{b}_i \otimes \boldsymbol{b}_i$ by*

$$\boldsymbol{a}(f, g) := \sum_{i=1}^{\infty} df(\boldsymbol{b}_i)dg(\boldsymbol{b}_i), \quad \text{for } f,\, g \in \mathscr{A}.$$

Notice that the series converges and it holds $|\boldsymbol{a}| \le \sum_{i=1}^{\infty} |\boldsymbol{b}_i|^2$, $\mathfrak{m}$-a.e. in X.

For every $g \in \mathscr{A}$, the map $\mathscr{A} \ni f \mapsto \boldsymbol{a}(f, g)$ defines a derivation, with $|\boldsymbol{a}(\cdot, dg)| \le |\boldsymbol{a}|\sqrt{\Gamma(g)}$. Thanks to this fact, large parts of the discussion on derivations above easily extends to the case of 2-tensors. For example, the chain rule entails, for any $\Phi \in C^1(\mathbb{R}^n, \mathbb{R})$ with $\Phi(0) = 0$, and $\boldsymbol{f} = (f_1, \ldots f_n) \in \mathscr{A}^n$, the identity

$$\boldsymbol{a}(d\Phi(\boldsymbol{f})) = \sum_{i,j=1}^{n} \boldsymbol{a}(f_i, f_j)\partial_i\Phi(\boldsymbol{f})\partial_j\Phi(\boldsymbol{f}) := [\boldsymbol{a}(\boldsymbol{f})](\nabla\Phi(\boldsymbol{f}), \nabla\Phi(\boldsymbol{f})), \quad \mathfrak{m}\text{-a.e. in } X,$$

where $[\boldsymbol{a}(\boldsymbol{f})]^{i,j} := \boldsymbol{a}(f_i, f_j)$, for $i, j \in \{1, \ldots, n\}$.

Remark 4.16. Let $q \in (1,\infty]$, r, $s \in (1,\infty)$ satisfy $q^{-1}+r^{-1}+s^{-1}=1$. If $\mathscr{A}$ is dense both in $\mathbb{V}^r$ and in $\mathbb{V}^s$, then any 2-tensor $\boldsymbol{a} \in L^q(\mathfrak{m})$ extends uniquely to a bilinear continuous operator, still denoted by $\boldsymbol{a}$, defined on $\mathbb{V}^r \times \mathbb{V}^s$, with values in $L^1(\mathfrak{m})$, which satisfies

$$|\boldsymbol{a}(f,g)| \leq |\boldsymbol{a}| \sqrt{\Gamma(f)}\sqrt{\Gamma(g)}, \qquad \mathfrak{m}\text{-a.e. in } X, \text{ for every } f \in \mathbb{V}^r,\ g \in \mathbb{V}^s.$$

Remark 4.17 (Dirichlet forms induced by 2-tensors)**.** Let $\boldsymbol{a} \in L^\infty(\mathfrak{m})$ be an elliptic symmetric 2-tensor and consider the bilinear form

$$\mathscr{A} \times \mathscr{A} \ni (f,g) \mapsto \mathcal{E}[\boldsymbol{a}](f,g) := \int \boldsymbol{a}(f,g) d\mathfrak{m}.$$

Since we assume that $\mathscr{A} \subseteq \mathbb{V}$ is dense, then $\boldsymbol{a}$ and $\mathcal{E}[\boldsymbol{a}]$ extend to continuous bilinear functional on $\mathbb{V} \times \mathbb{V}$. We claim that $(\mathcal{E}[\boldsymbol{a}], \mathbb{V})$ is a Dirichlet form on $L^2(\mathfrak{m})$, which induces a Markov semigroup $\mathsf{P}[\boldsymbol{a}]$. Indeed, ellipticity and the assumption $\boldsymbol{a} \in L^\infty(\mathfrak{m})$ entail that $\mathbb{V} \ni f \mapsto \|f\|_2^2 + \mathcal{E}[\boldsymbol{a}](f)$ defines a Hilbert norm equivalent to the natural one $\|f\|_{\mathbb{V}}$. Moreover, normal contractions operate on $\mathcal{E}[\boldsymbol{a}]$ because of the chain rule for $\boldsymbol{a}$. One may even go further and show that conditions (3.1) hold, replacing $\mathcal{E}$ with $\mathcal{E}[\boldsymbol{a}]$, with $\Gamma[\boldsymbol{a}](f) := \boldsymbol{a}(f)$, for $f \in \mathbb{V}$. Ellipticity and the assumption $\boldsymbol{a} \in L^\infty(\mathfrak{m})$ give that the spaces $\mathbb{V}^p[\boldsymbol{a}]$ (naturally built with respect to $\Gamma[\boldsymbol{a}]$) coincide with $\mathbb{V}^p$, for $p \in [1,\infty]$. Let us remark, however, that this is not necessary the case for the spaces $D^p(\Delta[\boldsymbol{a}])$, which may not coincide with $D^p(\Delta)$, even for $p=2$. ■

4.4 Diffusion operators

The main difficulty with introducing diffusion operators in the metric measure space setting is due to the fact that we prefer not to introduce Hessians of functions, although it would be certainly possible, e.g. as in [Bakry et al., 2014] or [Gigli, 2014], assuming curvature lower bounds. On the other side, Dirichlet forms allow for the introduction of a Laplacian operator almost immediately from their definition. Moreover, in the Euclidean space $\mathbb{R}^d$, one recovers the coefficients a of a diffusion operator $f \mapsto (a : \nabla^2) f$ simply choosing $f(x) = x^i$, $g(x) = x^j$, for i, $j \in \{1, \ldots, d\}$ in the identity

$$a : \nabla^2(fg) - (a : \nabla^2 f) g - f(a : \nabla^2 g) = 2a(\nabla f, \nabla g), \tag{4.6}$$

the same which defines the carré du champ Γ in terms of Δ, compare with (3.3). Therefore, our strategy consists in introducing diffusion operators in such a way that (4.6) holds.

Definition 4.18 (diffusion operators)**.** *A diffusion operator is a linear map $\mathcal{L} : \mathscr{A} \to L^0(\mathfrak{m})$ such that*

$$\mathscr{A} \times \mathscr{A} \ni (f,g) \mapsto \frac{1}{2}\left[\mathcal{L}(fg) - \mathcal{L}(f)g - f\mathcal{L}(g)\right] \tag{4.7}$$

is a non-negative symmetric 2-tensor.

Given a diffusion operator $\mathcal{L}$, we write $\boldsymbol{a}$ the 2-tensor defined by (4.7) (one should write $\boldsymbol{a}[\mathcal{L}]$, but there is no danger of confusion in what follows).

Example 4.19. *If $\mathscr{A} \subseteq D(\Delta)$, then $\mathcal{L} := \Delta$ is a diffusion operator with $\boldsymbol{a} = \Gamma$. Another example is $\mathcal{L} := \boldsymbol{b}$, for a derivation $\boldsymbol{b}$, so that $\boldsymbol{a} = 0$, by Leibniz rule. More generally, diffusion operators are stable with respect to sums and multiplications with non-negative functions $a \in L^0(\mathfrak{m})$, e.g. $\mathcal{L} := a\Delta + \boldsymbol{b}$ defines a diffusion operator, with associated 2-tensor $a\Gamma$.*

An important difference between diffusion operators and derivations is that we do not impose pointwise bounds directly on $\mathcal{L}f$, as they would require at least the introduction of Hessians for functions in $\mathscr{A}$, unless we deal with special cases, as in the example above. Instead, to extend calculus for diffusion operators, we require some continuity or closability assumption.

Example 4.20 (extension of diffusion operators)**.** *Let $q \in (1,\infty]$, r, $s \in (1,\infty)$ satisfy $q^{-1}+r^{-1}+s^{-1}=1$. If $\mathscr{A}$ is dense in $D^s(\Delta)$ and, for some constant $c \geq 0$, it holds*

$$\|\mathcal{L}(f)\|_{r'} \leq c\,\|f\|_{D^s(\Delta)}\,,$$

then $\mathcal{L}$ extends uniquely to a linear continuous operator on $D^s(\Delta)$, with values in $L^{r'}(\mathfrak{m})$. Other spaces may be considered as well, e.g. $\mathbb{V}^s$, or $D^s(\Delta)\cap\mathbb{V}^s$.

We are in the situation of the example above when $\mathcal{L} = a\Delta$, for $\mathscr{A} \subseteq D^r(\Delta)$ and $a \in L^q(\mathfrak{m})$. When $\mathcal{L} = \boldsymbol{b}$ is a derivation, with $\boldsymbol{b} \in L^q(\mathfrak{m})$, if $\mathscr{A}$ is dense in $\mathbb{V}^r$, we recover Remark 4.9.

Definition 4.21 (divergence)**.** *Let $\mathcal{L}$ be a diffusion operator such that $\mathcal{L}f \in L^1(\mathfrak{m})$, for every $f \in \mathscr{A}$. The distributional divergence* $\operatorname{div}\mathcal{L}$ *is defined as the linear functional*

$$\mathscr{A} \ni f \mapsto [\operatorname{div}\mathcal{L}](f) := -\int \mathcal{L}(f)d\mathfrak{m}.$$

For $q \in [1,\infty]$, we say that $\operatorname{div}\mathcal{L} \in L^q(\mathfrak{m})$ if there exists $g \in L^q(\mathfrak{m})$ such that

$$\int \mathcal{L}(f)d\mathfrak{m} = -\int fgd\mathfrak{m}, \quad \textit{for all } f \in \mathscr{A}.$$

Similarly, we say that $\operatorname{div}\boldsymbol{b}^- \in L^q(\mathfrak{m})$ if there exists a non-negative $g \in L^q(\mathfrak{m})$ such that

$$\int \mathcal{L}(f)d\mathfrak{m} \leq \int fgd\mathfrak{m}, \quad \textit{for all } f \in \mathscr{A},\ f \geq 0.$$

Example 4.22. *Let $a \in D^q(\Delta)$, $\mathscr{A} \subseteq D^{q'}(\Delta)$ and define $\mathcal{L}f := a\Delta f$, for $f \in \mathscr{A}$. Then, it holds $\operatorname{div}\mathcal{L} = -\Delta a \in L^2 \cap L^q(\mathfrak{m})$,*

By combining the definition of $\boldsymbol{a}$ and $\operatorname{div}\mathcal{L}$, if $\mathcal{L}$ maps $\mathscr{A}$ into $L^1(\mathfrak{m})$, we deduce the identity

$$\int \mathcal{L}(f)gd\mathfrak{m} = -\left[\operatorname{div}\mathcal{L}\right](fg) - \int \left[f\mathcal{L}(g) + 2\boldsymbol{a}(f,g)\right]d\mathfrak{m}, \quad \text{for all } f,\, g \in \mathscr{A}, \tag{4.8}$$

which could be used to provide an extension of $\mathcal{L}$, in a similar way as in Definition 4.12. In the next proposition, we use it to prove the chain rule for diffusion operators.

Proposition 4.23 (chain rule for diffusion operators)**.** *Let $\mathcal{L}$ be a diffusion operator, with $\mathcal{L}f \in L^1(\mathfrak{m})$ for every $f \in \mathscr{A}$. Let p, q, $s \in [1,\infty]$ satisfy $p^{-1}+q^{-1}+s^{-1}=1$, and $\boldsymbol{a}$,* $\operatorname{div}\mathcal{L} \in L^q(\mathfrak{m})$ *and assume that $\mathscr{A} \subseteq \mathbb{V}^p \cap \mathbb{V}^s$.*

For any $\Phi \in C^2(\mathbb{R}^n)$ with $\Phi(0)=0$ and $\boldsymbol{f} = (f_1,\ldots,f_n) \in \mathscr{A}^n$, it holds

$$\mathcal{L}(\Phi(\boldsymbol{f})) = \sum_{i,j=1}^n \partial^2_{i,j}\Phi(\boldsymbol{f})\boldsymbol{a}(f_i,f_j) + \sum_{i=1}^n \partial_i\Phi(\boldsymbol{f})\mathcal{L}(f_i), \quad \mathfrak{m}\textit{-a.e. in } X. \tag{4.9}$$

We may also write $\mathcal{L}(\Phi(\boldsymbol{f})) = \boldsymbol{a}(\boldsymbol{f}) : \nabla^2\Phi(\boldsymbol{f}) + \mathcal{L}(\boldsymbol{f}) \cdot \nabla\Phi(\boldsymbol{f})$, if we denote $[\boldsymbol{a}(\boldsymbol{f})]^{i,j} = \boldsymbol{a}(f_j, f_j)$, $\mathcal{L}(\boldsymbol{f})^i = \mathcal{L}(f_i)$ for i, $j \in \{1, \ldots, n\}$.

Proof. Notice first that both terms in (4.9) belong to $L^1(\mathfrak{m})$, thus it is sufficient to show that they coincide. By induction and identity (4.7), the chain rule (4.9) holds whenever Φ is a polynomial in n variables, with $\Phi(0) = 0$. To obtain the general case, we let $(p_k)_{k\geq 1}$ be a sequence of polynomial functions, with $p_k(0) = 0$, converging towards Φ locally uniformly towards in $C^2(\mathbb{R}^n)$. The assumptions, together with Proposition 3.1 entail that $p_k(\boldsymbol{f}) \to \Phi(\boldsymbol{f})$ in $\mathbb{V}^p \cap \mathbb{V}^s$ and the right hand side in (4.9), with p_k in place of Φ converges to some limit h in $L^1(\mathfrak{m})$, as $k \to \infty$. On the other side, we let $g \in \mathscr{A}$ be any function and consider (4.8), with $p_k(\boldsymbol{f})$ in place of f, which reads as

$$\int \mathcal{L}(p_k(\boldsymbol{f}))g d\mathfrak{m} = -\int p_k(\boldsymbol{f}) \left[g \operatorname{div} \mathcal{L} + \mathcal{L}g\right] + 2\boldsymbol{a}(p_k(\boldsymbol{f}), g) d\mathfrak{m}.$$

Passing to the limit in this expression, we deduce

$$\int hg d\mathfrak{m} = -\int \Phi(\boldsymbol{f}) \left[g \operatorname{div} \mathcal{L} + \mathcal{L}g\right] + 2\boldsymbol{a}(\Phi(\boldsymbol{f}), g) d\mathfrak{m} = \int \mathcal{L}(\Phi(\boldsymbol{f}))g d\mathfrak{m},$$

and being $g \in \mathscr{A}$ arbitrary, (4.9) holds. □

We conclude this section by introducing a class of diffusion operators that can be expressed in divergence form; the main example being $\mathcal{L} = \Delta$ and suitable perturbations. These play an important role especially in Chapter 9, since the are particularly well-suited for applications of Hilbert-space techniques, see also [Le Bris and Lions, 2008]. On the other side, let us remark that our general theory, in particular the validity of the superposition principle (Chapter 7), holds for diffusion operators not necessarily in divergence form.

Definition 4.24 (diffusion operators in divergence form)**.** *Let $\mathcal{L}$ be a diffusion operator such that $\mathcal{L}(f) \in L^{r'}(\mathfrak{m})$ for $f \in \mathscr{A}$, and assume that the associated 2-tensor $\boldsymbol{a} \in L^q(\mathfrak{m})$, and $\mathscr{A} \subseteq \mathbb{V}^s$, for q, r, $s \in [1, \infty]$, with $q^{-1} + r^{-1} + s^{-1} = 1$.*

We say that $\mathcal{L}$ is in divergence form if there exists a derivation $\boldsymbol{b} \in L^q(\mathfrak{m})$ such that, for every $f \in \mathscr{A}$, $g \in \mathbb{V}^r$ it holds

$$\int \mathcal{L}(f)g d\mathfrak{m} = -\int \boldsymbol{a}(f, g) d\mathfrak{m} + df(\boldsymbol{b})g d\mathfrak{m}.$$

If $\mathbb{V}^s$ is dense in $L^s(\mathfrak{m})$, then the derivation $\boldsymbol{b}$ above is unique. Notice that, when $\mathcal{L}$ is written in divergence form, then $\operatorname{div} \mathcal{L} = \operatorname{div} \boldsymbol{b}$.

Example 4.25. *Let q, r, $s \in [1, \infty]$, satisfy $q^{-1}+r^{-1}+s^{-1} = 1$, assume that $\mathscr{A} \subseteq D^r(\Delta) \cap \mathbb{V}^r$ and consider the diffusion $\mathcal{L} := a\Delta$. Then, $\mathcal{L}$ is in divergence form if $a \in \mathbb{V}^q$, and we let $\boldsymbol{b}$ be the gradient derivation associated to a. Indeed, if $g \in \mathbb{V}^s$, then it holds $ag \in \mathbb{V}^{r'}$ [Bouleau and Hirsch, 1991, Proposition I.6.2.3] and*

$$\int (a\Delta f) g d\mathfrak{m} = -\int \Gamma(f, ag) d\mathfrak{m} = -\int \left[a\Gamma(f, g) + \Gamma(a, f)g\right] d\mathfrak{m}.$$

Chapter 5

Adding time to the framework

In this chapter, we extend the framework of chapters 3 and 4, allowing for a "time-dependent" setting, i.e. on the product space $(0,T)\times X$, for some $T>0$. There is indeed a straightforward way to perform this operation, i.e. simply by replacing the space X with the product $(0,T)\times X$, the measure $\mathfrak{m}$ with $\mathscr{L}^1\otimes\mathfrak{m}$, and the Dirichlet form $\mathcal{E}$ with

$$L^2(\mathscr{L}^1\otimes\mathfrak{m})\ni f\mapsto\int_0^T\mathcal{E}(f_t)dt\in[0,\infty], \tag{5.1}$$

where $f_t(x):=f(t,x)$, for $(t,x)\in(0,T)\times X$.

Although this extension is a rather natural operation, more details are provided in the next section. Then, in Section 5.2, we recall some technical facts on Sobolev and absolutely continuous curves with values in Banach spaces, which are useful in Part III.

5.1 The time-extended framework

For fixed $T\in(0,\infty)$ we endow the space $\tilde{X}:=X\times(0,T)$ with the product topology and the product measure $\tilde{\mathfrak{m}}:=\mathscr{L}^1\otimes\mathfrak{m}$. We introduce the notation L^p_t for integration over $(0,T)$, so that e.g. $L^p_t(L^p_x):=L^p((0,T);L^p(\mathfrak{m}))=L^p(\widetilde{\mathfrak{m}})$. An important role in our deductions is played by norms on Lebesgue spaces with different exponents with respect to the variables t and x, e.g. $L^\infty_t(L^p_x)$ or $L^1_t(L^p_x)$: this already provides a sufficient reason for which the time-extended framework is not fully recovered by the general metric measure space point of view on $(0,T)\times X$.

We endow the space $(\tilde{X},\widetilde{\mathfrak{m}})$ with the form $\tilde{\mathcal{E}}$ given by (5.1), which is clearly quadratic, lower semicontinuous with respect to convergence in $L^2_t(L^2_x)$, by Fatou's lemma, and that normal contractions operate, i.e. the analogue of (3.2) holds, for every 1-Lipschitz function $\eta:\mathbb{R}\mapsto\mathbb{R}$, with $\eta(0)=0$.

Actually, the form $\tilde{\mathcal{E}}$ satisfies all the assumptions (3.1), its domain being $L^2_t(\mathbb{V})$ and the carré du champ being $\tilde{\Gamma}(f)(t,\cdot):=\Gamma(f_t)(\cdot)$, $\mathscr{L}^1$-a.e. $t\in(0,T)$. It is also easy to prove that the heat semigroup associated to $\tilde{\mathcal{E}}$, that we denote by $\tilde{\mathsf{P}}$, acts on $f\in L^2_t(L^2_x)$ by $f\mapsto[\tilde{\mathsf{P}}_\alpha f](t,\cdot)=(\mathsf{P}_\alpha f_t)(\cdot)$, $\mathscr{L}^1$-a.e. $t\in(0,T)$, for $\alpha\in[0,\infty)$ (here and below, we obviously avoid to use the variable $t\in[0,\infty)$ as a subscript for semigroups). The domain of the generator $\tilde{\Delta}$ is clearly $D(\tilde{\Delta})=L^2_t(D(\Delta))$.

Spaces $\tilde{\mathbb{V}}^p$ and $D^p(\tilde{\Delta})$ can be defined as well, but we do not use them extensively, as we are more interested on their variants with respect to mixed norms, e.g. $L^1_t(\mathbb{V}^p)$, or $L^1_t(D^p(\Delta))$.

Thanks to the time-extended framework we may study more general derivations, 2-tensors and diffusion operators, provided that a suitable algebra of functions $\tilde{\mathscr{A}}$ is fixed. This is a rather delicate part, since one would like to introduce minimal conditions on $\tilde{\mathscr{A}}$, but at the same time require regularity with respect to t for elements in $\tilde{\mathscr{A}}$: so far we are simply "gluing together" the sections $\{t\} \times X$ in a Borel way for $t \in (0,T)$, but for studying diffusion processes, we need some regularity also with respect this variable, thus we summarize our standing assumptions on $\tilde{\mathscr{A}}$ as follows:

$$\begin{gathered}\Phi(\cdot, f_1, \ldots, f_n) \in \tilde{\mathscr{A}}, \text{ for } \Phi \in C_b^{1,2}((0,T) \times \mathbb{R}^n) \text{ with } \Phi(\cdot, 0) = 0 \text{ and } f_1, \ldots, f_n \in \tilde{\mathscr{A}},\ n \geq 1, \\ \tilde{\mathscr{A}} \subseteq L^2_t(\mathbb{V}) \text{ is dense, } \tilde{\mathscr{A}} \subseteq W^{1,2}((0,T); L^p_x) \\ \text{and for every } t \in [0,T], \text{ the image of } \tilde{\mathscr{A}} \ni f \mapsto f_t \in L^p(\mathfrak{m}), \text{ is dense for } p \in [1,\infty),\end{gathered} \tag{5.2}$$

Arguing as in Section 4.1, one obtains density of $\tilde{\mathscr{A}}$ in $L^p_t(L^p_x)$, for $p \in [1,\infty)$ as well as density in spaces with mixed norms. The Sobolev assumption (see Definition 5.2) on $t \mapsto f_t$, for $f \in \tilde{\mathscr{A}}$, instead of a more natural C^1_b assumption is introduced here to obtain more degrees of freedom and becomes useful e.g. when dealing with elliptic diffusion operators. Clearly, writing f_t at fixed $t \in [0,T]$, we always mean the continuous representative for f, see Proposition 5.6.

We let $\tilde{\mathscr{A}_c}$ be the class of functions $f \in \tilde{\mathscr{A}}$ such that $f_t = 0$, for $t \in \{0, T\}$. Thanks to (5.2), $\tilde{\mathscr{A}_c}$ is a class of functions large enough in order to deduce whether a curve belongs to some Sobolev space with respect to t, see Remark 5.3 in the next section.

Remark 5.1. An alternative, but not equivalent, approach to the time-dependent framework consists in defining a Dirichlet form as the closure of

$$C^1_b((0,T); L^2(\mathfrak{m})) \cap L^2((0,T); \mathbb{V}) \ni f \mapsto \int_0^T \left[\|\partial_t f_t\|_2^2 + \mathcal{E}(f_t) \right] dt.$$

However, the variable $t \in (0,T)$ plays a distinguished role in all what follows, that we prefer to highlight from its very introduction. ■

5.2 Sobolev and absolutely continuous Banach-valued curves

In this section, we recall basic facts on weakly differentiable curves, defined on the interval $(0,T)$, with values in some Banach space B. Although we are interested mainly in the case $B = L^r(\mathfrak{m})$, we first provide some general definitions and results, closely following [Showalter, 1997, §III.1], and then we consider particular cases of our interest.

Let B be a Banach space (not necessarily separable) with norm $\|\cdot\|$ and denote the duality pairing $B \times B^* \to \mathbb{R}$ by $(f, \varphi) \mapsto \langle f, \varphi \rangle := \varphi(f)$. We say that a curve $u = (u_t)_{t \in (0,T)} \subseteq B$ is measurable if it can be obtained as a $\mathscr{L}^1$-a.e. limit of a sequence of simple (i.e., with finite range) Borel curves: this notion becomes particularly useful when dealing with non-necessarily separable spaces. We write $u \in L^1_t(B)$ if u is measurable and $|u| = (\|u_t\|)_t \in L^1_t$, i.e., $L^1((0,T), \mathscr{L}^1)$. We consider Lebesgue spaces $L^p_t(B)$ of equivalence classes of functions, in a similar way as for $B = \mathbb{R}$: these are complete Banach spaces, and for $p \in [1,\infty]$, the continuous dual $(L^p_t(B))^*$ contains $L^{p'}_t(B^*)$, via the map

$$L^{p'}_t(B^*) \ni \varphi = (\varphi_t)_{t \in (0,T)} \mapsto \left[L^p_t(B) \ni f \mapsto \int_0^T \langle f_t, \varphi_t \rangle \, dt \right].$$

These spaces are separable B is separable and $p < \infty$ as simple Borel curves are dense (when $p = \infty$ and B is a dual space, simple functions provide only a weakly-* dense set).

Spaces of weakly differentiable curves with values in B are defined by means of distributional derivatives, and precisely in duality with the space of test functions $C^1_c((0,T); M)$, where we let $M \subseteq B^*$ be some closed subspace (this allows to deal with the case of B being a dual space, and letting M be the primal space embedded in the bidual). For the sake of simplicity, we often omit to write M in all what follows, and we consider it as fixed; moreover, we assume that M is sufficiently large, namely we require $\|f\| = \sup_{\varphi \in M} \langle f, \varphi \rangle / \|\varphi\|_{B^*}$, for every $f \in B$. We say that $\varphi : (0,T) \to M$ is differentiable at $t \in (0,T)$ with derivative $\varphi'_t \in M$ if, for every $f \in B$,

$$\lim_{\varepsilon \to 0} \left| \frac{\langle f, \varphi_{t+\varepsilon} \rangle - \langle f, \varphi_t \rangle}{\varepsilon} - \langle f, \varphi'_t \rangle \right| = 0.$$

Definition 5.2 (weak derivatives)**.** *Let $p \in [1,\infty]$, and let $u \in L^p_t(B)$. We say that $h \in L^p_t(B)$ is (the) weak derivative of u if it holds*

$$\int_0^T \langle u_t, \varphi'_t \rangle \, dt = - \int_0^T \langle h_t, \varphi_t \rangle \, dt, \quad \textit{for every } \varphi \in C^1_c((0,T); M), \tag{5.3}$$

and write $h = \partial_t u$. We denote by $W^{1,p}_t(B)$ the Banach space of functions $u \in L^p_t(B)$ for which $\partial_t u \in L^p_t(B)$ exists, endowed with the norm

$$\|u\|_{W^{1,p}_t(B)} := \|u\|_{L^p_t(B)} + \|\partial_t u\|_{L^p_t(B)} .$$

From (5.3), we deduce that if $(u_n)_n \subseteq W^{1,p}_t(B)$ is a sequence with $u_n \to u$, $\partial_t \to h$ weakly in $L^p_t(B)$, then $u \in W^{1,p}_t(B)$ with $\partial_t u = h$.

Remark 5.3. It is not difficult to show that $u \in W^{1,p}_t(B)$ still holds if the identity (5.3) is satisfied for φ belonging to a smaller set, e.g. as those functions φ of the form $\varphi_t = \psi_t \phi$, for some $\psi \in C^1_c((0,T); \mathbb{R})$ and $\phi \in \mathcal{D}$, with $\mathcal{D}$ weakly-* dense in $M \subseteq B^*$. Indeed, it is sufficient to approximate any $\varphi \in C^1_c((0,T); \mathbb{R})$ with a pointwise convergent, uniformly bounded sequence of linear combinations of such functions. In turn, this is possible writing $\varphi = \varphi_0 + \int_0^t \varphi'_s ds$, and approximating φ_0 and $\varphi' \in L^\infty_t(B^*)$, with respect to weak-* convergence, with bounded simple functions whose range belongs to $\mathcal{D}$. ■

Proposition 5.4 ($H = W$)**.** *Let $p \in [1,\infty)$ and $u \in W^{1,p}_t(B)$. There exists $(u^n)_{n \geq 1} \subseteq C^1_b((0,T); B)$ with $u^n \to u$ in $W^{1,p}_t(B)$. For $p = \infty$, one can find $(u^n)_n$ such that $u^n \to u$, $\partial_t u^n \to \partial_t u$ in duality with $L^1_t(B^*)$, as $n \to \infty$.*

Proof. Extend u to a curve on $(-T,T)$ by reflection, i.e. let $\overline{u}(t) = u(-t)$ for $t \in (-T,0)$. It can be checked that $\overline{u} \in W^{1,p}((-T,T); B)$, with $\partial_t \overline{u}_t = \partial_t u_{-t}$, a.e. $t \in (-T,0)$.Let $\rho \in C^1_c(0,T)$ be a convolution kernel, i.e. $\rho \geq 0$, and $\int \rho(s) ds = 1$ and for $n \geq 1$ define (via Bochner's integral) the convolution

$$u^n_t := \int_0^T \overline{u}(t - s/n) \rho(s) ds, \text{for } t \in (0,T).$$

For $n \geq 1$, it holds $u^n \in C^1_b((0,T); B)$, with

$$(u^n)'_t = \int_0^T (\partial_t \overline{u})(t - s/n) \rho(s) ds = \frac{1}{n} \int_0^T \overline{u}(t - s/n) \rho'(s) ds.$$

where continuity follows for strong continuity of translations in $L^p_t(B)$. Again by strong continuity of translations, as $n \to \infty$, one obtains the claimed convergence $u^n \to u$ in $W^{1,p}_t(B)$. $\square$

Definition 5.5 (absolutely continuous curves)**.** *Let $p \in [1,\infty]$. We say that a curve $u : (0,T) \mapsto B$ is absolutely continuous and it belongs to $AC^p((0,T);B)$ (or briefly $AC^p_t(B)$) if there exists $h \in L^p(0,T)$ such that*

$$\|u_t - u_s\| \le \int_s^t h_r dr, \quad \textit{for every } s,\, t \in (0,T),\ s \le t. \tag{5.4}$$

If $u \in AC^p_t(B)$, then the $\mathscr{L}^1$-essentially smallest among all h's such that (5.4) holds is called the metric speed of u and denoted by $(\|\dot{u}_t\|)_{t\in(0,T)}$. Notice that, if $u \in AC^p_t(B)$, then it is uniformly continuous so the limits $u_0 := \lim_{t\downarrow 0} u_t$ and $u_0 := \lim_{t\uparrow T} u_t$ exist.

In the case $B = \mathbb{R}$, it is known that $AC^p(0,T) = W^{1,p}(0,T)$, so in particular every Sobolev function admits a continuous representative belonging to $AC^p(0,T)$. We provide a weak generalization of this (only in one direction) to the case of a general space B, see e.g. [Showalter, 1997, Proposition III.1.1].

Proposition 5.6. *Let $u \in W^{1,p}_t(B)$. Then, there exists a (unique) representative $\tilde{u} \in AC^p_t(B)$, with $\|\dot{\tilde{u}}\| \le \|\partial_t u\|$, $\mathcal{L}^1$-a.e. in $(0,T)$, and*

$$\|\tilde{u}\|_{C([0,T];B)} := \sup_{t\in(0,T)} \|\tilde{u}\| \le C \|u\|_{W^{1,p}_t(B)}, \tag{5.5}$$

where C is some constant depending on p and T only.

It actually holds $\|\dot{\tilde{u}}\| = \|\partial_t u\|$, $\mathscr{L}^1$-a.e. in $(0,T)$, but we do not need this fact. Thanks to this proposition, we always identify $u = \tilde{u}$ whenever $u \in W^{1,p}_t(B)$.

Proof. It is sufficient to establish (5.5) for $p = 1$ and $u \in C^1_b((0,T);B)$ arguing then by density, using Proposition 5.4. Let s, $t \in (0,T)$, with $s \le t$ and $\varphi \in M$. Then, it holds

$$\langle u_t, \varphi\rangle - \langle u_s, \varphi\rangle = \int_s^t \langle u'_r, \phi\rangle\, dr,$$

so that we have

$$\|u_t - u_s\| = \sup_{\varphi\in M, \|\varphi\|_{B^*}\le 1} |\langle u_t, \varphi\rangle - \langle u_s, \varphi\rangle| \le \sup_{\varphi\in M, \|\varphi\|_{B^*}\le 1} \int_s^t |\langle u'_r, \varphi\rangle|\, dr, \le \int_s^t \|u'_r\|\, dr$$

and by averaging with respect to s, we deduce, for $t \in (0,T)$,

$$T\|u_t\| \le \int_0^T \|u_t - u_s\|\, ds + \int_0^T \|u_s\|\, ds \le T\int_0^T \|u'_s\|\, ds + \int_0^T \|u_s\|\, ds,$$

which entails (5.5). $\square$

We conclude this section by providing some variants of Proposition 5.6, first in a Gelfand triple setting [Showalter, 1997, Proposition III.1.2] and then in the case of B being a Lebesgue space.

Recall that a Gelfand triple is defined by fixing some separable Hilbert space H and a dense (continuous) inclusion $i : B \hookrightarrow H$. Then, by identifying $H \sim H^*$ via the Riesz map, one obtains

$$B \overset{i}{\hookrightarrow} H \sim H^* \overset{i^*}{\hookrightarrow} B^*,$$

where $i^* \circ i : B \hookrightarrow B^*$ is a dense inclusion (i^* is injective because the range of i is dense). As a consequence, the spaces $L^p_t(B)$ are embedded into $L^p_t(B^*)$. Let $b \in B$, $h \in H$. Then, by definition of adjoint, it holds $\langle i(b), h\rangle_H = \langle b, i^*(h)\rangle_B$, thus $|\langle h, i(b)\rangle_H| \leq \|b\|_B \|i^*(h)\|_{B^*}$ (we naturally introduce a specific notation for the norms and duality maps in the various spaces). In this case, we naturally let $M := B$.

Proposition 5.7. *In the Gelfand triple setting introduced above, let $p \in [1, \infty]$ and $u \in W^{1,p}_t(B^*) \cap L^{p'}_t(B)$. Then, $u \in C([0,T]; B^*)$ actually belongs to $C([0,T]; H)$ and $(\|u_t\|^2_H)_{t\in(0,T)}$ is absolutely continuous, with*

$$\partial_t \|u_t\|^2_H = 2\langle u_t, \partial_t u_t\rangle_B, \quad a.e.\ t \in (0,T).$$

Proof. The proof is similar to that of the previous proposition. The proof of Proposition 5.4 clearly shows that it is sufficient to argue for $u \in C^1_b([0,T]; B)$ and provide a quantitative bound that entails the general case by density. For s, $t \in (0,T)$, with $s \leq t$ and $\varphi \in B \subseteq H$, it holds

$$\langle u_t, \varphi\rangle_H - \langle u_s, \varphi\rangle_H = \int_s^t \langle u'_r, \varphi\rangle_H \, dr,$$

thus, letting $\varphi = u_t - u_s \in B \subseteq H$, we obtain

$$|u_t - u_s|^2_H = \int_s^t \langle u_t - u_s, u'_r\rangle_H \, dr \leq \int_s^t \|u_t - u_s\|_B \|u'_r\|_{B^*} \, dr.$$

By averaging over $s \in (0,T)$, we obtain

$$\begin{aligned} \frac{T}{2}\|u_t\|^2_H &\leq \int_0^T \|u_t - u_s\|^2_H \, ds + \int_0^T \|u_t\|^2_H \, ds \\ &\leq T\int_0^T \|u_s\|_B \|u'_s\|_{B^*} \, ds + \int_0^T \|u_s\|^2_B \, ds, \\ &\leq T\|u\|_{L^{p'}_t(B)} \|\partial_t u\|_{L^p_t(B^*)} + T\|u\|^2_{L^\infty((0,T);B^*)}, \end{aligned}$$

where the last inequality is a consequence of Hölder inequality. By the previous proposition, the last term is estimated from above by $\|u\|_{W^{1,1}_t(B^*)}$, thus providing the required bound. The second statement follows again by density, arguing first for $u \in C^1_b((0,T); B)$, thus $(\|u_t\|^2_H)_{t\in(0,T)} \in C^1_b(0,T)$, with

$$\frac{d}{dt}\|u_t\|^2_H = 2\langle u_t, u'_t\rangle_H, \quad \text{for } t \in (0,T).$$

$\square$

We deduce more precise results in the case of spaces of functions in our framework, i.e. we let $H = L^2(\mathfrak{m})$ and $B = \mathbb{V}$.

Proposition 5.8. *Let $p \in [1,\infty]$, $u \in L^p_t(\mathbb{V}^*) \cap L^{p'}_t(\mathbb{V})$, $\beta \in C^2_b(\mathbb{R})$, with $\beta(0) = \beta'(0) = 0$.*
Then, the curve $(0,T) \ni t \mapsto \int \beta(u_t) d\mathfrak{m}$ is $AC^1(0,T)$, with

$$\partial_t \int \beta(u_t) d\mathfrak{m} = \left\langle \beta'(u_t), \partial_t u_t \right\rangle_{\mathbb{V}}, \quad \mathscr{L}^1\text{- a.e. } t \in (0,T).$$

Recall that we tacitly identify u with its continuous representative, which by the previous result is continuous with values in $H = L^2(\mathfrak{m})$, so that the curve above is well-defined.

Proof. Again, we argue by density, letting $u \in C^1_b((0,T);\mathbb{V})$. Notice that the map $t \mapsto \int \beta(u_t) d\mathfrak{m}$ is well defined and continuous, by the assumptions on β. By a Taylor expansion, it is not difficult to prove that it is $C^1_b(0,T)$, with

$$\frac{d}{dt} \int \beta(u_t) d\mathfrak{m} = \int \beta'(u_t) u_t' d\mathfrak{m}, \text{for } t \in (0,T). \tag{5.6}$$

Let $u^n \in C^1_b((0,T);\mathbb{V})$ converge towards u in $L^p_t(\mathbb{V}^*) \cap L^{p'}_t(\mathbb{V})$. Then, the regularity and growth assumptions on β give that for every $t \in (0,T)$, $\beta(u^n_t)$ converge towards $\beta(u_t)$, which also provides a continuous curve. To show its absolute continuity, $AC^1(0,T)$, it is sufficient to prove that it belongs to $W^{1,1}(0,T)$, passing to the limit in (5.6) with u^n in place of u, with respect to convergence in $L^1(0,T)$. This is a consequence of $\beta'(u^n) \to \beta'(u)$ in $L^p_t(\mathbb{V})$, by the chain rule, the uniform bound on β'' and Proposition 3.1. □

We conclude this section with another generalization of Proposition 5.6, where we deal with Lebesgue spaces. Its easy proof is just a variant of the arguments provided so far.

Proposition 5.9. *Let p, q, r, $s \in [1,\infty]$, satisfy $p^{-1}+r^{-1}+s^{-1} = 1$ and let $u \in W^{1,p}_t(L^{s'}(\mathfrak{m})) \cap L^p_t(L^r(\mathfrak{m}))$.*
Then, for every $\beta \in C^1_b(\mathbb{R})$ with $\beta(0) = \beta'(0) = 0$, the curve $(\beta(u_t))_{t\in(0,T)}$ is $AC^1_t(L^{s'}(\mathfrak{m}))$, with

$$\partial_t \beta(u_t) = \beta'(u_t)\partial_t u_t, \quad \mathscr{L}^1\text{-a.e. } t \in (0,T).$$

Remark 5.10. It is possible to state and prove variants where β' has controlled growth. In particular, for $|\beta'(z)| \le |z|^{s/r}$ we obtain that the curve $\beta(u)$ belongs to $AC^1_t(L^1(\mathfrak{m}))$. Further variants can be devised, e.g. dealing with sums and intersections of Lebesgue spaces.

Chapter 6

Fokker-Planck equations, martingale problems and their equivalence

In this chapter, we introduce Fokker-Planck equations, martingale problems and flows, for diffusions in metric measure spaces, analogously to Chapter 1: although the arguments are perfectly parallel to those developed for diffusion processes in $\mathbb{R}^d$, the theory is different in a crucial aspect, besides its framework. Indeed, there is no hope for pointwise investigations, and we always look for results averaged with respect to $\mathfrak{m}$.

Throughout all this section, we let (X, τ), $\mathfrak{m}$ and $\mathcal{E}$ satisfy (3.1), fix $T > 0$ and consider the time-extended framework as introduced in the previous chapter, letting in particular $\tilde{\mathscr{A}}$ satisfy (5.2). Finally, we let $r \in (1, \infty]$ and $\mathcal{L}$ be a diffusion operator on $(0, T) \times X$, such that

$$\mathcal{L}f = (\mathcal{L}_t f)_{t \in (0,T)} \in L^1_t(L^{r'}_x), \quad \text{for every } f \in \tilde{\mathscr{A}}. \tag{6.1}$$

For simplicity, as there is no danger of confusion, we also write $\mathscr{A}$ in place of $\tilde{\mathscr{A}}$.

6.1 Definition and basic facts

In this section, which parallels Section 1.1, we introduce suitable notions of solutions to Fokker-Planck equations, martingale problems and flows associated to $\mathcal{L}$, together with basic results.

Fokker-Planck equations

Definition 6.1 (solutions in $L^\infty_t(L^r_x)$ to FPE's)**.** *A Borel function u on $(0, T) \times X$ is said to be a solution in $L^\infty_t(L^r_x)$ to the Fokker-Planck equation (FPE)*

$$\partial_t u_t = \mathcal{L}^*_t(u_t), \quad \text{in } (0, T) \times X, \tag{6.2}$$

if it holds $u = (u_t)_t \in L^\infty_t(L^r_x)$ and

$$\int_0^T \int [\partial_t f_t + \mathcal{L}_t f] \, u_t d\mathfrak{m} dt = 0, \quad \text{for every } f \in \mathscr{A}_c. \tag{6.3}$$

Let us explicitly remark that we prefer to consider a solution u as a function and not an equivalence class with respect to $\tilde{\mathfrak{m}}$-negligible sets in $(0,T)\times X$: to keep a perfect correspondence with Section 1.1, the natural object to consider should be the Borel curve of measures $\nu := (u_t\mathfrak{m})_{t\in(0,T)}$, but we prefer to work directly with densities.

As usual with weak formulations for PDE's, the main advantage is that no regularity assumption on u, but only integrability, is imposed.

Remark 6.2 (extension of the weak formulation)**.** Assume that $\mathcal{L}$ is extended on a space $F \subseteq W^{1,1}_t(L^{r'}(\mathfrak{m}))$, containing $\mathscr{A}_c$, and such that any $f \in F$ can be approximated by sequence $(f_n)_n \subseteq \mathscr{A}_c$ such that

$$f_n \to f, \quad \partial_t f_n \to \partial_t f, \quad \mathcal{L} f_n \to \mathcal{L} f, \quad \text{weakly in } L^1_t(L^{r'}_x). \tag{6.4}$$

Then, any solution u in $L^\infty_t(L^r_x)$ to the FPE (6.2) clearly satisfies (6.3) also with $f \in F$. ■

Remark 6.3 (equivalent formulation)**.** By (5.2), $\mathscr{A}$ is stable with respect to multiplication by elements in $C^1_c(0,T)$, letting fg in place of $f \in \mathscr{A}$ in (6.3), with $g \in C^1_c(0,T)$, we obtain that any solution u in $L^\infty_t(L^r_x)$ to the FPE (6.2) satisfies

$$\int_0^T \partial_t g_t \left[\int f_t u_t d\mathfrak{m}\right] dt = \int_0^T g_t \int \left[\partial_t f_t + \mathcal{L}_t f\right] u_t d\mathfrak{m} dt$$

thus the curve $(0,T) \ni t \mapsto \int f_t u_t d\mathfrak{m}$ is $W^{1,1}(0,T)$, with

$$\partial_t \int f_t u_t d\mathfrak{m} = \int \left[\partial_t f_t + \mathcal{L}_t f\right] u_t d\mathfrak{m}, \quad \mathscr{L}^1\text{-a.e. } t \in (0,T). \tag{6.5}$$

Clearly, by integration over $t \in (0,T)$, requiring (6.5) to hold for every $f \in \mathscr{A}_c$ is equivalent to the original formulation (6.3). ■

We next address the validity of an analogue of Remark 1.4, i.e., of [Ambrosio et al., 2008, Lemma 8.1.2] in this setting. The assumption $r > 1$ plays a role here; the main difficulty is to formulate the correct "separability" condition on $\mathscr{A}$, which we add as a hypothesis.

Lemma 6.4 (existence of a continuous representative)**.** *Let u be a solution in $L^\infty_t(L^r_x)$ to the FPE* (6.2)*. Assume that there exists a countable set $\mathscr{A}^* \subseteq \mathscr{A}$ such that every $f \in \mathscr{A}$ is approximated by a sequence $(f_n)_n \subseteq \mathscr{A}^*$ in the sense of* (6.4)*.*

Then, there exists a (unique) weakly- continuous curve $\tilde{u} = (\tilde{u}_t)_{t\in[0,T]} \subseteq L^r(\mathfrak{m})$, with $u_t = \tilde{u}_t$ $\mathfrak{m}$-a.e. in X, $\mathscr{L}^1$-a.e. $t \in (0,T)$, such that*

$$\int f_{t_2}\tilde{u}_{t_2} d\mathfrak{m} - \int f_{t_1}\tilde{u}_{t_1} d\mathfrak{m} = \int_{t_1}^{t_2}\int \left[\partial_t f_s + \mathcal{L}_s f\right] u_s d\mathfrak{m} ds, \quad \textit{for } t_1,\, t_2 \in [0,T], \textit{ with } t_1 \le t_2. \tag{6.6}$$

Proof. For every $f \in \mathscr{A}^*$, let $A_f \subseteq (0,T)$ be the set of Lebesgue points for the Borel curve $t \mapsto \int f_t u_t d\mathfrak{m}$ and let $A := \cap_{f\in\mathscr{A}^*} A_f$. Notice that the function u, restricted to $A \times X$ provides a bounded Borel curve $(u_t)_{t\in A} \subseteq L^r(\mathfrak{m})$.

Moreover, such a curve $(u_t)_{t\in A}$ can be also be seen as a family of linear functionals on $\mathscr{A}$. Indeed, it holds, for every t_1, $t_2 \in A$ with $t_1 \le t_2$ and $f \in \mathscr{A}^*$,

$$\int f_{t_2} u_{t_2} d\mathfrak{m} = \int f_{t_1} u_{t_1} d\mathfrak{m} + \int_{t_1}^{t_2}\int \left[\partial_t f_s + \mathcal{L}_s f\right] u_s d\mathfrak{m} ds,$$

and the right hand side is continuous with respect to the convergence in (6.4), thus we may define uniquely a linear functional $\ell_{t_2} : \mathscr{A} \to \mathbb{R}$. Notice the the precise choice for $t_1 \in A$ is not relevant, thus the condition $t_1 \le t_2$ can be dropped. The right hand side can be used to define similarly ℓ_{t_2} for $t_2 \in (0, T]$ (again, choosing any $t_1 \le t_2$ with $t_1 \in A$) and even ℓ_0, by letting

$$\ell_0(f) = \int f_{t_2} u_{t_2} d\mathfrak{m} - \int_0^{t_2} \int \left[\partial_t f_s + \mathcal{L}_s f\right] u_s d\mathfrak{m} ds,$$

for any $t_2 \in A$. Notice that, for every $f \in \mathscr{A}$, $[0, T] \ni t \mapsto \ell_t(f)$ is a continuous curve. The left hand side above provides the bound

$$|\ell_t(f)| \le \sup_{s \in A} \|u_s\|_r \|f_t\|_{r'}, \quad \text{for } t \in A,$$

whose validity can be extended, by taking limits, for $t \in [0, T]$. By density of the range of $f \mapsto f_t \in L^{r'}(\mathfrak{m})$ in (5.2), we see that ℓ_t induces a linear continuous functional on $L^{r'}(\mathfrak{m})$, with norm smaller than $\sup_{s \in A} \|u_s\|_r$ and so, by duality between Lebesgue spaces (where we use $r > 1$), we conclude that for every $t \in [0, T]$, there exists $\ell_t(f) = \int f_t \tilde{u}_t d\mathfrak{m}$ and moreover, $t \mapsto \tilde{u}_t \in L^r(\mathfrak{m})$ is weakly-* continuous and satisfies (6.6). □

From a classical PDE point of view, Definition 6.1 is very weak, since e.g. in the case of $\mathcal{L} = \Delta$, we integrate by parts twice. The usual definition of weak solution, i.e. integrating by parts once, can be recovered whenever $\mathcal{L}$ is in divergence form and $u \in L^2_t(\mathbb{V})$, as the next Proposition shows: its proof is straightforward from the definition of diffusion in divergence form.

Proposition 6.5 (solutions to FPE's in divergence form)**.** *Let $r \in [2, \infty]$ and $\mathcal{L}$ be in divergence form, with*

$$|\boldsymbol{a}|, |\boldsymbol{b}|, \in L^2_t(L^q_x), \qquad \text{with } 1/q = 1/2 + 1/r.$$

Then, $u \in L^\infty_t(L^r_x) \cap L^2_t(\mathbb{V})$ is a solution to the FPE (6.2) if and only if it holds

$$\int \left[(\partial_t f - df(\boldsymbol{b}))\, u - \boldsymbol{a}(f, u)\right] d\tilde{\mathfrak{m}} = 0, \quad \text{for every } f \in \mathscr{A}_c.$$

In the next section, we prove that a satisfactory correspondence between well-posedness of FPE's and martingale problems can be established, dealing with solutions in $L^\infty_t(L^r_x)$, not only in $L^2_t(\mathbb{V})$.

Martingale problems

Before we define solutions to the martingale problem associated to a diffusion operator $\mathcal{L}$, let us introduce the following notation, which parallels that in Section 1.1: on the Polish space $C([0, T]; X)$, we define $e_t : \gamma \mapsto \gamma_t := \gamma(t) \in X$ be the evaluation map, for $t \in [0, T]$. The natural filtration on $C([0, T]; X)$ is the increasing family of σ-algebras $\mathcal{F} = (\mathcal{F}_t)_{t \in [0,T]}$ with $\mathcal{F}_t := \sigma(e_s : s \in [0, t])$. Given $\boldsymbol{\eta} \in \mathscr{P}(C([0, T]; X))$, we let $\eta_t := (e_t)_\sharp \boldsymbol{\eta}$ be the 1-marginal law at t. Notice that the family $\eta := (\eta_t)_{t \in [0,T]}$ is Borel and actually narrowly continuous. We also write $\eta \in L^\infty_t(L^r_x)$ when $\eta_t \ll \mathfrak{m}$, with density belonging to $L^r(\mathfrak{m})$, with a uniform bound in $t \in [0, T]$. With a slight abuse of notation we then write $\|\eta_t\|_r$ for the $L^r(\mathfrak{m})$-norm of $d\eta_t/d\mathfrak{m}$ and $\|\eta\|_{L^\infty_t(L^r_x)}$ for $\sup_{t \in [0,T]} \|\eta_t\|_r$.

Definition 6.6 (L^r-regular solutions to MP's). *A probability measure $\boldsymbol{\eta} \in \mathscr{P}(C([0,T];X))$ is a L^r-regular solution to the martingale problem (MP) associated to $\mathcal{L}$ if $\eta \in L^\infty_t(L^r_x)$ and, for every $f \in \mathscr{A}$, the process*

$$[0,T] \ni t \mapsto f_t \circ e_t - \int_0^t [\partial_t f_s + \mathcal{L}_s f] \circ e_s ds \tag{6.7}$$

is a martingale with respect to the natural filtration on $C([0,T];X)$.

The assumption $\eta \in L^\infty_t(L^r_x)$ entails, for $t \in [0,T]$, the bound

$$\int |f_t| \circ e_t d\boldsymbol{\eta} + \int \int_0^t [|\partial_s f_s| \circ e_s + |\mathcal{L}_s f| \circ e_s] \, d\boldsymbol{\eta} \le C \, \||\partial_t f| + |\mathcal{L} f|\|_{L^1_t(L^{r'}_x)} \|\eta\|_{L^\infty_t(L^r_x)}, \tag{6.8}$$

where C is some constant depending on T and r only (due to the norm of the trace operator $f \mapsto f_t$). The process defined in (6.7) is therefore uniformly bounded in $L^1(\boldsymbol{\eta})$, so that one is always reduced to check orthogonality of increments, and in particular that for every $t \in [0,T]$ and bounded $\mathcal{F}_t$-measurable $g : C([0,T];X) \to \mathbb{R}$ it holds

$$\int g \left[f_T \circ e_T - \int_0^T [\partial_t f_s + \mathcal{L}_s f] \circ e_s ds \right] d\boldsymbol{\eta} = \int g \left[f_t \circ e_t - \int_0^t [\partial_t f_s + \mathcal{L}_s f] \circ e_s ds \right] d\boldsymbol{\eta}.$$

We also notice that the process $[0,T] \ni t \mapsto \int_0^t [\partial_t f_s + \mathcal{L}_s f] \circ e_s ds$ can be defined as a progressively measurable, a.s. continuous process in $L^\infty_{loc}(\boldsymbol{\eta}, (\mathcal{F}_t)_t)$, see right after Definition 1.5.

Remark 6.7 (the deterministic case). When $\mathcal{L} = \boldsymbol{b}$ is a derivation, the martingales

$$t \mapsto f_t \circ e_t - \int_0^t [\partial_t f_s + df_s(\boldsymbol{b}_s)] \circ e_s ds, \quad \text{for } f \in \mathscr{A}$$

reduce to constant processes, since their quadratic variation, which in general is given by $2\boldsymbol{a}(f)$ is identically zero, see also Section 2.2.2. Therefore, martingale solutions are probability measures $\boldsymbol{\eta}$ concentrated on curves for which $t \mapsto f_t \circ e_t$ is Sobolev with $\partial_t f_t \circ e_t = [\partial_t f_t + df_t(\boldsymbol{b}_t)] \circ e_t$, $\boldsymbol{\eta}$-a.e., which recovers the notion of solution to the ODE induced by $\boldsymbol{b}$ given in [Ambrosio and Trevisan, 2014, Definition 7.3]. ■

Remark 6.8 (solutions to MP's induce solutions to FPE's). Integrating with respect to $\boldsymbol{\eta}$, i.e., taking expectation, any L^r-regular solution $\boldsymbol{\eta}$ to the MP provides, by means of its 1-marginals, a weakly-* continuous solution $\eta = (\eta_t)_t \subseteq L^r(\mathfrak{m})$ to the FPE (6.2). ■

As in Section 1.1, we investigate stability properties for solutions to martingale problems, remarking that all the definitions above can be given also with respect to any interval $[t_1, T]$, in place of $[0,T]$.

Proposition 6.9. *Let $t_1 \in [0,T)$, $\boldsymbol{\eta} \in \mathscr{P}(C([t_1,T];X)$ be a L^r-regular solution to the martingale problem associated to $\mathcal{L}$. Let $t_2 \in [t_1,T]$ and let $\rho : C([t_1,T];X) \to [0,\infty)$ be a probability density (with respect to $\boldsymbol{\eta}$) belonging to $L^\infty(\boldsymbol{\eta})$, measurable with respect to $\mathcal{F}_{t_2}$. Let π denote the natural restriction map*

$$C([t_1,T];X) \ni \gamma \mapsto (\gamma(t))_{t \in [t_2,T]} \in C([t_2,T];X).$$

Then, $\pi_\sharp(\rho\boldsymbol{\eta}) \in \mathscr{P}(C([t_2,T];X)$ is a L^r-regular solution to the martingale problem associated to $\mathcal{L}$, on the space $C([t_2,T];X)$.

Proof. The proof goes exactly as that of Proposition 1.8, noticing also that the inequality

$$(\pi_\sharp(\rho\boldsymbol{\eta})) \le \|\rho\|_{L^\infty(\boldsymbol{\eta})}\,\pi_\sharp(\boldsymbol{\eta})$$

entails the bound in $L^r(\mathfrak{m})$ for the 1-marginals of $\pi_\sharp(\rho\boldsymbol{\eta})$, uniformly in $t \in [t_2, T]$. $\square$

Stability with respect to convex combinations can be proved exactly as in Proposition 1.9; the proof is even simpler, as the integrability assumptions are automatically fulfilled because of (6.8). Finally, L^r-regularity follows by Jensen's inequality.

Proposition 6.10. *Let $(Z, \mathcal{A})$ be a measurable space, $\overline{\nu} \in \mathscr{P}(Z)$ and $(\boldsymbol{\eta}_z)_{z\in Z} \subseteq \mathscr{P}(C[0,T];X)$ be a Borel family of probability measures, such that $\boldsymbol{\eta}_z$ is a L^r-regular solution to the MP associated to $\mathcal{L}$, for $\overline{\nu}$-a.e. $z \in Z$. Then, $\boldsymbol{\eta} := \int \boldsymbol{\eta}_z d\overline{\nu}(z) \in \mathscr{P}(C[0,T];X)$ is a L^r-regular solution to the MP associated to $\mathcal{L}$.*

Martingale flows

The notion of martingale flows consists in a Borel selection of solutions to martingale problems, so that L^r-regularity is preserved: this idea originates in DiPerna-Lions theory for ODE's with weakly differentiable coefficients and was subsequently put in a more convenient formulation by [Ambrosio, 2004].

Definition 6.11 (L^r-regular martingale flows)**.** *A family $(\boldsymbol{\eta}(s,x))_{(s,x)\in[0,T]\times X}$ of probability measures on $C([0,T];X)$ is said to be a L^r-regular martingale flow (MF) associated to $\mathcal{L}$ if, for every $s \in [0,T]$, $x \mapsto \boldsymbol{\eta}(s,x)$ is Borel and for every $\overline{u} \in L^r(\mathfrak{m})$, with $\overline{u}\mathfrak{m}$ probability,*

$$\boldsymbol{\eta} := \int \boldsymbol{\eta}(s,x)\overline{u}(x)d\mathfrak{m}(x) \in \mathscr{P}(C([0,T];X)) \tag{6.9}$$

is a L^r-regular solution to the martingale problem on $[0,T]$, associated to the diffusion $\chi_{[s,T]}\mathcal{L}$, with $\eta_s = \overline{u}\mathfrak{m}$.

Remark 6.12. Differently from Chapter 1, we are not allowed to let $\overline{u}\mathfrak{m} = \delta_x$ for $x \in X$ and deduce that $\boldsymbol{\eta}(s,x)$ is a solution the martingale problem in $[s,T]$ with law at s given by δ_x. Still, from the identity

$$\overline{u}\mathfrak{m} = \int \boldsymbol{\eta}(s,x)_s \overline{u}(x) d\mathfrak{m}(x), \quad \overline{u} \in L^r(\mathfrak{m}), \text{ with } \overline{u}\mathfrak{m} \text{ probability},$$

we obtain $\eta(s,x)_s = \delta_x$, $\mathfrak{m}$-a.e. $x \in X$, for every $s \in [0,T]$. ■

Remark 1.11 as well as Remark 1.12 can be rephrased also in this setting. Notice that Chapman-Kolmogorov equations, which read as

$$\eta(s,x)_t = \int_X \eta(r,y)_t \eta(s,x)_r(dy), \quad \mathfrak{m}\text{-a.e. } x \in X, \text{ for every } r,\ s,\ t \in [0,T] \text{ with } s \le r \le t, \tag{6.10}$$

are not a condition for defining a L^r-regular martingale flow, as we obtain them as a consequence of well-posedness.

The superposition principle

Definition 6.13 (superposition principle)**.** *Let $u = (u_t)_{t\in(0,T)}$ be a solution in $L^\infty_t(L^r_x)$ to the FPE (6.2). We say that the superposition principle holds for u if there exists a superposition solution $\boldsymbol{\eta} \in \mathscr{P}(C([0,T];X))$, i.e., a L^r-regular solution to the MP associated to $\mathcal{L}$ such that $\eta_t = u_t\mathfrak{m}$, $\mathscr{L}^1$-a.e. $t \in (0,T)$.*

Clearly, existence of a superposition solution for u entails existence of a weakly-* continuous representative and that $u_t\mathfrak{m}$ is a probability measure, $\mathscr{L}^1$-a.e. $t \in [0,T]$. Moreover, if $u_t\mathfrak{m}$ is already weakly-* continuous, one has $u_t\mathfrak{m} = \eta_t$ for every $t \in [0,T]$. We address the validity of the superposition principle in metric measure spaces, under suitable assumptions, in the next chapter.

6.2 Correspondence between FPE's, MP's and flows

Arguing similarly as in Section 1.2, we show that the validity of the superposition principle entails abstract correspondences between well-posedness for the notions introduced above.

Fokker-Planck equations $\Leftrightarrow$ martingale problems

If the superposition principle holds, then existence results are transferred easily both ways: in particular existence of solutions ν in $L^\infty_t(L^r_x)$ to the FPE associated to $\mathcal{L}$, to which the superposition principle applies, entails existence of solutions to the MP. With an identical proof, the following analogue of Lemma 1.14 holds.

Lemma 6.14 (transfer of uniqueness for 1-marginals)**.** *Let $\overline{u} \in L^r(\mathfrak{m})$, with $\overline{u}\mathfrak{m}$ probability. Assume that the superposition principle holds for every weak-* continuous solution in $L^\infty_t(L^r_x)$ to the FPE*

$$\partial_t u_t = \mathcal{L}^*_t u_t, \quad \text{in } (0,T)\times X, \text{ with } u_0 = \overline{u}. \tag{6.11}$$

Then, the following conditions are equivalent:

i) there exists at most one weak- continuous solution in $L^\infty_t(L^r_x)$ to the FPE (6.11)*

ii) any two L^r-regular solutions $\boldsymbol{\eta}^1$, $\boldsymbol{\eta}^2$ to the MP associated to $\mathcal{L}$, with $\eta^1_0 = \eta^2_0 = \overline{u}\mathfrak{m}$, have identical 1-marginals, i.e. $\eta^1_t = \eta^2_t$ for $t \in [0,T]$.

Next, we provide an analogue of Lemma 1.15. Let us point out that the proof goes identically and relies on Proposition 6.9, which in turn exploits

$$0 \le u \le v \in L^r(\mathfrak{m}) \quad \Rightarrow \quad u \in L^r(\mathfrak{m}). \tag{6.12}$$

Indeed, as explicitly remarked in [Ambrosio and Crippa, 2008, §3] and [Figalli, 2008, §3.1], besides minor technicalities, one could restate all these results, at least in the Euclidean framework, for solutions to FPE's in $L^\infty_t(L)$ and L-regular martingale problems, where L is any class of measures such that the analogue (6.12) holds, i.e.

$$0 \le u \le v \in L \quad \Rightarrow \quad u \in L.$$

One may also consider mixed integrability conditions, both with respect to t and x.

Lemma 6.15 (transfer of uniqueness). *For every $s \in [0,T]$, let the superposition principle hold for every weakly-* continuous solution in $L^\infty_t(L^r_x)$ to the FPE*

$$\partial_t u_t = \mathcal{L}^*_t u_t, \quad \text{in } (s,T) \times X.$$

Then, the following conditions are equivalent:

i) for every $s \in [0,T]$ and probability density $\overline{u} \in L^r(\mathfrak{m})$, there exists at most one weakly- continuous solution u in $L^\infty_t(L^r_x)$ to the FPE*

$$\partial_t u_t = \mathcal{L}^*_t u_t, \quad \text{in } (s,T) \times \mathbb{R}^d, \text{ with } u_s = u;$$

ii) for every $s \in [0,T]$, if $\boldsymbol{\eta}^1$, $\boldsymbol{\eta}^2 \in \mathscr{P}(C([s,T];X))$ are L^r-regular solutions to the MP associated to $\mathcal{L}$ on $C([s,T];X)$, with $\eta^1_s = \eta^2_s$, then $\boldsymbol{\eta}^1 = \boldsymbol{\eta}^2$.

Martingale problems $\Leftrightarrow$ martingale flows

We provide the analogue of Lemma 1.16. We introduced the notation $C_{s,\overline{u}} \subseteq \mathscr{P}(C([0,T];X))$ for the set of L^r-regular solutions to the MP associated to $\chi_{[0,s]}\mathcal{L}$, with $\eta_s = \overline{u}\mathfrak{m}$. Notice that we are not able to prove the implication *ii) $\Rightarrow$ i)*, except for the case $r = \infty$.

Lemma 6.16 (well-posedness of MF's). *Consider the following two conditions:*

i) for every $s \in [0,T]$, $\overline{u} \in L^r(\mathfrak{m})$, with $\overline{u}\mathfrak{m}$ probability, there exists a unique $u \in C_{s,\overline{u}}$;

ii) there exists a $\mathfrak{m}$-essentially unique L^r-regular martingale flow associated to $\mathcal{L}$, i.e. given any two flows $\boldsymbol{\eta}^1$, $\boldsymbol{\eta}^2$, there holds

$$\boldsymbol{\eta}^1(s,x) = \boldsymbol{\eta}^2(s,x), \quad \mathfrak{m}\text{-a.e. in } X, \text{ for every } s \in [0,T].$$

Then, it always holds i) $\Rightarrow$ ii)*, while* ii) $\Rightarrow$ i) *holds if $r = \infty$. Moreover, if* i) *holds, then the L^r-regular martingale flow satisfies* (1.7).

Proof. i) $\Rightarrow$ ii). Uniqueness of a L^r-regular martingale flow assuming uniqueness of L^r-regular martingale problems is trivial. To prove existence, we argue at fixed $s \in [0,T]$. For any $\overline{u} \in L^r(\mathfrak{m})$ with $\overline{u}\mathfrak{m}$ probability density, we consider (the unique) $\boldsymbol{\eta}^{\overline{u}} \in C_{s,\overline{u}}$ and disintegrate it with respect to e_s, i.e., we consider a regular conditional probability for the identity map in $C([s,T];X)$, given e_s. We thus obtain a Borel family of probability measures $\boldsymbol{\eta}^{\overline{u}}(s,x)$. We claim that, for any $\overline{v} \in L^r(\mathfrak{m})$ with $\overline{v}\mathfrak{m}$ probability, it holds

$$\boldsymbol{\eta}^{\overline{u}}(s,x) = \boldsymbol{\eta}^{\overline{v}}(s,x), \quad \mathfrak{m}\text{-a.e. } x \in X \text{ such that } \overline{u}(x) > 0 \text{ and } \overline{v}(x) > 0.$$

To this aim, it is sufficient to fix $\varepsilon > 0$, let $\rho \in L^1 \cap L^\infty(\mathfrak{m})$ be any probability density concentrated on $\{\overline{u} > \varepsilon, \overline{v} > \varepsilon\}$, and show that

$$\int \boldsymbol{\eta}^{\overline{u}}(s,x)\rho(x)d\mathfrak{m}(x) = \int \boldsymbol{\eta}^{\overline{v}}(s,x)\rho(x)d\mathfrak{m}(x) \quad \text{as measures on } C([0,T];X),$$

and this follows from uniqueness, as we show that both members belong to $C_{s,\rho}$. By definition of disintegration of measure, one can rewrite both sides above as

$$\int \boldsymbol{\eta}^{\overline{u}}(s,x)\rho(x)d\mathfrak{m}(x) = (\rho \circ e_s)\boldsymbol{\eta}^{\overline{u}} \quad \text{and} \quad \int \boldsymbol{\eta}^{\overline{v}}(s,x)\rho(x)d\mathfrak{m}(x) = (\rho \circ e_s)\boldsymbol{\eta}^{\overline{v}}.$$

Since it holds
$$\rho \leq \|\rho\|_\infty \chi_{\{\overline{u}>\varepsilon\}} < \varepsilon^{-1} \|\rho\|_\infty \overline{u}, \quad \mathfrak{m}\text{-a.e. in } X,$$
we obtain $\rho \circ e_s \leq \|\rho\|_\infty /\varepsilon$, $\boldsymbol{\eta}^{\overline{u}}$-a.s., and similarly $\rho \circ e_s \leq \|\rho\|_\infty /\varepsilon$, $\boldsymbol{\eta}^{\overline{v}}$-a.s.. Moreover, as $\rho \circ e_s$ is clearly $\mathcal{F}_s$-measurable, by Proposition 6.9, the claim is proved.

Now fix any $\overline{v} \in L^r(\mathfrak{m})$, with $\overline{v}\mathfrak{m}$ probability and $\overline{v} > 0$, $\mathfrak{m}$-a.e. in X, and define $\boldsymbol{\eta}(s,x) := \boldsymbol{\eta}^{\overline{v}}(s,x)$. Our aim is to show that
$$\boldsymbol{\eta}^{\overline{u}} = \int \boldsymbol{\eta}(s,x)\overline{u}(x)d\mathfrak{m}(x). \tag{6.13}$$
For $n \geq 1$, let $\overline{u}_n := c_n(\overline{u}(x) \wedge n)$, where c_n ensures that $\overline{u}_n$ is a probability density. Due to the claim, it holds
$$\int \boldsymbol{\eta}^{\overline{u}}(s,x)c_n(\overline{u}(x) \wedge n)d\mathfrak{m}(x) = \boldsymbol{\eta}^{\overline{u}_n} = \int \boldsymbol{\eta}^{\overline{v}}(s,x)c_n(\overline{u}(x) \wedge n)d\mathfrak{m}(x),$$
thus
$$\int \boldsymbol{\eta}^{\overline{u}}(s,x)(\overline{u}(x) \wedge n)d\mathfrak{m}(x) = \int \boldsymbol{\eta}(s,x)(\overline{u}(x) \wedge n)d\mathfrak{m}(x),$$
and so, as $n \to \infty$, by monotone convergence we obtain (6.13).

ii) $\Rightarrow$ i) . Let us consider the general case first, highlight then where the assumption $r = \infty$ seems crucial. Existence of L^r-regular solutions to the martingale problem, given the existence of a L^r-regular martingale flow follows trivially from (6.9). To prove uniqueness we argue by contradiction: assume that there exists $\overline{s} \in [0,T]$ and $\overline{v} \in L^r(\mathfrak{m})$, with $\overline{v}\mathfrak{m}$ probability, and $\boldsymbol{\eta}^1$, $\boldsymbol{\eta}^2 \in C_{\overline{s},\overline{v}}$ with $\boldsymbol{\eta}^1 \neq \boldsymbol{\eta}^2$. By disintegration with respect to $e_{\overline{s}}$, we obtain two Borel families of probability measures $(\boldsymbol{\eta}^1(\overline{s},x))_{x\in X}$, $(\boldsymbol{\eta}^2(\overline{s},x))_{x\in X}$ that differ on a Borel set $A \subseteq X$ with $\mathfrak{m}(A) > 0$ and $\overline{v} > \varepsilon$ on A, for some fixed $\varepsilon > 0$.

By the existence assumption, let $(\boldsymbol{\eta}(s,x))_{s,x}$ be a martingale flow and then modify it only on $\{s\} \times A$, letting $\boldsymbol{\eta}(\overline{s},x) = \boldsymbol{\eta}^i(\overline{s},x)$, for $x \in A$, $i \in \{1,2\}$. Clearly, the two maps obtained are Borel, and we would like to argue that they provide two different L^r-regular martingale flows, in contrast with the uniqueness assumption. To check that they define indeed L^r-regular martingale flows, we rely on the identity, valid for any $\overline{u} \in L^r(\mathfrak{m})$, with $\overline{u}\mathfrak{m}$ probability,
$$\begin{aligned}\int \boldsymbol{\eta}^i(s,x)\overline{u}(x)d\mathfrak{m}(x) &= \int_{A^c} \boldsymbol{\eta}(s,x)\overline{u}(x)d\mathfrak{m}(x) + \int_A \boldsymbol{\eta}^i(s,x)\overline{u}(x)d\mathfrak{m}(x)\\ &= (1-p)\int_{A^c} \boldsymbol{\eta}(s,x)\frac{\overline{u}(x)}{1-p}d\mathfrak{m}(x) + p\int_A \boldsymbol{\eta}^i(s,x)\frac{\overline{u}(x)}{p}d\mathfrak{m}(x)\end{aligned}$$
where $p = \int_A \overline{u}d\mathfrak{m}$, provided that this quantity is non-null, otherwise there is nothing to prove. This identity follows from Proposition 6.10, with Z being the two point space. However, to apply this last result, one has to show that both terms above are L^r-regular solutions to the martingale problem: for the first one, we use the definition of L^r-regular martingale flow, while for the second one we must restrict ourselves to the case $r = \infty$, and use the inequality
$$\frac{\overline{u}}{p}\chi_A \leq \frac{\overline{u}}{p}\chi_{\{\overline{v}>\varepsilon\}} \leq \frac{\|\overline{u}\|_\infty}{p\varepsilon}\overline{v}, \quad \mathfrak{m}\text{-a.e. in } X,$$
that gives, by definition of disintegration with respect to $e_{\overline{s}}$,
$$\int_A \boldsymbol{\eta}^i(s,x)\frac{\overline{u}(x)}{p}d\mathfrak{m}(x) \leq \frac{\|\overline{u}\|_\infty}{p\varepsilon}\boldsymbol{\eta}^i,$$

and so by Proposition 6.9, we conclude.

To prove (6.10) assuming *i)*, it is enough fix $\overline{u} \in L^r(\mathfrak{m})$, with $\overline{u}\mathfrak{m}$ probability, and repeat the corresponding argument in the proof of Lemma 1.16, to obtain

$$\pi_\sharp(\int \boldsymbol{\eta}(s,x)\overline{u}(x)d\mathfrak{m}(x)) = \int_X \boldsymbol{\eta}(r,y)\left[\int_X \eta(s,x)_r(dy)\overline{u}(x)d\mathfrak{m}(x)\right]$$

which entails (6.10), being $\overline{u}$ arbitrary. □

Remark 6.17 (Regular flows in the deterministic case)**.** When $\mathcal{L} = \boldsymbol{b}$ is a derivation, as observed in Remark 6.7, solutions to the martingale problem reduce to curves concentrated on solutions to the ODE induced by $\boldsymbol{b}$ as in [Ambrosio and Trevisan, 2014, Definition 7.3]. It is then natural to question whether uniqueness of L^∞-regular martingale flows entails that the flow is actually deterministic, i.e. it holds $\boldsymbol{\eta}(s,x) = \delta_{\mathbb{X}(s,x)}$, for some Borel function $\mathbb{X}(s,\cdot) : X \times [0,T] \to X$ (for every fixed $s \in [0,T]$), thus fully recovering the results from [Ambrosio and Trevisan, 2014, §8], which generalize Ambrosio's approach to DiPerna-Lions theory to metric measure spaces. Indeed, it is sufficient to prove that the marginal $\eta(s,x)_t$ is a degenerate measure for every (rational) $t \in [0,T]$, and this can be performed by arguing verbatim as in [Ambrosio and Crippa, 2008, Theorem 18], where the original argument of [Ambrosio, 2004, Theorem 5.4] is slightly improved. For the sake of completeness, we report it right below. ■

Theorem 6.18 (no splitting criterion)**.** *Assume that $\mathcal{L} = \boldsymbol{b}$ is a derivation, and assume that there exists a $\mathfrak{m}$-essentially unique L^r-regular martingale flow $(\boldsymbol{\eta}(s,x))_{s,x}$. Then, $\boldsymbol{\eta}(s,x)$ is a Dirac measure at a single curve $\mathbb{X}(s,x) \in C([0,T];X)$, for $\mathfrak{m}$-a.e. $x \in X$, for every $s \in [0,T]$. Moreover,* (6.10) *reads as $\mathbb{X}(s,x)(t) = \mathbb{X}(r,\mathbb{X}(s,x))(t)$, for $\mathfrak{m}$-a.e. $x \in X$, r, s, $t \in [0,T]$, with $s \le r \le t$.*

Proof. By contradiction, we assume that there exists $s \in [0,T]$ such that $\boldsymbol{\eta}(s,x)$ is not a Dirac measure at a single curve in a Borel set of positive measure $A \subseteq X$: without any loss of generality, we let $s = 0$ and omit to write it in what follows. A simple argument (see [Ambrosio and Crippa, 2008, Lemma 15]) gives that there exists disjoint sets E, $E' \subseteq X$ and $t \in [0,T]$ with $\eta(x)_t(E)\eta(x)_t(E') > 0$, for $x \in A$. Possibly reducing to a smaller set, we may let both $\eta(x)_t(E) > \varepsilon$ and $\eta(x)_t(E') > \varepsilon$, for $x \in A$, for some $\varepsilon \in (0,1)$. Then, we consider the Borel functions on $X \times C([0,T];X)$,

$$\rho_E(x)(\gamma) := \frac{\chi_{\gamma(t)\in E}}{\eta(x)_t(E)}, \quad \rho_{E'}(x)(\gamma) := \frac{\chi_{\gamma(t)\in E'}}{\eta(x)_t(E')},$$

and the Borel family of probability measures $\boldsymbol{\eta}_E$ (respectively, $\boldsymbol{\eta}_{E'}$), obtained by the flow $\boldsymbol{\eta}$ by replacing $\boldsymbol{\eta}(0,x)$ with $\rho_E(x)\boldsymbol{\eta}(0,x)$ (respectively $\rho_{E'}(x)\boldsymbol{\eta}(0,x)$). By construction, it holds $\int_A \boldsymbol{\eta}_E(0,x)d\mathfrak{m}(x) \neq \int_A \boldsymbol{\eta}_{E'}(0,x)d\mathfrak{m}(x)$, since the marginals at t are concentrated respectively at E and E' which are disjoint.

To obtain a contradiction, we only have to prove that $\boldsymbol{\eta}_E$ and $\boldsymbol{\eta}_{E'}$ provide L^r-regular martingale flows. Thus, we let $\bar{u} \in L^r(\mathfrak{m})$, with $\bar{u}\mathfrak{m}$ probability. Since $\rho_E \le \varepsilon^{-1}\boldsymbol{\eta}$, we obtain that, the marginals of $\boldsymbol{\eta}_E$ are bounded above by those of $\boldsymbol{\eta}$, thus L^r-regularity holds for $\int \boldsymbol{\eta}_E(0,x)\bar{u}d\mathfrak{m}$ (and the same argument applies to E'). To show that the martingale property holds, one would like to apply Proposition 6.9 with $\rho(\gamma) := \rho_E(e_0)(\gamma)$, $\gamma \in C([0,T];X)$, but this clearly violates the condition of being $\mathcal{F}_0$-measurable (in general, it is only $\mathcal{F}_t$ measurable).

It is precisely at this point that the deterministic assumption enters the picture: by Remark 6.7, the martingales associated to $\boldsymbol{\eta}$ are in fact constant, or equivalently, they are martingales with respect to the constant filtration equal to the full Borel σ-algebra: thus the measurability assumption on ρ in Proposition 6.9 can be removed and we conclude indeed that $\boldsymbol{\eta}_E$ provide an L^r-regular martingale flow (and similarly, $\boldsymbol{\eta}_{E'}$). $\square$

Chapter 7

The superposition principle in metric measure spaces

We conclude the second part of this thesis by establishing the validity of a superposition principle for diffusion processes in metric measure spaces, under suitable assumptions. The argument we employ is close to that of [Ambrosio and Trevisan, 2014, Theorem 7.6], where the deterministic case is settled, in turn influenced by a change of variables appearing in [Kolesnikov and Röckner, 2014]. The strategy roughly consists in establishing first a superposition principle in the space $\mathbb{R}^\infty$, in Section 7.1, obtained as a limit of Euclidean spaces, and then of transferring it to general spaces, after some technical preliminaries, that we address in Section 7.2. Let us point out that, since we are dealing with diffusion processes, there appears some connection between these techniques and those employed in the proof of existence for Markov processes associated to (quasi-)regular Dirichlet forms.

7.1 Diffusions in $\mathbb{R}^\infty$

In this section, we establish a superposition principle for diffusions on $\mathbb{R}^\infty$, following the strategy in [Ambrosio and Trevisan, 2014, Theorem 7.2], it is also pointed out towards the end of [Ambrosio et al., 2008, §8.2] that one can extend the superposition principle, in the deterministic case, from $\mathbb{R}^d$ to Hilbert spaces.

Let us briefly describe the setting of $\mathbb{R}^\infty$, in particular suitable notions of diffusion operators, Fokker-Planck equations and martingale problems.

Ambient space

We endow $\mathbb{R}^\infty := \mathbb{R}^\mathbb{N}$ with the complete and separable distance

$$(x, y) \mapsto \sum_{i=1}^{\infty} 2^{-i} \min\left\{1, |x^i - y^i|\right\},$$

inducing the product topology, and we let

$$\pi^d : \mathbb{R}^\infty \to \mathbb{R}^d, \quad \pi^d(x) = (x^1, \ldots, x^d)$$

denote the canonical projections from $\mathbb{R}^\infty$ to $\mathbb{R}^d$, for $d \geq 1$. We also endow the space $C([0,T];\mathbb{R}^\infty)$ with the distance

$$(\gamma, \tilde{\gamma}) \mapsto \sum_{i=1}^{\infty} 2^{-i} \max_{t\in[0,T]} \min\left\{1, |\gamma(t)^i - \tilde{\gamma}(t)^i|\right\},$$

which makes $C([0,T];\mathbb{R}^\infty)$ complete and separable as well. Notice that a set $K \subseteq C([0,T];\mathbb{R}^\infty)$ is compact if and only if the set $\{x^n \circ \gamma : \gamma \in K\}$ is compact in $C([0,T];\mathbb{R})$, for every $n \geq 1$. Therefore, if for $i \geq 1$ we let $\Psi^i : C([0,T];\mathbb{R}) \to [0,\infty]$ be any coercive functional, then

$$C([0,T];\mathbb{R}^\infty) \ni \gamma \quad \mapsto \quad \Psi(\gamma) := \sum_{i=1}^{\infty} \Psi^i(x^i \circ \gamma_t)$$

is coercive.

Test functions and diffusion operators

We let $\mathscr{A}^\infty := \mathcal{F}C_b^{1,2}((0,T)\times\mathbb{R}^\infty)$ be the class of $C_b^{1,2}$ cylindrical functions, i.e. those $f : (0,T)\times\mathbb{R}^\infty \to \mathbb{R}$ that can be written in the form

$$f_t(x) = f_t(\pi^n(x)) = f_t\left(x^1, \ldots, x^n\right) \qquad x \in \mathbb{R}^\infty,$$

for some $n \geq 1$ and $f \in \mathscr{A}^n := C_b^{1,2}((0,T)\times\mathbb{R}^n)$. We write

$$\partial_i f_t := (\partial_i f_t)\circ\pi^n, \quad \partial_{i,j} f_t := (\partial_{i,j} f_t)\circ\pi^n \quad \text{for } i,j \geq 1,\, t\in(0,T),$$

so that $\partial_i f = 0$ for $i > n$ and we say that f is n-cylindrical.

Given Borel maps

$$a = (a^{i,j})_{i,j=1}^\infty : (0,T)\times\mathbb{R}^\infty \to \mathrm{Sym}_+(\mathbb{R}^\infty), \quad b = (b^i)_{i=1}^\infty : (0,T)\times\mathbb{R}^\infty \to \mathbb{R}^\infty, \tag{7.1}$$

where $\mathrm{Sym}_+(\mathbb{R}^\infty)$ is defined as the set of double sequences $(a^{i,j})$ such that,

$$\sum_{i,j=1}^{n} a^{i,j}\xi^i\xi^j \geq 0, \quad \text{for every } n \geq 1, \text{ for every } \xi \in \mathbb{R}^\infty,$$

we define the diffusion operator $\mathcal{L} := \mathcal{L}(a,b)$, mapping $\mathscr{A}$ into Borel functions on $(0,T)\times\mathbb{R}^\infty$,

$$\mathcal{L}_t f := \sum_{i,j=1}^{\infty} a_t^{i,j}\partial_{i,j}^2 f_t + \sum_{i=1}^{\infty} b_t^i \partial_i f,$$

and notice that the series are actually finite sums.

FPE's and MP's

Solutions $\nu = (\nu_t)_{t\in(0,T)} \subseteq \mathscr{M}(\mathbb{R}^\infty)$ to the FPE associated to $\mathcal{L}(a,b)$ are introduced exactly as in Definition 6.1, where we replace the condition of $\mathcal{L}$ being locally in $L^1(\nu)$ with

$$\int_0^T \int \left[|a_t^{i,j}| + |b_t^i|\right] d\,|\nu_t|\, dt, \qquad \text{for } i,j \geq 1.$$

Solutions to the martingale problem associated to $\mathcal{L}(a,b)$ are introduced as in Definition 6.6.

The superposition principle

In this framework, superposition solutions are defined exactly as in Definition 6.13.

Theorem 7.1 (superposition principle for diffusions in $\mathbb{R}^\infty$)**.** *Let a, b be Borel maps as in (7.1). Then, the superposition principle holds for every solution $\nu = (\nu_t)_{t\in(0,T)} \subseteq \mathscr{P}(\mathbb{R}^\infty)$ to the FPE associated to $\mathcal{L}(a,b)$ provided that, for some $p \in (1,\infty)$, it holds*

$$\int_0^T \int \left[|a_t^{i,j}|^p + |b_t^i|^p\right] d\nu_t dt, \quad \textit{for } i,j \geq 1. \tag{7.2}$$

Proof. We rely once again on the scheme introduced in Section 2.2.

Step 1 (approximation)**.** By means of push-forwards with cylindrical maps, we reduce the problem in $\mathbb{R}^\infty$ to a sequence of problems in $\mathbb{R}^d$, for $d \geq 1$. Precisely, we let $\tilde\nu_t^d := \pi^d_\sharp(\nu_t) \in \mathscr{P}(\mathbb{R}^d)$, for $t \in [0,T]$, $d \geq 1$. Arguing as in the first part of Section 2.2.1, $\tilde\nu^d$ is a solution to the FPE associated to $\mathcal{L}(\tilde a^d, \tilde b^d)$, where

$$(\tilde a^d)_t^{i,j}(y) := \mathbb{E}_{\nu_t}\left[a_t^{i,j} \,|\, \pi^d = y\right], \quad (\tilde b^d)_t^i(y) := \mathbb{E}_{\nu_t}\left[b_t^i \,|\, \pi^d = y\right], \quad \text{for } i,j \geq 1,\, y \in \mathbb{R}^d,.$$

By Theorem 2.14, there exists some superposition solution $\tilde{\boldsymbol\eta}^d \in \mathscr{P}(C([0,T];\mathbb{R}^d))$.

To address tightness, it is better to embed $\tilde{\boldsymbol\eta}^d$ into $\mathscr{P}(C([0,T];\mathbb{R}^\infty))$, by means of the inclusion

$$J^d : [0,T]\times\mathbb{R}^d \ni (t,y) \mapsto (t,y,0,\ldots) \in [0,T]\times\mathbb{R}^\infty, \quad \text{for } d \geq 1.$$

One obtains at once that $\boldsymbol\eta^d := J^d_\sharp(\tilde{\boldsymbol\eta}^d) \in \mathscr{P}(C([0,T];\mathbb{R}^\infty))$ is a solution to the martingale problem associated to $\mathcal{L}(a^d, b^d)$, where

$$(a^d)_t^{i,j} := \mathbb{E}_{\nu_t}\left[a_t^{i,j} \,|\, \pi^d\right], \quad (b^d)_t^i := \mathbb{E}_{\nu_t}\left[b_t^i \,|\, \pi^d\right], \quad \text{for } i,j \geq 1.$$

Step 2 (tightness)**.** By Corollary 2.11, there exists some constant C depending only on p, $\delta \in (0, 1-1/p)$ and T only such that

$$\mathbb{E}\left[\mathcal{A}^i\right] \leq C \int \left[|(b^d)^i|^p + |(a^d)^{i,i}|^p\right] d\nu^d, \text{ for } i \geq 1, \tag{7.3}$$

where we let $\mathcal{A}^i(\gamma) := \mathcal{A}(x^i\circ\gamma)$ and $\mathcal{A}$ defined by (2.7).

To conclude, we notice that measures ν_0^d are tight, as they narrowly converge towards ν_0, so there exists a coercive functional $\psi : \mathbb{R}^\infty \mapsto [0,\infty]$ such that $\sup_{d\geq1}\int \psi d\nu_0^d < \infty$. Finally, we let

$$\mathscr{P}(C([0,T];\mathbb{R}^d)) \ni \gamma \mapsto \Psi(\gamma) := \psi\circ e_0 + \sum_{i=1}^\infty c_i \mathcal{A}^i,$$

where c_i is defined as 2^{-i} times right hand side in (7.3), for $i \geq 1$. Therefore, the functional Ψ is coercive and

$$\sup_{d\geq1}\int \Psi d\boldsymbol\eta^d < \infty,$$

thus the family $(\boldsymbol\eta^d)_d$ is tight.

Step 2 (limit). We argue as in the first part of Section 2.2.3, with two notable differences. First, we let $\tilde{a}$, $\tilde{b}$ be bounded continuous n-cylindrical maps, i.e.

$$\tilde{a} = J^n \circ a^n \circ \pi^n, \quad \tilde{b} = J^n \circ b^n \circ \pi^n,$$

for some $n \geq 1$ and bounded continuous maps

$$a^n : (0,T) \times \mathbb{R}^n \to \mathrm{Sym}_+(\mathbb{R}^n), \quad b^n : (0,T) \times \mathbb{R}^n \to \mathbb{R}^n.$$

Then, we notice that in the derivation of (2.13), the sum does not need to be replaced by a full series (which we are not assuming to converge) but it can be extended only up to some $k \geq 1$, depending on the fact that $f \in \mathscr{A}^\infty$ is k-cylindrical. Thanks to these remarks the thesis follows by density of continuous cylindrical functions in $L^1(\nu)$. □

7.2 Diffusions in metric measure spaces

We consider now the setting of the previous chapter, i.e. we let (X, τ), $\mathfrak{m}$, $\mathcal{E}$ be as in Section 3.1 fix $T > 0$, and consider the time-extended framework described in Section 5.1, letting in particular $\mathscr{A} := \tilde{\mathscr{A}}$ satisfy (5.2). We fix $r > 1$ and let $\mathcal{L}$ be a diffusion operator such that (6.1) holds.

Together with the superposition principle, we with to obtain Hölder continuity for the paths of solution to the martingale problem. However, X enjoys only a topology τ, a σ-finite measure $\mathfrak{m}$, and a Dirichlet form $\mathscr{E}$, with associated P, Δ and Γ, but there appears to be no distance. To overcome this situation, we introduce a further condition in the framework, namely existence of a countable set $\mathscr{A}^* = \{f_1, f_2, \ldots\} \subset \mathscr{A}$ of functions admitting a continuous representative. Since we assume that $\operatorname{supp} \mathfrak{m} = X$, the continuous representative of a Borel function, if it exists, is unique, and for this reason we do not use above or in the sequel a distinguished notation. For simplicity, we also assume that functions in $\mathscr{A}^*$ are constant with respect to $t \in (0,T)$.

As a first condition, we require that the family $\mathscr{A}^*$ separates the points:

$$\text{for every } x,\, y \in X, \text{ there exists } f \in \mathscr{A}^* \text{ with } f(x) \neq f(y). \tag{7.4}$$

Starting from $\mathscr{A}^*$, we may build several distances on X. To discuss diffusions, we introduce $d : X \times X \to [0, \infty)$, given by

$$d(x,y) := \sum_{i=1}^{\infty} 2^{-i} \min\left\{ |f_i(x) - f_i(y)|, 1 \right\},$$

which clearly satisfies the axioms of distances: in particular, it holds $d(x,y) = 0$ if and only if $x = y$, because of (7.4). Notice that, for $i \geq 1$, each f_i is uniformly continuous with respect to d, since it holds

$$d(x,y) < 2^{-i} \Rightarrow |f_i(x) - f_i(y)| < 2^i d(x,y), \quad \text{for all } x,\, y \in X,$$

and the topology induced by d is coarser than τ.

Remark 7.2. In the deterministic case, it is more natural to introduce the (possibly extended) distance

$$d_\infty(x,y) := \sup\left\{|f(x)-f(y)| \,:\, f \in \mathscr{A}^*\right\}. \tag{7.5}$$

Then, d_∞ induces a topology, in general, finer than that of d, and one is naturally led to consider extended Polish space in the sense of [Ambrosio et al., 2014a, Definition 2.3]. Under suitable assumptions, one proves that solutions to the ODE are concentrated on absolutely continuous curves with respect to d_∞, see Lemma 7.6, which reports [Ambrosio and Trevisan, 2014, Lemma 7.4]. For general diffusions, such a result clearly does not have an immediate counterpart, already because of the consequent absolute continuity of paths, but there also are subtler reasons, due to the difficulty of exchanging expectation and supremum. ■

As another condition on $\mathscr{A}^*$, we require that

$$\exists \lim f(x_n) \text{ in } \mathbb{R} \text{ for all } f \in \mathscr{A}^*, \quad \Rightarrow \quad \exists \lim x_n \text{ in } X, \qquad \text{for every } (x_n)_n \subseteq X, \tag{7.6}$$

i.e. d induces the topology τ on X, see Remark 7.4.

We are in a position to state and prove the main result in this chapter, which provides an extension to diffusions of [Ambrosio and Trevisan, 2014, Theorem 7.6]. In the determisitic case, it seems that less assumptions are required, and this is indeed the case, mainly because of two reasons: the chain rule for diffusion operators does not hold for general diffusions and we prove it, in Proposition 4.23, under an integrability assumption on $\operatorname{div}\mathcal{L}$; our superposition principle for diffusions in $\mathbb{R}^\infty$ require L^p-integrability on coefficients, for some p strictly greater than 1. Despite this fact, it provides a satisfactory result, in particular because our well-posedness results for Fokker-Planck equations rely bounds on $\operatorname{div}\mathcal{L}$.

Theorem 7.3 (superposition principle for diffusions)**.** *Assume that* $\operatorname{div}\mathcal{L} \in L^1_t(L^1_x)$, $\boldsymbol{a} \in L^1_t(L^r_x)$ *and that* $\mathscr{A} \subseteq L^\infty_t(\mathbb{V}^\infty)$. *Let* $\mathscr{A}^* \subseteq \mathscr{A}$ *satisfy* (7.4), (7.6), *and the set of functions*

$$g := \Phi(\cdot, f_1, \ldots, f_n), \textit{ for } \Phi \in C^2_b((0,T)\times\mathbb{R}^n) \textit{ with } \Phi(\cdot,0)=0, \textit{ for } f_1,\ldots,f_n \in \mathscr{A}^*,$$

for $n \geq 1$, *be dense in* $\mathscr{A}$ *in the following sense: for every* $g \in \mathscr{A}$ *one can find a* $(g_k)_k$ *as above such that, as* $k \to \infty$,

$$g_k \to g, \quad \partial_t g_k \to \partial_t g, \quad \mathcal{L} g_k \to \mathcal{L} g, \quad \textit{weakly in } L^1_t(L^{r'}_x). \tag{7.7}$$

Then, every solution $u = (u_t)_{t\in(0,T)} \subseteq L^\infty_t(L^r_x)$ *to the FPE*

$$\partial_t u_t = \mathcal{L}^*_t u_t, \quad \textit{in } (0,T)\times X.$$

with $u_t\mathfrak{m}$ *probability, for* $\mathscr{L}^1$*-a.e.* $t \in (0,T)$, *and, for some* $p > 1$,

$$\int_0^T \int \left[|\mathcal{L}_t f|^p + |\boldsymbol{a}_t(f,g)|^p\right] u_t d\mathfrak{m} dt < \infty, \quad \textit{for every } f,\, g \in \mathscr{A}^*,$$

admits a superposition solution $\boldsymbol{\eta}$.

Proof. The proof relies first on a transfer argument from the superposition principle in $\mathbb{R}^\infty$, namely Theorem 7.1, and then on a density argument. When we perform the transfer by means of the push-forward, actually, parts of our deductions are obtained exactly as in Section

2.2.1, but we repeat them here for the sake of clarity and because the setting there is restricted to $\mathbb{R}^d$.

We define a map $J : X \to \mathbb{R}^\infty$, letting

$$J(x) := \big(f^1(x), f^2(x), f^3(x), \ldots\big),$$

where we endow $\mathbb{R}^\infty = \mathbb{R}^\mathbb{N}$ with the topology induced by the distance introduced in Section 7.1. A simple consequence of (7.4) is that J injective, while (7.6) entails that $J(X)$ is a closed subset of $\mathbb{R}^\infty$ and that J^{-1} is continuous from $J(X)$ to X: actually, J is an isometry.

We define $\nu_t \in \mathscr{P}(\mathbb{R}^\infty)$ by $\nu_t := J_\#(u_t\mathfrak{m})$, a vector field $b : (0,T) \times \mathbb{R}^\infty \to \mathbb{R}^\infty$ by

$$b^i_t := \begin{cases} \mathcal{L}_t(f^i) \circ J^{-1} & \text{on } J(X); \\ \\ 0 & \text{otherwise,} \end{cases}$$

and a map $a : (0,T) \times \mathbb{R}^\infty \to \mathrm{Sym}_+(\mathbb{R}^\infty)$ by

$$a^{i,j}_t := \begin{cases} \boldsymbol{a}_t(f^i, f^j) \circ J^{-1} & \text{on } J(X); \\ \\ 0 & \text{otherwise,} \end{cases}$$

and we notice that

$$|b^i| \circ J \le |\mathcal{L}(f^i)|, \quad |a^{i,j}| \circ J \le \boldsymbol{a}(f^i, f^j) \quad \tilde{\mathfrak{m}}\text{-a.e. in } (0,T) \times X,$$

thus (7.2) holds, with $p = q$. The chain rule for diffusions, Proposition 4.23 shows that $\nu = (\nu_t)_{t\in[0,T]}$ is a solution to the FPE associated to the diffusion operator in $\mathbb{R}^\infty$ with coefficients a, b, thus the assumptions of Theorem 7.1 are satisfied with $\nu = (\nu_t)_t$.

As a consequence, we can apply Theorem 7.1 to obtain a superposition solution $\tilde{\boldsymbol{\eta}} \in \mathscr{P}(C([0,T]; \mathbb{R}^\infty))$. Since all measures ν_t are concentrated on $J(X)$, one has

$$\gamma(t) \in J(X) \text{ for } \tilde{\boldsymbol{\eta}}\text{-a.e. } \gamma, \text{ for all } t \in [0,T] \cap \mathbb{Q}.$$

Then, the closedness of $J(X)$ and the continuity of γ give $\gamma([0,T]) \subset J(X)$ for $\tilde{\boldsymbol{\eta}}$-a.e. γ. For this reason, it makes sense to define

$$\boldsymbol{\eta} := \Theta_\# \tilde{\boldsymbol{\eta}},$$

where $\Theta : C([0,T]; J(X)) \to C([0,T]; X)$ is the map $\gamma \mapsto \Theta(\gamma) := J^{-1} \circ \gamma$. Since $(J^{-1})_\#\nu_t = u_t\mathfrak{m}$, $\mathcal{L}$-a.e. $t \in (0,T)$, we obtain immediately that $(e_t)_\#\boldsymbol{\eta} = \mu_t$, $\mathcal{L}^1$-a.e. $t \in (0,T)$.

We complete the proof by showing that $\boldsymbol{\eta}$ provides a superposition solution for u. Since $f^i \circ \Theta(\gamma) = x^i \circ \gamma$, for $\gamma \in J(X)$, taking the definition of a and b into account, we obtain that the process

$$[0,T] \ni t \mapsto M^i_t := f^i \circ e_t - \int_0^t \mathcal{L}_s(f^i) \circ e_s \, ds \tag{7.8}$$

is a continuous martingale, and that the covariation processes are

$$[M^i, M^j]_t = 2\int_0^t \boldsymbol{a}_s(f^i, f^j) \circ e_s ds, \quad \text{for } i,\ j \ge 1.$$

To extend the martingale property from $f^i \in \mathscr{A}^*$ to $f \in \mathscr{A}$, by Itô formula and the chain rule for diffusions, (7.8) from f^i to compositions $(t,x) \mapsto \Phi(t, f^1(x), \ldots, f^n(x))$. By the density assumptions (7.7), we pass to the limit the martingale property, in the integral formulation, and conclude. □

Remark 7.4. So far, it is clear that the topology τ played no significant role in our deductions. We claim therefore that, up to replacing the ambient space (X,τ) with $(\overline{X},\tau_d)$, where τ_d is the topology generated by d on the abstract completion of X with respect to the metric d, condition (7.6) holds. We denote by $\iota:(X,\tau)\subseteq(\overline{X},\tau_d)$ be the continuous inclusion.

Every $f\in\mathscr{A}^*$, being uniformly continuous, extends uniquely to a continuous function on $\overline{X}$, that we still denote by f, and the the distance d on $\overline{X}$ is still represented by the same series. By taking push-forwards of the measure $\mathfrak{m}$ (i.e. $\overline{\mathfrak{m}}:=\iota_\sharp\mathfrak{m}$) the Dirichlet form $\mathcal{E}$ (i.e. $\overline{\mathcal{E}}(f):=\mathcal{E}(f\circ\iota)$, for $f\in L^2(\overline{\mathfrak{m}})$) and the algebra $\mathscr{A}$, (i.e. $\left\{f\circ\iota^{-1}\,|\,f\in\mathscr{A}\right\}$) we obtain an enlarged but equivalent structure, at least with respect to the structural conditions (3.1) and (5.2); but then, (7.6) holds too. Notice how this construction reminds the way one reduces quasi-regular Dirichlet forms to regular ones, see e.g. [Chen and Fukushima, 2012, Theorem 1.4.3].

By replacing the ambient space X with $\overline{X}$, the only remarkable difference is that provide solutions to the MP as probability measures on continuous curves with values in $\overline{X}$: we discuss this problem in the examples of Part IV, in particular with respect to infinite dimensional spaces. ■

One may also look for criteria entailing stronger continuity properties, for every $t\in[0,T]$, $\boldsymbol{\eta}$-a.s., as discussed in the following results.

Lemma 7.5. *Let $\boldsymbol{\eta}\in\mathscr{P}(C([0,T];X)$ be a L^r-regular solution to the MP associated to $\mathcal{L}$ and define, for $p\in(1,\infty)$,*

$$\|\boldsymbol{b}\|_p^*:=\sup\left\{\int_0^T\int|\mathcal{L}_t(f)|^p\,d\eta_t dt\,:\,f\in\mathscr{A}^*\right\},\quad \|\boldsymbol{a}\|_p^*:=\sup\left\{\int_0^T\int\boldsymbol{a}(f)^p d\eta_t dt\,:\,f\in\mathscr{A}^*\right\}.$$

Then, for every $\delta\in(0,1-1/p)$ it holds, for some constant C depending on p, δ and T only,

$$\int\mathcal{A}(\gamma)d\boldsymbol{\eta}(\gamma)\le C\left(\|\boldsymbol{b}\|_p^*+\sqrt{\|\boldsymbol{a}\|_p^*}\right)^{1/p},$$

where $\mathcal{A}$ is defined as in (2.7), replacing $\mathbb{R}$ with X, endowed with the distance d.

Proof. The arguments in Lemma 2.8 and the subsequent remarks, leading to Corollary 2.11, entail the bound

$$\int\mathcal{A}(\gamma)d\boldsymbol{\eta}(\gamma)\le C\sum_{i=1}^\infty 2^{-i}\left\{\left[\int_0^T\int|\mathcal{L}_t f_i|^p\,d\eta_t\,dt\right]^{1/p}+\left[\int_0^T\int\boldsymbol{a}_t(f_i)^p d\eta_t\,dt\right]^{1/2p}\right\},$$

so that the conclusion is immediate. □

In the deterministic case, the situation is much clearer and we prove a precise result, with respect to the distance d_∞, which is expected to induced a topology equal or even finer than τ, when $\mathscr{A}^*$ is chosen appropriately, see [Ambrosio and Trevisan, 2014, Lemma 9.2].

Lemma 7.6. *Let $\boldsymbol{\eta}\in\mathscr{P}(C([0,T];X))$ be a solution to the MP associated to $\mathcal{L}=\boldsymbol{b}$ a derivation with $|\boldsymbol{b}|\in L^1_t(L^p_x)$, for some $p\in[1,\infty]$. Then, $\boldsymbol{\eta}$ is concentrated on $AC^p([0,T];(X,d_\infty))$, with*

$$|\dot\gamma|(t)=|\boldsymbol{b}_t|_*(\gamma_t)\quad \textit{for a.e. } t\in(0,T),\textit{ for }\boldsymbol{\eta}\textit{-a.e. }\gamma,$$

where $|\boldsymbol{b}_t|^*:=\sup_{f\in\mathscr{A}^*}\{|df(\boldsymbol{b}_t)|\}$.

Proof. Given $f \in \mathscr{A}^*$, for $\boldsymbol{\eta}$-a.e. γ, the map $t \mapsto f \circ \gamma_t$ is absolutely continuous, with

$$f \circ \gamma_t - f \circ \gamma_s = \int_s^t df(\boldsymbol{b}_r)(\gamma_r)dr, \quad \text{for all } s,\, t \in [0,T].$$

In particular one has $df(\boldsymbol{b}_t)(\gamma_t) = (f \circ \gamma)'(t)$ a.e. in $(0,T)$, for $\boldsymbol{\eta}$-a.e. γ.

By Fubini's theorem and the fact that the marginals of $\boldsymbol{\eta}$ are absolutely continuous w.r.t. $\mathfrak{m}$ we obtain that, for $\boldsymbol{\eta}$-a.e. γ, one has

$$\sup_{f \in \mathscr{A}^*} \left|(f \circ \gamma)'(t)\right| = \sup_{f \in \mathscr{A}^*} |df(\boldsymbol{b}_t)(\gamma_t)| = \left|\boldsymbol{b}_t\right|_* (\gamma_t), \quad \text{for a.e. } t \in (0,T),$$

and therefore

$$d_{\mathscr{A}^*}(\gamma_t, \gamma_s) = \sup_{f \in \mathscr{A}^*} \left|(f \circ \gamma)(t) - (f \circ \gamma)(s)\right| \le \int_s^t \left|\boldsymbol{b}_r\right|_* (\gamma_r)dr, \quad \text{for all } s,\, t \in [0,T],$$

proving that $\gamma \in AC([0,T];(X, d_{\mathscr{A}^*}))$, with $|\dot\gamma|(t) \le \left|\boldsymbol{b}_t\right|_* (\gamma_t)$, for a.e. $t \in (0,T)$. The converse inequality follows from the fact that every $f \in \mathscr{A}^*$ is 1-Lipschitz with respect to $d_{\mathscr{A}^*}$, thus for $\boldsymbol{\eta}$-a.e. γ one has

$$\left|\boldsymbol{b}_t\right|_* (\gamma_t) = \sup_{f \in \mathscr{A}^*} \left|(f \circ \gamma)'(t)\right| \le |\dot\gamma|(t), \quad \text{for a.e. } t \in (0,T).$$

$\square$

Part III

Fokker-Planck equations in metric measure spaces

Chapter 8

Formal energy estimates

In this chapter we study formal estimates satisfied by solutions to the Fokker-Planck equation

$$\partial_t u_t = \mathcal{L}_t^* u_t, \quad \text{in } (0,T)\times X, \quad u_0 = \overline{u}, \tag{8.1}$$

extending the usual a-priori estimates in the theory of parabolic equations. In the deterministic case, i.e. when $\mathcal{L} = \boldsymbol{b}$ is a derivation, the key observation in [DiPerna and Lions, 1989] is that, to obtain $L^\infty_t(L^r_x)$ bounds for the solution, it is sufficient for $\operatorname{div} \boldsymbol{b}$ to be bounded. Moreover, since we consider forward solutions, the assumption $\operatorname{div} \boldsymbol{b}^- \in L^1_t(L^\infty_x)$ is sufficient, at least formally. Variants of this arguments can be devised, in the elliptic case, where bounds on $\operatorname{div} \boldsymbol{b}$ can be dropped in favour of bounds on $|\boldsymbol{b}|$ (Section 8.2), or when Sobolev inequalities hold (Section 8.3).

8.1 General case

We formally consider the equation satisfied by some "energy" $t \mapsto \int \beta(u_t) d\mathfrak{m}$, where $\beta : \mathbb{R} \mapsto \mathbb{R}$ is a smooth function,

$$\partial_t \int \beta(u) d\mathfrak{m} = \int \beta'(u)\partial_t u \, d\mathfrak{m} = \int \mathcal{L}(\beta'(u)) u \, d\mathfrak{m},$$

where, for simplicity, we omit to write $t \in (0,T)$, and we used the chain rule for $t \mapsto \beta(u_t)$, by the fact that u is a solution to (8.1). By the definition of $\boldsymbol{a}$ in terms of $\mathcal{L}$, we have

$$\int \mathcal{L}(\beta'(u)) u \, d\mathfrak{m} = \int \left[\mathcal{L}(\beta'(u)u) - \beta'(u)\mathcal{L}(u) - 2\boldsymbol{a}(\beta'(u), u) \right] d\mathfrak{m},$$

and recalling the definition of $\operatorname{div} \mathcal{L}$ and the chain rule for $\boldsymbol{a}$,

$$\int \left[\mathcal{L}(\beta'(u)u) - \beta'(u)\mathcal{L}(u) - 2\boldsymbol{a}(\beta'(u), u) \right] d\mathfrak{m} = -\int \left[u\beta'(u) \operatorname{div} \mathcal{L} + \beta'(u)\mathcal{L}(u) + 2\beta''(u)\boldsymbol{a}(u) \right] d\mathfrak{m}.$$

Finally, we use the chain rule for diffusions, Proposition 4.23, that gives

$$\beta'(u)\mathcal{L}(u) = \mathcal{L}(\beta(u)) - \beta''(u)\boldsymbol{a}(u), \quad \mathfrak{m}\text{-a.e. in } X,$$

and we conclude that

$$\partial_t \int \beta(u) d\mathfrak{m} = -\int \left(u\beta'(u) - \beta(u) \right) \operatorname{div} \mathcal{L} \, d\mathfrak{m} - \int \beta''(u)\boldsymbol{a}(u) \, d\mathfrak{m}. \tag{8.2}$$

Unfortunately, all these steps are not rigorous, because our notion of solution to (8.1) is in duality with $\mathscr{A}$ and we do not know whether $\beta'(u)$ can be chosen as a test function.

However, depending on the choice of β, identity (8.2) entails several estimates, essentially by means of Gronwall inequality. For example, if β is convex, then it holds

$$\beta(\tilde{z}) \geq \beta(z) + (\tilde{z} - z)\beta'(z) \quad \text{for every } z,\, \tilde{z} \in \mathbb{R},$$

and, if $\beta(0) = 0$, letting $\tilde{z} = 0$, we obtain $z\beta'(z) - \beta(z) \geq 0$. Moreover, the term $\beta''(u)\boldsymbol{a}(u)$, is non-negative, so

$$\partial_t \int \beta(u) d\mathfrak{m} \leq \int \big(u\beta'(u) - \beta(u)\big) \operatorname{div} \mathcal{L}^- d\mathfrak{m} \leq \left\|\operatorname{div} \mathcal{L}^-\right\|_\infty \int \big(u\beta'(u) - \beta(u)\big)\, d\mathfrak{m}.$$

We choose $\beta(z) = |z|^r$, for $r \in (1, \infty)$, and deduce

$$\partial_t \int |u|^r \, d\mathfrak{m} \leq (r-1) \left\|\operatorname{div} \mathcal{L}^-\right\|_\infty \int |u|^r \, d\mathfrak{m},$$

which entails, by Gronwall inequality, the uniform bound

$$\sup_{t\in[0,T]} \|u_t\|_r \leq \exp\left\{ \left(1 - \frac{1}{r}\right) \left\|\operatorname{div} \mathcal{L}^-\right\|_{L^1_t(L^\infty_x)} \right\} \|\overline{u}\|_r . \tag{8.3}$$

Letting $r \to \infty$, we deduce the bound

$$\sup_{t\in[0,T]} \|u_t\|_\infty \leq \exp\left\{ \left\|\operatorname{div} \mathcal{L}^-\right\|_{L^1_t(L^\infty_x)} \right\} \|\overline{u}\|_\infty ,$$

while letting $r \to 1$ we obtain

$$\sup_{t\in[0,T]} \|u_t\|_1 \leq \|\overline{u}\|_1 .$$

These estimates, when regarded as a-priori bounds, lead to existence by compactness arguments, but they also entail uniqueness, since the equation is linear. Given two solutions u_1, u_2, with respect to initial data $\overline{u}_1$, $\overline{u}_2 \in L^r(\mathfrak{m})$, we let $u = u_1 - u_2$, which solves (8.1) with $\overline{u} = \overline{u}_1 - \overline{u}_2$, and deduce

$$\|u_1 - u_2\|_{L^\infty_t(L^r_x)} \leq \exp\left\{ \left(1 - \frac{1}{r}\right) \left\|\operatorname{div} \mathcal{L}^-\right\|_{L^1_t(L^\infty_x)} \right\} \|\overline{u}_1 - \overline{u}_2\|_r .$$

Arguing similarly, one can consider $\beta_+(z) = (z^+)^r$ and deduce

$$\sup_{t\in[0,T]} \left\|u_t^+\right\|_r \leq \exp\left\{ \left(1 - \frac{1}{r}\right) \left\|\operatorname{div} \mathcal{L}^-\right\|_{L^1_t(L^\infty_x)} \right\} \left\|\overline{u}^+\right\|_r ,$$

and correspondingly for $\beta_-(z) = (z^-)^r$. In particular, we deduce that $\overline{u} \geq 0$ entails $u_t \geq 0$, for $t \in [0, T]$ and a comparison principle.

8.2 The elliptic case

When $\boldsymbol{a}$ is elliptic, i.e. for some $\lambda > 0$, $\boldsymbol{a}(u) \geq \lambda\Gamma(u)$, by keeping the term $\beta''(u)\boldsymbol{a}(u)$ in (8.2), we deduce stronger estimates. Indeed, let again β be convex, non-negative, with $\beta(0) = 0$, and integrate (8.2) over $t \in [0, T]$, to deduce

$$\int \beta(u_T) d\mathfrak{m} - \int \beta(\overline{u}) d\mathfrak{m} \leq \int \left[u\beta'(u) - \beta(u)\right] \operatorname{div} \mathcal{L}^- d\widetilde{\mathfrak{m}} - \int \beta''(u)\boldsymbol{a}(u) d\widetilde{\mathfrak{m}},$$

where we recall the notation $\widetilde{\mathfrak{m}} := \mathscr{L}^1 \otimes \mathfrak{m}$. From this, we obtain a bound for $\beta''(u)\boldsymbol{a}(u)$, choosing e.g. $\beta(z) = z^2$, thus

$$2\int_0^T \int \boldsymbol{a}(u_t) d\mathfrak{m}\, dt \leq \int |\overline{u}|^2 d\mathfrak{m} + \int_0^T \int |u_t|^2 \operatorname{div} \mathcal{L}_t^- d\mathfrak{m}\, dt.$$

If $\operatorname{div} \mathcal{L}^- \in L^1_t(L^\infty_x)$, by (8.3) above we deduce $\boldsymbol{a}(u) \in L^2_t(L^1_x)$, with

$$\int_0^T \int \boldsymbol{a}(u_t) d\mathfrak{m}\, dt \leq \frac{1}{2}\left[1 + \exp\left\{\left\|\operatorname{div} \mathcal{L}^-\right\|_{L^1_t(L^\infty_x)}\right\} \left\|\operatorname{div} \mathcal{L}^-\right\|_{L^1_t(L^\infty_x)}\right] \int |\overline{u}|^2 d\mathfrak{m},$$

which leads to $u \in L^2_t(\mathbb{V})$, in the elliptic case.

The quantity $\beta''(u)\boldsymbol{a}(u)$ can be exploited also in a different way, providing a PDE counterpart of Girsanov theorem. Precisely, we "perturb" a given diffusion operator $\mathcal{L}$ by adding a derivation $\boldsymbol{b}$ and consider the FPE associated to $\mathcal{L} + \boldsymbol{b}$,

$$\partial_t u_t = (\mathcal{L} + \boldsymbol{b})^* u_t, \quad \text{in } (0, T) \times X. \tag{8.4}$$

Notice that the associated 2-tensor $\boldsymbol{a}$ remains unchanged. In place of (8.2), we obtain

$$\partial_t \int \beta(u) d\mathfrak{m} \leq \int \big(u\beta'(u) - \beta(u)\big) \operatorname{div} \mathcal{L}^- d\mathfrak{m} + \int \beta''(u) \left[du(\boldsymbol{b})u - \boldsymbol{a}(u)\right] d\mathfrak{m}.$$

When $\boldsymbol{a}$ is λ-elliptic and $|\boldsymbol{b}| \in L^1_t(L^\infty_x)$, the second integral in the right hand side above can be bounded from above,

$$\begin{aligned}
\int \beta''(u) \left[du(\boldsymbol{b})u - \boldsymbol{a}(u)\right] d\mathfrak{m} &\leq \int \beta''(u) \left[|u|\, |\boldsymbol{b}| \sqrt{\Gamma(u)} - \lambda\Gamma(u)\right] d\mathfrak{m} \\
&\leq \int \beta''(u) \left[\frac{|u|^2 \|\boldsymbol{b}\|_\infty^2}{2\lambda} - \frac{\lambda}{2}\Gamma(u)\right] d\mathfrak{m} \\
&\leq \frac{\|\boldsymbol{b}\|_\infty^2}{2\lambda} \int \beta''(u) |u|^2 d\mathfrak{m} - \frac{\lambda}{2} \int \beta''(u)\Gamma(u) d\mathfrak{m}
\end{aligned}$$

Letting once again $\beta(z) = |z|^r$, for $r \in (1, \infty)$, and assuming that $\overline{u} \in L^r(\mathfrak{m})$ Gronwall inequality leads to $u \in L^\infty_t(L^r_x)$, since we have

$$\partial_t \int |u|^r d\mathfrak{m} \leq \left[r(r-1) \|\boldsymbol{b}\|_\infty^2 / (2\lambda) + (r-1) \left\|\operatorname{div} \mathcal{L}^-\right\|_\infty\right] \int |u|^r d\mathfrak{m} - \frac{\lambda r(r-1)}{2} \int \Gamma(u) d\mathfrak{m}.$$

We also obtain $u \in L^2_t(\mathbb{V})$ if $\overline{u} \in L^2(\mathfrak{m})$. Let us remark again that we impose no regularity assumption on $\boldsymbol{b}$, only bounds on $|\boldsymbol{b}|$.

8.3 Energy estimates and Sobolev inequalities

In this section, we show that the validity of Sobolev inequalities allows for improving the bounds studied above, at least in the elliptic case. Before addressing their derivation, we remark the following connection between energy estimates and lower bounds on spectra of suitable Schrödinger operators, with potential energies related to $\operatorname{div} \mathcal{L}$. Letting $\beta(z) = |z|^r$, for $r \in (1, \infty)$, in (8.2), we obtain

$$\partial_t \int |u|^r \, d\mathfrak{m} = -(r-1) \int \left[|u|^r \operatorname{div} \mathcal{L} + r \, |u|^{r-2} \, \boldsymbol{a}(u) \right] d\mathfrak{m}. \tag{8.5}$$

The chain rule for $\boldsymbol{a}$ entails

$$r \, |u|^{r-2} \, \boldsymbol{a}(u) = \frac{4}{r} \boldsymbol{a}(|u|^{r/2}),$$

thus

$$\partial_t \int |u|^r \, d\mathfrak{m} = -(r-1) \mathcal{E}_r \left(|u|^{r/2} \right),$$

where we introduce be the quadratic form

$$f \mapsto \mathcal{E}_r(f) := \int \left[\operatorname{div} \mathcal{L} \, |f|^2 + \frac{4}{r} \boldsymbol{a}(f) \right] d\mathfrak{m},$$

which corresponds to a Schrödinger operator with kinetic energy $\frac{4}{r} \int \boldsymbol{a}(f)$ and potential energy $\operatorname{div} \mathcal{L}$. Notice that the contribution of the kinetic energy is infinitesimal, as $r \to \infty$.

If the form $\mathcal{E}_r$ is bounded from below, i.e. for some $c = (c_t)_t \in L^1(0, T)$ it holds

$$\int \left[\operatorname{div} \mathcal{L} \, |f|^2 + \frac{4}{r} \boldsymbol{a}(f) \right] d\mathfrak{m} \geq c(t) \int |f|^2 \, d\mathfrak{m}, \quad \text{for every } f \in L^2_x, \tag{8.6}$$

then, we obtain the bound from above

$$\partial_t \int |u|^r \, d\mathfrak{m} \leq -(r-1) c \int |u|^r \, d\mathfrak{m},$$

and by Gronwall inequality, we conclude

$$\sup_{t \in [0,T]} \|u_t\|_r \leq \exp \left\{ \left(1 - \frac{1}{r} \right) \|c\|_{L^1_t} \right\} \|\overline{u}\|_r \, .$$

Notice also that, if it holds

$$\int \left[\operatorname{div} \mathcal{L} \, |f|^2 + \frac{4}{r} \boldsymbol{a}(f) \right] d\mathfrak{m} \geq a \int \Gamma(f) d\mathfrak{m} + c_t \int |f|^2 \, d\mathfrak{m}, \text{for every } f \in L^2(\mathfrak{m}),$$

for some constant $a > 0$, we obtain $u \in L^2_t(\mathbb{V})$, arguing as in the first part of the previous section.

Being $\boldsymbol{a}(f)$ non-negative, we may always let $c_t := \|\operatorname{div} \mathcal{L}^-\|_\infty$. However, if some Sobolev inequality holds, it is well known that (8.6) is true for a wide class of potentials, possibly unbounded from below, see e.g. [Kato, 1995]. Sobolev inequalities can be formulated as follows, see e.g. the monograph [Varopoulos et al., 1992].

Definition 8.1 (d-dimensional Sobolev inequality)**.** *Let $d \geq 2$. We say that the d-dimensional Sobolev inequality holds if there exists some constant $c_d \geq 0$, depending on d only, such that*

$$\|f\|_{2d/(d-2)} \leq c_d \|f\|_{\mathbb{V}}, \quad \textit{for every } f \in \mathbb{V}.$$

Analogous inequalities can be stated, with exponents $r \in [1, \infty]$ in place of 2, replacing $\mathbb{V}$ with $\mathbb{V}^r$ and also for $d \leq 2$, using the $L^\infty(\mathfrak{m})$ norm in the left hand side.

Here, we fix some $d \geq 2$, assume that the d-dimensional Sobolev inequality holds, in the form above and investigate a lower bound for the energy

$$f \mapsto \int \Gamma(f) + V |f|^2 \, d\mathfrak{m},$$

assuming that $V^- \in L^\infty(\mathfrak{m}) + L^{d/2}(\mathfrak{m})$. We claim that there exists some constant $c \in \mathbb{R}$ such that

$$\int \Gamma(f) + V |f|^2 \, d\mathfrak{m} \geq c \int |f|^2 \, d\mathfrak{m}, \quad \text{for every } f \in L^2(\mathfrak{m}).$$

Indeed, decomposing

$$V = V\chi_{\{V \leq -\alpha\}} + V\chi_{\{V \geq -\alpha\}}, \qquad \text{for } \alpha \geq 0 \text{ large enough},$$

without loss of generality, we assume that $\|V\|_{d/2} < \varepsilon$, for any but fixed $\varepsilon > 0$: here, $\varepsilon = c_d^{-1}$ is sufficient. Then, applying Hölder inequality, with exponent $d/2$, and the d-dimensional Sobolev inequality, we obtain

$$\int |V| \, |f|^2 \, d\mathfrak{m} \leq \|V\|_{d/2} \|f\|^2_{\frac{2d}{d-2}} \leq \varepsilon c_d \|f\|^2_{\mathbb{V}}$$

Letting $\varepsilon = c_d^{-1}$, we conclude that

$$-\int |V| \, |f|^2 \, d\mathfrak{m} \geq -\int \Gamma(f) - \int |f|^2 \, d\mathfrak{m}.$$

Letting $\varepsilon = (2c_n)^{-1}$ instead, we maintain some ellipticity, and prove

$$\int \Gamma(f) + V |f|^2 \, d\mathfrak{m} \geq \frac{1}{2} \int \Gamma(f) d\mathfrak{m} - C \int |f|^2 \, d\mathfrak{m},$$

where C is some constant depending only on c_d.

The validity of the d-dimensional Sobolev inequality allows also for improving the Girsanov-type argument, in the previous section. Indeed, consider a solution u to (8.4), where $\mathcal{L}$ in a λ-elliptic diffusion operator and $\boldsymbol{b}$ is a derivation. The chain rule gives

$$r |u|^{r-2} u \, du(\boldsymbol{b}) = 2 |u|^{r/2} \, d |u|^{r/2} (\boldsymbol{b}),$$

and in this case the problem of estimates in $L^\infty_t(L^r_x)$ is reduced to lower bounds for the energy

$$\int \operatorname{div} \mathcal{L} \, |f|^2 + \frac{4}{r} \boldsymbol{a}(f) - 2 |f| \, |df(\boldsymbol{b})| \, d\mathfrak{m}.$$

We claim that we can weaken the condition $|\boldsymbol{b}| \in L^1_t(L^\infty_x)$ to $|\boldsymbol{b}| \in L^1_t(L^\infty_x + L^d_x)$. Indeed, Hölder inequality gives

$$\int |f| \, |df(\boldsymbol{b})| \, d\mathfrak{m} \leq \int |f| \, |\boldsymbol{b}| \sqrt{\Gamma(f)} d\mathfrak{m} \leq \|\boldsymbol{b}\|_d \, \|f\|_{\frac{2d}{d-2}} \, \|f\|_{\mathbb{V}} \leq c_d \, \||\boldsymbol{b}|\|_d \, \|f\|^2_{\mathbb{V}} \, .$$

By decomposing

$$\boldsymbol{b} = \chi_{\{|\boldsymbol{b}| \leq \alpha\}} \boldsymbol{b} + \chi_{\{|\boldsymbol{b}| > \alpha\}} \boldsymbol{b}, \quad \text{for } \alpha \text{ large enough,}$$

we may assume with no loss of generality $\|\boldsymbol{b}\|_d < \varepsilon$ for $\varepsilon > 0$ small enough, to be chosen in terms of c_d, λ and r. From this we deduce the lower bound

$$\int \operatorname{div} \mathcal{L} \, |f|^2 + \frac{4\lambda}{r} \Gamma(f) - 2 \, |f| \, |df(\boldsymbol{b})| \, d\mathfrak{m} \geq c \int |f|^2 \, d\mathfrak{m},$$

for some constant $c \in \mathbb{R}$.

Summing up, if we assume the d-dimensional Sobolev inequality holds, and let $\mathcal{L}$ be an elliptic operator, such that $\operatorname{div} \mathcal{L}^- \in L^1_t(L^\infty_x + L^{d/2}_x)$, $\boldsymbol{b} \in L^1_t(L^\infty_x + L^d_x)$ be a derivation, then energy estimates for any solution u to (8.4) formally hold, leading to bounds in $L^\infty_t(L^r_x)$, as well as $L^2_t(\mathbb{V})$.

We investigate yet another consequence of the d-dimensional Sobolev inequality, namely ultra-contractivity for the associated semigroup P, i.e. the inequality

$$\|\mathsf{P}_t f\|_\infty \leq \tilde{c}_d t^{-d/2} \, \|f\|_1 \, , \text{for every } t \in (0, \infty), \; f \in L^1(\mathfrak{m}),$$

for some constant $\tilde{c}_d \geq 0$, independent of f, t and related to c_d only.

Following an argument from [Varopoulos et al., 1992] for which credit is given to Nash, we prove analogous results where, in place of $\mathsf{P}_t f$, we let u_t be a solution to (8.1).

We begin with the identity (8.5), for $r = 2$,

$$\partial_t \int |u_t|^2 \, d\mathfrak{m} = -\mathcal{E}_2(u_t),$$

we assume that, for some constant $a > 0$ and $c \in L^1(0, T)$, we have a lower bound

$$\mathcal{E}_2(f) \geq a \int \Gamma(f) d\mathfrak{m} + c \int |f|^2 \, d\mathfrak{m}, \quad \text{for every } f \in \mathbb{V}.$$

Then, we use the d-dimensional Sobolev inequality and, up to taking $c - a$ in place of c, we obtain the inequality

$$\mathcal{E}_2(f) \geq a \, \|f\|^2_{\mathbb{V}} + c \, \|f\|^2_2 \geq a c_d \, \|f\|^2_{2d/d-2} + c \, \|f\|^2_2 \, .$$

Letting $f = u_t$, we find therefore

$$\partial_t \, \|u\|^2_2 \leq -a c_d \, \|u\|^2_{2d/(d-2)} - c \, \|u\|^2_2 \, . \tag{8.7}$$

For any $p \leq q \leq r \in [1, \infty]$, Hölder inequality gives

$$\|f\|_q \leq \|f\|_p^{\frac{p(r-q)}{r-p}} \, \|f\|_r^{\frac{r(q-p)}{r-p}} \, , \quad \text{for every } f \in L^r \cap L^p(\mathfrak{m}).$$

In this case, we choose $p = 1$, $q = 2$ and $r = 2d/(d-2)$, to deduce

$$\|u\|_2 \leq \|u\|_1^{\frac{4}{d+2}} \|u\|_{2d/(d-2)}^{\frac{2d}{d+2}}.$$

Without loss of generality, assuming e.g. some bound on $\operatorname{div} \mathcal{L}^-$, the quantity $\sup_{t\in[0,T]} \|u_t\|_1$ is bounded in terms of $\|\overline{u}\|_1$, so that for $t \in [0,T]$, and some constant C depending on $\mathcal{L}$, and d only, it holds

$$\|u_t\|_{2d/(d-2)}^2 \leq C \|\overline{u}\|_1^{-4/d} \|u_t\|_2^{2+4/d}.$$

Substituting this inequality in (8.7), we obtain

$$\partial_t \|u\|_2^2 \leq -C \|\overline{u}\|_1^{-4/d} \|u\|_2^{2+4/d} - c \|u\|_2^2,$$

thus $v(t) := \exp\left\{\int_0^t c_s ds\right\} \|u_t\|_2^2$ satisfies the inequality

$$\partial_t v \leq -C \exp\left\{4/d \|c\|_{L^1_t}\right\} \|\overline{u}\|_1^{-4/d} v^{1+2/d},$$

and by comparison with $t^{-d/2}$, we conclude that the ultracontractive bound

$$\|u_t\|_2^2 \leq C t^{-d/2} \|\overline{u}\|_1$$

holds true.

Let us conclude this chapter by remarking that similar arguments can be devised in presence of a logarithmic Sobolev inequality.

Definition 8.2 (log-Sobolev inequality)**.** *We say that the* log-*Sobolev inequality holds if there exists some constant* $c \geq 0$ *such that*

$$\|f\|_{L^2 \log L(\mathfrak{m})} \leq c \|f\|_{\mathbb{V}}, \quad \text{for every } f \in \mathbb{V}.$$

The only technical difficulty that appears is to deal with Orlicz spaces such as $L^2 \log L(\mathfrak{m})$, defined as the set of Borel functions f such that

$$\int f^2 \log(|f|^2 + 1) d\mathfrak{m} < \infty,$$

endowed with the associated Luxemburg norm,

$$\|f\|_{L^2 \log L(\mathfrak{m})} := \sup\left\{\lambda > 0 \,|\, \int (\lambda f)^2 \log(|\lambda f|^2 + 1) d\mathfrak{m} \leq 1\right\}.$$

From the validity of the log-Sobolev inequality, lower bounds on spectra of Schrödinger operators whose potential have negative parts exponentially integrable can be deduced, see e.g. [Shigekawa, 2007], and similarly one can provide bounds for additive perturbations of elliptic diffusion operators by means of derivations such that $|\boldsymbol{b}|^2$ is exponentially integrable.

Chapter 9

Existence of solutions to FPE's

In this section, we focus on existence results for solutions to the FPE

$$\partial_t u_t = \mathcal{L}_t^* u_t, \quad \text{in } (0,T) \times X, \tag{9.1}$$

with prescribed initial datum $u_0 = \overline{u}$, when $\mathcal{L}$ is in divergence form, see Definition 4.24. In view of the validity of the superposition principle, see Chapter 7, existence for FPE's settles also the problem of existence for martingale problems. Recall that our standing assumptions on X are (3.1) and on $\mathcal{A}$ are (5.2).

Here, we follow a strategy similar to [Ambrosio and Trevisan, 2014, §4], but indeed classical: it consists in dealing first with the elliptic case, in Section 9.1, where existence follows by Hilbert space techniques (Lions-Lax-Milgram theorem) and we actually provide solutions in $L^2_t(\mathbb{V})$. The general case follows then compactness, where the extra regularity $u \in L^2_t(\mathbb{V})$ is exploited to rigorously settle the energy estimates described in the previous chapter. Indeed, to reduce to the elliptic case, we add a viscosity, i.e. we solve

$$\partial_t u_t = (\mathcal{L}_t + \sigma\Delta)^* u_t \quad \text{in } (0,T) \times X,$$

for $\sigma > 0$, and then we let $\sigma \downarrow 0$. Let us remark that, in order to prove the mass-conservation property of solutions to the continuity equation we assume the existence of $(f_n)_n \subset \mathscr{A}$ satisfying

$$0 \le f_n \le 1, \quad f_n \uparrow 1 \ \tilde{\mathfrak{m}}\text{-a.e. in } \tilde{X}, \quad \mathcal{L} f_n \to 0 \text{ weakly in } L^1_t(L^{r'}_x). \tag{9.2}$$

9.1 The elliptic case

In this section, we prove deal with the following result.

Theorem 9.1 (existence, elliptic case.)**.** *Let $\mathscr{A}$ be dense in $W^{1,2}_t(L^2_x) \cap L^2_t(\mathbb{V})$, let $\mathcal{L}$ be a diffusion operator in divergence form, with*

$$|\boldsymbol{a}|, |\boldsymbol{b}| \in L^\infty_t(L^\infty_x), \quad \operatorname{div} \mathcal{L}^- \in L^1_t(L^\infty_x), \quad \boldsymbol{a} \textit{ elliptic.}$$

Then, there exists a (unique) solution $u \in C([0,T]; L^2(\mathfrak{m})) \cap L^2_t(\mathbb{V})$ to (9.1), with $u_0 = \overline{u}$, in the following weak sense:

$$\int [-\partial_t f - df(\boldsymbol{b})]\, u + \boldsymbol{a}(f,u) d\widetilde{\mathfrak{m}} = \int f_0 \overline{u} d\mathfrak{m} \qquad \forall f \in \mathscr{A}_T \tag{9.3}$$

where we let $\mathscr{A}_T$ be the subset of functions $f \in \mathscr{A}$ with $f_T = 0$. Furthermore, if $\overline{u} \in L^2 \cap L^r(\mathfrak{m})$, then $u \in L^\infty_t(L^2_x \cap L^r_x)$ for $r \geq 0$ and if $\overline{u} \geq 0$, then $u_t \geq 0$ for $\mathscr{L}^1$-a.e. $t \in (0,T)$.

We remark that the formulation (9.3) differs from the notion of weak solution in $L^\infty_t(L^r_x)$ to (9.1) and their precise link is discussed in Lemma 6.5. Notice that the advantage of the formulation above is that no integrability assumption on $\mathcal{L}f$, for $f \in \mathscr{A}$, has to be imposed, but we pay the price of extra regularity for u.

Proof. With no loss of generality, we assume $\sigma \in (0,1]$. Existence in $L^2(I;\mathbb{V})$ is a consequence of J.-L. Lions' extension of Lax-Milgram Theorem, whose statement is recalled below, for which we refer to [Showalter, 1997, Theorem II.2.1, Corollary III.2.3].

Together with existence, we obtain the a priori estimate:

$$\left\| e^{-\lambda t} u \right\|_{L^2_t(\mathbb{V})} \leq \frac{2\left\|\overline{u}\right\|_2}{\sigma} \quad \text{with} \quad \lambda := \frac{\sigma}{2} + \frac{2}{\sigma}\left\|\boldsymbol{b}\right\|^2_{L^\infty_t(L^\infty_x)}. \tag{9.4}$$

To this aim, we make a preliminary change of variables $h_t := e^{-\lambda t} u_t$ and we study the following weak formulation,

$$\int \left[-\partial_t f + \lambda f - df(\boldsymbol{b})\right] h + \boldsymbol{a}(f,h) d\widetilde{\mathfrak{m}} = \int f_0 u d\mathfrak{m} \qquad \forall f \in \mathscr{A}_T. \tag{9.5}$$

Existence is a consequence of the next theorem, applied with $H = L^2_t(\mathbb{V})$, $V = \mathscr{A}_T$, endowed with the norm

$$\|f\|^2_V = \|f\|^2_{L^2_t(\mathbb{V})} + \|f_0\|^2_2, \quad \text{for } f \in \mathscr{A}_T, \tag{9.6}$$

and

$$B(f,h) = \int \left[-\partial_t f + \lambda f - df(\boldsymbol{b})\right] h + \boldsymbol{a}(f,h) d\widetilde{\mathfrak{m}}, \qquad \ell(f) = \int f_0 \overline{u} d\mathfrak{m}.$$

Theorem 9.2 (Lions)**.** *Let V, H be respectively a normed and a Hilbert space, with V continuously embedded in H, $\|v\|_H \leq \|v\|_V$ for all $v \in V$, and let $B : V \times H \to \mathbb{R}$ be bilinear, with $B(v,\cdot)$ continuous for all $v \in V$. If B is coercive, namely there exists $c > 0$ satisfying $B(v,v) \geq c\|v\|^2_V$ for all $v \in V$, then for all $\ell \in V'$ there exists $h \in H$ such that $B(\cdot,h) = \ell$ and*

$$\|h\|_H \leq \frac{\|\ell\|_{V'}}{c}. \tag{9.7}$$

Let us start by proving continuity, thus let $f \in \mathscr{A}_T$. The linear functional $h \mapsto B(f,h)$ is $L^2_t(\mathbb{V})$-continuous, since we estimate $|B(f,h)|$ from above with

$$\|h\|_{L^2(I;\mathbb{V})}\left[\|\partial_t f\|_{L^2_t(L^2_x)} + \lambda\|f\|_{L^2_t(L^2_x)} + (\|\boldsymbol{b}\|_{L^\infty_t(L^\infty_x)} + \|\boldsymbol{a}\|_{L^\infty_t(L^\infty_x)}\|\sqrt{\Gamma(f)}\|_{L^2_t(L^2_x)}\right].$$

The functional ℓ satisfies $\|\ell\|_{V'} \leq \|\overline{u}\|_2$, immediately from the definition of $\|\cdot\|_V$ in (9.6).

To conclude the verification of the assumptions of Theorem 9.2, we show coercivity (here the change of variables we did and the choice of λ play a role). It holds

$$\begin{aligned}\int \left[\lambda f - df(\boldsymbol{b})\right] f d\widetilde{\mathfrak{m}} &\geq \lambda\|f\|^2_{L^2_t(L^2_x)} - \|\boldsymbol{b}\|_{L^\infty_t(L^\infty_x)}\|f\|_{L^2_t(L^2_x)}\|\sqrt{\Gamma(f)}\|_{L^2_t(L^2_x)} \\ &\geq \frac{\sigma}{2}(\|f\|^2_{L^2_t(L^2_x)} - \|\sqrt{\Gamma(f)}\|^2_{L^2_t(L^2_x)}).\end{aligned} \tag{9.8}$$

Since $f \in \mathscr{A} \subseteq W^{1,2}_t(L^2_x)$, the chain rule $\partial_t f^2 = 2f\partial_t f$ holds and we integrate by parts

$$2\int f\partial_t f d\widetilde{\mathfrak{m}} = \int f_T^2 d\mathfrak{m} - \int f_0^2 d\mathfrak{m} = -\int f_0^2 d\mathfrak{m}, \quad \text{using } f_T = 0.$$

Hence, inequality (9.8) entails that

$$\int \left[-\partial_t f + \lambda f - df(\boldsymbol{b})\right] f + \boldsymbol{a}(f) d\widetilde{\mathfrak{m}} \geq \frac{1}{2}\int f_0^2 d\mathfrak{m} + \frac{\sigma}{2}\|f\|^2_{L^2_t(L^2_x)} + \frac{\sigma}{2}\|\sqrt{\Gamma(f)}\|^2_{L^2_t(L^2_x)}.$$

Since $\sigma \leq 1$, it follows from these two inequalities that

$$B(f,f) \geq \sigma\|f\|^2_V. \tag{9.9}$$

Finally, (9.4) follows at once from (9.7) and (9.9), taking into account that $\|\ell\|_{V'} \leq \|\overline{u}\|_2$. Existence for the formulation (9.5) is settled.

We prove now that the solution to (9.5) belonging in $L^2_t(\mathbb{V})$ is actually unique and enjoys the properties stated in the theorem. By the assumptions of $\mathscr{A}$, we may extend the formulation to $f = W^{1,2}_t(L^2_x) \cap L^2_t(\mathbb{V})$, thus letting $f \in C^{1,2}((0,T);\mathbb{V})$, the equation itself entails that $\partial_t u$ belongs to $L^2_t(\mathbb{V}^*)$, thus $u \in L^2_t(\mathbb{V}) \cap W^{1,2}_t(\mathbb{V}^*)$. As a consequence of Proposition 5.7, u admits a representative in $\in C([0,T];L^2_t)$ and from the formulation (9.5) we immediately deduce that $u_0 = \overline{u}$, as elements in $L^2(\mathfrak{m})$.

We also prove that some energy estimates holds. In particular, for $r \in [1,\infty]$, if $\overline{u} \in L^2 \cap L^r(\mathfrak{m})$, then

$$\sup_{t\in[0,T]} \|u_t\|_r \leq \exp\left\{\left(1 - \frac{1}{r}\right)\left\|\operatorname{div}\mathcal{L}^-\right\|_{L^1_t(L^\infty_x)}\right\}\|\overline{u}\|_r \tag{9.10}$$

and, if $\overline{u} \geq 0$, then $u_t \geq 0$, for every $t \in [0,T]$. As a consequence, we obtain uniqueness for (9.5) in $L^2_t(\mathbb{V})$.

To show the validity of (9.10), let $r \in (1,\infty)$, as the inequality for the endpoints follow by suitable limits. To make rigorous the arguments in the previous chapter, we rely on Proposition 5.8. Notice however that $\beta(z) = |z|^r$ is not allowed, since $|\beta'(z)| / |z|$ and $|\beta''(z)|$ could be unbounded. However, one can build an approximating sequence of convex functions β_n, such that the two conditions are satisfied, and it holds $\beta_n'(z)z - \beta(z) \leq (r-1)\beta_n(z)$, for every $z \in \mathbb{R}$. Precisely, if $r \geq 2$, we let

$$\beta_n(z) := \begin{cases} \beta(-n) + (r-1)(-n)^{r-1}(z+n) & \text{if } z < -n; \\ \beta(z) & \text{if } -n \leq z \leq n; \\ \beta(n) + (r-1)(n)^{r-1}(z-n) & \text{if } z > n, \end{cases}$$

while if $r < 2$, we let

$$\beta_n(z) := \begin{cases} \dfrac{(z^+)^2}{2\epsilon^{2-r}} & \text{if } z \leq \epsilon; \\ \\ (z^+)^r - \dfrac{\epsilon^r}{2} & \text{if } z \geq \epsilon. \end{cases}$$

for $\varepsilon = 1/n$. To be rigorous, one should also consider a slightly smoothed version, since $\beta''(z)$ does not exist at $z = n$ in the first case, $z = \varepsilon$ in the second case.

Therefore, we deduce that the curve $[0,T] \ni t \mapsto \int \beta_n(u_t) d\mathfrak{m}$ is absolutely continuous, with

$$\frac{d}{dt}\int \beta(u_t) d\mathfrak{m} = \left\langle \beta'(u_t), \partial_t u_t \right\rangle_{\mathbb{V}}.$$

Arguing as in (8.2), we conclude that

$$\frac{d}{dt}\int \beta(u_t) d\mathfrak{m} \leq (r-1)\left\|\operatorname{div}\mathcal{L}_t^-\right\|_\infty \int \beta_n(u_t) d\mathfrak{m}, \quad \text{a.e. } t \in (0,T),$$

and by Gronwall's lemma we deduce an approximated energy inequality (9.10). Letting $n \to \infty$, we conclude by Fatou's lemma that (9.10) holds.

Finally, letting $\beta(z) = z^2$ and arguing as in Section 8.2, we deduce the bound

$$\|u\|_{L^2_t(\mathbb{V})} \leq \frac{1}{\sigma} C(\left\|\operatorname{div}\mathcal{L}^-\right\|_{L^1_t(L^\infty_x)}) \left\|\overline{u}\right\|_2, \tag{9.11}$$

where the notation remarks that C is some quantity depending on $\|\operatorname{div}\mathcal{L}^-\|_{L^1_t(L^\infty_x)}$ only. □

Remark 9.3 (uniqueness in $L^2_t(\mathbb{V})$)**.** The argument above actually entails existence and uniqueness in $L^2_t(\mathbb{V})$ for solutions to

$$\partial_t u_t = \mathcal{L}^* u_t + \ell, \quad \text{in } (0,T)\times X,$$

where ℓ is any continuous functional on $W^{1,2}_t(L^2_x) \cap L^2_t(\mathbb{V})$, where the notion of solution is given by the following formulation:

$$\int \left[-\partial_t f - df(\boldsymbol{b})\right] u + \boldsymbol{a}(f,u) d\widetilde{\mathfrak{m}} = \ell(f) \quad \text{for every } f \in \mathscr{A}_T.$$

holds. Uniqueness is a consequence of the fact that the difference solves the equation with $\ell = 0$ and $\overline{u} = 0$. Notice that, when $\boldsymbol{a}$ is not elliptic, e.g., in the deterministic case, only existence in this class becomes an issue. ■

9.2 General case

We prove the following existence theorem, as an application of the vanishing viscosity strategy sketched at the beginning of the chapter.

Theorem 9.4 (existence, general case)**.** *Let $r \in (1,\infty]$, $\mathscr{A}$ be dense in $W^{1,2}_t(L^2(\mathfrak{m})) \cap L^2_t(\mathbb{V})$, $\mathcal{L}$ be a diffusion operator in divergence form, with*

$$|\boldsymbol{a}|, |\boldsymbol{b}| \in L^\infty_t(L^\infty_x), \quad \operatorname{div}\mathcal{L}^- \in L^1_t(L^\infty_x),$$

and $\mathcal{L} : \mathscr{A} \to L^1_t(L^r_x)$. Then, for every $\overline{u} \in L^r(\mathfrak{m})$, there exists a weak solution $u \in L^\infty_t(L^r_x)$ to the FPE (9.1), whose weakly- continuous representative satisfies $u_0 = \overline{u}$.*

Furthermore, if $\overline{u} \geq 0$, such a solution satisfies $u_t \geq 0$, $\mathscr{L}^1$-a.e. $t \in (0,T)$ and, provided that (9.2) holds, if $\overline{u}\mathfrak{m}$ is a probability measure, the same holds for $u_t\mathfrak{m}$, $\mathscr{L}^1$-a.e. $t \in (0,T)$.

Proof. The proof is based on a series of approximations, based on the fact that the bound (9.10) involves only $\|\operatorname{div} \mathcal{L}^-\|_{L^1_t(L^\infty_x)}$: we aim at expressing the given FPE as a limit of a sequence of bounded elliptic FPE's so that, up to extracting a weakly-* convergent subsequence, existence is settled.

For $\sigma \in (0,1]$, we consider the weak solution $u^\sigma \in L^2_t(\mathbb{V})$ to the FPE

$$\partial_t u = (\mathcal{L} + \sigma\Delta)^* u, \quad \text{in } (0,T) \times X,$$

provided by Theorem 9.1 (as already remarked just after its statement, it is not necessary to assume $\mathscr{A} \subseteq D(\Delta)$).

Since u^σ satisfies 9.10, we deduce that it is uniformly bounded in $L^\infty_t(L^r_x)$, thus there exists some weak-* limit point $u \in L^\infty_t(L^r_x)$. To show that the limit is a weak solution as required, we have to face the technical problem that $\mathscr{A}$ is not necessarily contained in $D(\Delta)$, thus we cannot pass to the limit directly in the formulation

$$\int \left[-\partial_t f - \mathcal{L} f\right] u^\sigma - \sigma(\Delta f) u^\sigma d\widetilde{\mathfrak{m}} = \int f_0 u d\mathfrak{m} \quad \text{for every } f \in \mathscr{A}_T.$$

However, σu^σ is bounded in $L^2_t(\mathbb{V})$, by (9.11), thus it weakly converges towards 0 and we may pass to the limit as $\sigma \downarrow 0$ in the formulation

$$\int \left[-\partial_t f - \mathcal{L} f\right] u + \sigma\Gamma(f,u) d\widetilde{\mathfrak{m}} = \int f_0 u d\mathfrak{m} \qquad \forall f \in \mathscr{A}.$$

Finally, to prove the mass preservation property, it is sufficient to choose f_n in place of f above, and then let $n \to \infty$, to get that $\partial_t \int u = 0$. □

When compared with existence results in specific settings, such as Euclidean or Gaussian spaces, the assumptions above are restrictive, as one expects existence of solutions in $L^\infty_t(L^r_x)$ assuming only

$$|\boldsymbol{a}|, |\boldsymbol{b}| \in L^1_t(L^{r'}_x), \quad \text{and} \quad \operatorname{div} \mathcal{L}^- \in L^1_t(L^\infty_x).$$

Here, it is precisely the Hilbert space technique that forces the introduction of stronger assumptions than those known in particular classes of spaces: we trade some strength in the result in favour of generality. However, let us stress the fact that the crucial a priori bounds involving $\operatorname{div} \mathcal{L}^-$ only are comparable to those obtainable in specific settings: what is lacking here is an approximating procedure by means of "smooth" and bounded diffusion operators. We may state this as an existence criterion, as follows.

Proposition 9.5 (existence by approximation)**.** *Let $r \in (1,\infty]$, let $(\mathcal{L}^n)_n$ be a sequence of diffusion operators such that $\mathcal{L}^n f \to \mathcal{L} f$ strongly in $L^1_t(L^{r'}_x)$ for $f \in \mathscr{A}$, as $n \to \infty$.*

If $(u^n)^n \subseteq L^\infty_t(L^r_x)$ is a bounded sequence, with u^n weak solutions in $L^\infty(L^r_x)$ to the FPE associated with $\mathcal{L}^n$, then any weakly- limit in $L^\infty_t(L^r_x)$ is a weak solution to the FPE associated with $\mathcal{L}^n$. Furthermore, if u^n are weakly-* continuous and $u^n_0 \to \overline{u}$, then the weakly-* continuous for u satisfies $u_0 = \overline{u}$.*

Let us finally remark that in [Ambrosio and Trevisan, 2014], a slightly different route is taken, obtaining in particular existence for solutions in $L^\infty_t(L^r_x)$ in case $r \geq 2$, $\mathcal{L} = \boldsymbol{b}$ is a derivation with $|\boldsymbol{b}| \in L^1_t(L^{r'}_x)$ and $\operatorname{div} \boldsymbol{b}^- \in L^1_t(L^\infty_x)$. By refining the arguments therein, it seems reasonable to extend those abstract existence results if $\mathcal{L}$ is in divergence form, with $|\boldsymbol{a}| \in L^1_t(L^{r'}_x)$, $|\boldsymbol{b}| \in L^1_t(L^{r'}_x)$ and $\operatorname{div} \boldsymbol{b}^- \in L^1_t(L^\infty_x)$.

Chapter 10

Uniqueness of solutions to FPE's

The formal energy estimates established in Chapter 8 seem to entail uniqueness for solutions to FPE's in $L^\infty_t(L^r_x)$, assuming $\operatorname{div}\mathcal{L}^- \in L^1_t(L^\infty_x)$ only. However, it is well-known already in the deterministic case, see [DiPerna and Lions, 1989, §IV.2], that a rigorous deduction of these inequalities requires some additional assumption on $\mathcal{L}$, although an optimal class is presently not known. We follow a by-now classical strategy, initiated by DiPerna-Lions to show uniqueness for the transport equation, assuming Sobolev regularity for the driving vector field: it consists in the study of the equations satisfied by suitable approximations of a given solution and a careful analysis of the error terms appearing, by means of the so-called commutator lemmas.

Before we address our uniqueness results, we give a more detailed description of the scheme that we follow, in Section 10.1, highlighting the crucial role of commutators. In Section 10.2, we establish, in different situations, useful commutator lemmas that we employ in Section 10.3 to deduce uniqueness for FPE's.

10.1 The smoothing scheme

In order to establish uniqueness, our aim is to rigorously establish the deductions that lead to the identity

$$\partial_t \int \beta(u) d\mathfrak{m} = -\int \big(u\beta'(u) - \beta(u)\big) \operatorname{div}\mathcal{L}\, d\mathfrak{m} - \int \beta''(u)\boldsymbol{a}(u)\, d\mathfrak{m}. \tag{10.1}$$

which, as already noticed in Chapter 8, is not precise: the first identity leading to (10.1) is

$$\partial_t \int \beta(u) d\mathfrak{m} = \int \beta'(u_t)\partial_t u d\mathfrak{m},$$

and uses the chain rule for the derivative of $\beta(u_t)$ while, in the second identity,

$$\int \beta'(u)\partial_t u d\mathfrak{m} = \int \mathcal{L}(\beta'(u)) u d\mathfrak{m},$$

we should at least ensure that $\mathcal{L}(\beta'(u))$ is well-defined. Recall that the diffusion operator $\mathcal{L}$ is initially defined on $\mathscr{A}$ and solutions to the Fokker-Planck equation are understood in duality with $\mathscr{A}$: under suitable assumptions, as in Remark 4.9 we may extend by continuity $\mathcal{L}$ to larger spaces such as $L^\infty_t(\mathbb{V}^s)$, $L^\infty_t(D^s(\Delta))$ or their intersections, but still there is no reason for $\beta'(u_t)$ to be sufficiently smooth, at least in the non-elliptic case.

The main problem therefore involves the regularity of u, both with respect to $t \in (0,T)$, in order to use the chain rule, and with respect to $x \in X$. The smoothing scheme that we follow consists in introducing, for $\alpha \in (0,1)$, some linear operator R_α, acting on functions defined in $(0,T) \times X$ (symmetric with respect to $\widetilde{\mathfrak{m}}$) such that, if we define $u^\alpha := \mathsf{R}_\alpha u$, then u^α is sufficiently regular so that (8.2) should become rigorous. Of course, the disadvantage is now that u^α is not a solution to the original FPE (8.1). However, we expect for u^α to solve an approximate FPE, namely

$$\partial_t u^\alpha = \mathcal{L}^* u^\alpha + w^\alpha, \tag{10.2}$$

where we let

$$w^\alpha := [(\partial_t + \mathcal{L})^*, \mathsf{R}_\alpha] u = (\partial_t + \mathcal{L})^* \mathsf{R}_\alpha u - \mathsf{R}_\alpha (\partial_t + \mathcal{L})^* u$$

be the formal commutator between (the dual of) $\partial_t + \mathcal{L}$ and R_α. This identity is crucial for our deductions. Indeed, in many cases, we prove that, for $\alpha > 0$, the right hand side in (10.2) is actually a function, deducing that $t \mapsto u^\alpha_t$ is a Sobolev curve. At this point, Gronwall inequality rigorously applies, entailing approximate versions of the energy estimates, which contain an the error term involving w^α.

The key point is then to study general properties for the commutator, in particular its convergence to 0, as $\alpha \downarrow 0$, assuming also that $u^\alpha \to u$, in a sense to be made precise. To ensure that the terms involving the commutator give no contribution in the limit, it turns out that strong convergence in some Lebesgue space is sufficient, and in many cases we are able to prove this, as a consequence of suitable assumptions on R and $\mathcal{L}$.

So far, the strategy described is the same as in DiPerna-Lions original approach. Our novel contribution, initiated in [Ambrosio and Trevisan, 2014], consists in choosing R to be a Markov semigroup on functions on $(0,T) \times X$, argue by duality and Bakry-Emméry interpolation, writing

$$\mathsf{R}_\alpha(\partial_t + \mathcal{L}) f - (\partial_t + \mathcal{L})\mathsf{R}_\alpha f = \int_0^\alpha \frac{d}{d\sigma}\left[\mathsf{R}_\sigma(\partial_t + \mathcal{L})\mathsf{R}_{\alpha-\sigma} f\right] d\sigma = \int_0^\varepsilon \mathsf{R}_\sigma [\mathsf{D}, (\partial_t + \mathcal{L})] \mathsf{R}_{\alpha-\sigma} f d\sigma, \tag{10.3}$$

where we let D be the generator of R. Estimates on the infinitesimal commutator $[\mathsf{D}, (\partial_t + \mathcal{L})]$ can be given in terms of natural quantities reflecting the relative regularity of $\partial_t + \mathcal{L}$ with respect to D. For example, a natural choice is letting $\mathsf{R} f(t,x) := (\mathsf{P} f_t)(x)$, for $(t,x) \in (0,T) \times X$, thus $[\mathsf{P}_\alpha, \partial_t] f = 0$. The decisive computation in [Ambrosio and Trevisan, 2014] is then to link $[\Delta, \boldsymbol{b}]$ with the symmetric part of the derivative of $\boldsymbol{b}$, see also Section 10.2.2 below. However, one can perform different choices for R, allowing for genuine t-dependent semigroups, see Section 10.2.4.

10.2 Commutator estimates

This is certainly the most technical section throughout all the thesis: as introduced above, rigorous deductions of energy estimates and then uniqueness results strongly rely on the study of commutators terms.

Although the single identity (10.3) lies at the heart of our technique, we are currently not able to give a unified result, and in this section we deal with different cases of commutators, roughly at increasing technical complexity: thus, we introduce gradually various features and problems encountered.

First, it is proficient to study commutator estimates where $\partial_t + \mathcal{L}$ is replaced by some linear continuous operator A, Section 10.2.1, providing the ideal case to which we reduce in the other ones. Also, in many (but not all) situations we let $\mathsf{R} = \mathsf{P}$, and our abstract framework allows for considering the time-independent case, see at the beginning of Section 10.2.4. Then, in Section 10.2.2, we study the case of a Sobolev derivation, which is the case considered in [Ambrosio and Trevisan, 2014] and, in Section 10.2.3, we consider diffusion operators of the form $\mathcal{L} = a\Delta$, where a belongs to a second order Sobolev space: in this case, we perform an interpolation based on a second-order Taylor expansion, instead of (10.3). Finally, in Section 10.2.4, we address the commutator between ∂_t and the semigroup R associated to a bounded elliptic form $\boldsymbol{a}$ on $(0,T) \times X$: in this case, we assume Sobolev regularity for $t \mapsto \boldsymbol{a}_t$, in the spirit of [Figalli, 2008, Proof of Theorem 4.3, Step 3.3].

Throughout this section, we fix $q \in (1, \infty]$, r, $s \in (1, \infty)$, satisfying $q^{-1} + r^{-1} + s^{-1} = 1$. We avoid to deal with the endpoint cases $q = 1$, $r = s = \infty$, since we argue often by duality and density, and for some results we also use the L^p-Δ inequality, for $p \in \{r, s\}$ (which may hold true in some settings for $p = \infty$, e.g. in Chapter 11).

10.2.1 The commutator with a continuous linear map

In this section, we let

$$\mathsf{A} : L^s(\mathfrak{m}) \mapsto L^{r'}(\mathfrak{m})$$

be a linear continuous operator. For $\alpha > 0$, we introduce the commutator between the heat semigroup P_α and A,

$$L^s(\mathfrak{m}) \ni f \mapsto [\mathsf{P}_\alpha, \mathsf{A}] f := \mathsf{P}_\alpha (\mathsf{A} u) - \mathsf{A} (\mathsf{P}_\alpha u) \in L^{r'}(\mathfrak{m}).$$

Since we argue by duality, it is useful to introduce the bilinear formulation

$$L^r(\mathfrak{m}) \cap L^s(\mathfrak{m}) \ni (u, f) \mapsto \int u [\mathsf{P}_\alpha, \mathsf{A}] f d\mathfrak{m} = \int (\mathsf{P}_\alpha u) (\mathsf{A} f) - u \mathsf{A} (\mathsf{P}_\alpha f) \, d\mathfrak{m},$$

and the following infinitesimal version (with respect to α):

$$D^r(\Delta) \times D^s(\Delta) \ni (u, f) \mapsto \int u [\Delta, \mathsf{A}] f d\mathfrak{m} := \int (\Delta u) (\mathsf{A} f) - u \, \mathsf{A} \Delta f d\mathfrak{m}. \tag{10.4}$$

It is natural in this setting to regard any bound on $\int u[\Delta, \mathsf{A}] f d\mathfrak{m}$ in terms of norms on u and f that are weaker than the trivial ones, i.e. $D^r(\Delta)$ and $D^s(\Delta)$, as expression of some regularity of A, relative to Δ: to fix ideas, let us consider the following example.

Example 10.1 (Multiplication operator)**.** Let $A \in L^q(\mathfrak{m})$ and define $\mathsf{A} f := A f$, for $f \in L^s(\mathfrak{m})$. If $A \in \mathbb{V}^q$, we can integrate by parts in (10.4) and obtain, for $f \in D^s(\Delta)$, $u \in D^r(\Delta)$,

$$\int u [\Delta, \mathsf{A}] f d\mathfrak{m} = \int [\Gamma(f, Au) - \Gamma(Af, u)] \, d\mathfrak{m} = \int [\Gamma(f, A) u - f \Gamma(A, u)] \, d\mathfrak{m}.$$

In particular, it holds

$$\left| \int f [\Delta, \mathsf{A}] u d\mathfrak{m} \right| \le 2 \|A\|_{\mathbb{V}^q} \|f\|_{\mathbb{V}^s} \|u\|_{\mathbb{V}^r},$$

although this turns out to be a rather ineffective estimate for our purposes, see Remark 10.4.

The first lemma that we prove is a basic result to obtain bounds on $[\mathsf{P}_\alpha, \mathsf{A}]f$ exploiting the joint validity of the L^p-Γ inequality for $p \in \{r, s\}$, i.e., a smoothing effect of P, together with an inequality of the type

$$\left| \int v[\Delta, \mathsf{A}]g \, d\mathfrak{m} \right| \leq \|[\Delta, \mathsf{A}]\|_{r,s} \|v\|_{\mathbb{V}^r} \|g\|_{\mathbb{V}^s} \tag{10.5}$$

for some constant $\|[\Delta, \mathsf{A}]\|_{r,s}$, for sufficiently many functions v, g, i.e., some regularity of A. Let us remark that, since A is continuous, the difficulty is not to provide some bound, but a quantitative expression in terms of the infinitesimal commutator $[\Delta, A]$.

Lemma 10.2 (basic commutator inequality)**.** *Let $u \in D^r(\Delta)$, $f \in D^s(\Delta)$, let the L^p-Γ inequality hold for $p \in \{r, s\}$ and let* (10.5) *hold with $v = \mathsf{P}_t u$, $g = \mathsf{P}_\tau f$, for every t, $\tau \in (0, 1)$, where $\|[\Delta, \mathsf{A}]\|_{r,s}$ is some constant independent of t and τ.*

Then, for every $\alpha \in (0, 1)$, it holds

$$\left| \int u[\mathsf{P}_\alpha, \mathsf{A}]f d\mathfrak{m} \right| \leq c \|[\Delta, \mathsf{A}]\|_{r,s} \|u\|_{L^2 \cap L^r} \|f\|_{L^2 \cap L^s}, \tag{10.6}$$

where $c \geq 0$ depends only on the constants c_r^Γ, c_s^Γ in (3.12)*.*

Proof. We consider the curve

$$[0, \alpha] \ni \sigma \mapsto \mathsf{A}f^{\alpha-\sigma} \in L^{r'}(\mathfrak{m}),$$

where we introduced the notation $g^t := \mathsf{P}_t g$. Since A is a linear and continuous and $f \in D^s(\Delta)$, the curve belongs to $C^1([0, \alpha], L^{r'}(\mathfrak{m}))$, with derivative

$$\frac{d}{d\sigma}\mathsf{A}f^{\alpha-\sigma} = -\mathsf{A}\Delta f^{\alpha-\sigma} = -\mathsf{A}(\Delta f)^{\alpha-\sigma}, \quad \text{for } \sigma \in (0, \alpha).$$

Similarly, since $u \in D^r(\Delta)$, the curve $\sigma \mapsto u^\sigma$ is $C^1([0, \alpha], L^{s'}(\mathfrak{m}))$, thus

$$[0, \alpha] \ni \sigma \mapsto F(\sigma) := \int u^\sigma \mathsf{A}f^{\alpha-\sigma} d\mathfrak{m}$$

belongs to $C^1([0, \alpha], \mathbb{R})$. By the fundamental theorem of calculus, the identity

$$\int u[\mathsf{P}_\alpha, \mathsf{A}]f d\mathfrak{m} = F(\alpha) - F(0) = \int_0^\alpha F'(\sigma) d\sigma,$$

reduces the problem to provide bounds for $|F'(\sigma)|$, to be integrated over $\sigma \in [0, \alpha]$. Leibniz rule applies and gives

$$F'(\sigma) = \int (\Delta u^\sigma) \left(\mathsf{A}f^{\alpha-\sigma}\right) - u^\sigma \mathsf{A}\Delta f^{\alpha-\sigma} d\mathfrak{m} = \int u^\sigma [\Delta, \mathsf{A}] f^{\alpha-\sigma} d\mathfrak{m}, \quad \text{for } \sigma \in (0, \alpha). \tag{10.7}$$

Then, we argue at fixed $\sigma \in (0, \alpha)$ and use the assumption (10.5), with $v = u^{\alpha-\sigma}$, $g = f^\sigma$, obtaining the inequality

$$\left|F'(\sigma)\right| \leq \|[\Delta, \mathsf{A}]\|_{r,s} \|u^{\alpha-\sigma}\|_{\mathbb{V}^r} \|f^\sigma\|_{\mathbb{V}^s}.$$

The validity of the L^p-Γ inequality entails that $u \mapsto u^\sigma$ is continuous from $L^2 \cap L^r(\mathfrak{m})$ to $\mathbb{V}^r$, with norm smaller than $c_r/\sqrt{\sigma}$, where c_r is some constant depending on c_r^Γ only. A similar remark holds for $f \mapsto f^{\alpha-\sigma}$, thus we find

$$\left|F'(\sigma)\right| \leq \frac{c_r c_s}{\sqrt{\sigma(\alpha-\sigma)}} \left\|[\Delta, \mathsf{A}]\right\|_{r,s} \|u\|_{L^2 \cap L^r} \|f\|_{L^2 \cap L^s}.$$

To conclude, we integrate over $\sigma \in [0, \alpha]$, recalling that

$$\int_0^\alpha \frac{d\sigma}{\sqrt{\sigma(\alpha-\sigma)}} = \pi,$$

and we deduce the validity of (10.6). $\square$

Useful variants of Lemma 10.2 can be devised, an important one is based on the remark that to estimate $|F'(\sigma)|$, we are free to add and subtract some other total derivative $G'(\sigma)$, provided that the increment $|G(\alpha) - G(0)|$ is bounded. To show how this reasoning applies, we provide the following improvement of Lemma 10.2, where condition (10.5) is replaced with

$$\left|\int v[\Delta, \mathsf{A}] g d\mathfrak{m} + \int v \left[\mathsf{a}(\Delta g)\right] d\mathfrak{m}\right| \leq \left\|[\Delta, \mathsf{A}]\right\|_{r,s} \|v\|_{\mathbb{V}^r} \|g\|_{\mathbb{V}^s}, \tag{10.8}$$

where $\mathsf{a} \in \mathscr{L}(L^s(\mathfrak{m}), L^{r'}(\mathfrak{m}))$.

Lemma 10.3 (refined commutator inequality)**.** *Let $u \in D^r(\Delta)$, $f \in D^s(\Delta)$, let the L^p-Γ inequality hold for $p \in \{r, s\}$ and let* (10.8) *hold with $v = \mathsf{P}_t u$, $g = \mathsf{P}_\tau f$, for every t, $\tau \in (0,1)$, for some constant $\left\|[\Delta, \mathsf{A}]\right\|_{r,s}$ and $\mathsf{a} \in \mathscr{L}(L^s(\mathfrak{m}), L^{r'}(\mathfrak{m}))$, independent of t and τ.*

Then, for every $\alpha \in (0,1)$, it holds

$$\left|\int u[\mathsf{P}_\alpha, \mathsf{A}] f d\mathfrak{m}\right| \leq c\left[\left\|[\Delta, \mathsf{A}]\right\|_{r,s} + \|\mathsf{a}\|\right] \|u\|_{L^2 \cap L^r} \|f\|_{L^2 \cap L^s}, \tag{10.9}$$

where $c \geq 0$ is some constant depending only on c_r^Γ, c_s^Γ in (3.12)*, c_r^Δ and c_s^Δ in* (3.14)*.*

Proof. We argue by adding and subtract suitable functions $G'(\sigma)$, $H'(\sigma)$ to (10.7), in such a way that the correspondent increments $|G(\alpha) - G(0)|$, $|H(\alpha) - H(0)|$, are controlled, exploiting the validity of the L^p-Γ and the L^p-Δ inequalities.

Consider the curve

$$[0, \alpha] \ni \sigma \mapsto G(\sigma) = \int u^\alpha \mathsf{a}(f^{\alpha-\sigma}) d\mathfrak{m}.$$

By the assumption on a,

$$|G(\alpha) - G(0)| \leq 2\|\mathsf{a}\| \|u\|_r \|f\|_s,$$

thus we may add and subtract $G'(\sigma)$ in (10.7), and we are reduced to estimate the difference

$$F'(\sigma) - G'(\sigma) = \int u^\sigma [\Delta, \mathsf{A}] f^{\alpha-\sigma} d\mathfrak{m} + \int u^\alpha \mathsf{a} \Delta(f^{\alpha-\sigma}) d\mathfrak{m},$$

which is almost the left hand side in (10.8), with $v = u^\sigma$ and $g = f^{\alpha-\sigma}$. To obtain it, we add and subtract

$$H'(\sigma) = \int f^\sigma \mathsf{a}(\Delta u^{\alpha-\sigma}) d\mathfrak{m},$$

and we estimate both $|F'(\sigma)+H'(\sigma)|$ and $|G'(\sigma)-H'(\sigma)|$. Arguing as in the proof of Lemma 10.2, but using (10.8) in place of (10.5), we deduce at once that

$$\int_0^\alpha \left|F'(\sigma)+H'(\sigma)\right| d\sigma \leq c \left\|[\Delta, \mathsf{A}]\right\|_{r,s} \|u\|_{L^2\cap L^s} \|f\|_{L^2\cap L^s},$$

thus, we are left with the term

$$\left|G'(\sigma)+H'(\sigma)\right| = \left|\int (u^\alpha - u^\sigma)\mathsf{a}(\Delta f^{\alpha-\sigma}) d\mathfrak{m}\right| \leq \|\mathsf{a}\| \|u^\alpha - u^\sigma\|_r \|\Delta f^{\alpha-\sigma}\|_s.$$

We now use (3.14) and Corollary 3.5, to obtain

$$\|u^\alpha - u^\sigma\|_r \|\Delta f^{\alpha-\sigma}\|_s \leq \|u\|_r \|f\|_s \frac{c_s^\Delta}{\alpha-\sigma} \min\left\{2, c\log\left(1+\frac{\alpha-\sigma}{\sigma}\right)\right\}.$$

Finally, we integrate over $\sigma \in [0,\alpha]$,

$$\begin{aligned}\int_0^\alpha \min\left\{\frac{2}{\alpha-\sigma}, \frac{c}{\alpha-\sigma}\log\left(1+\frac{\alpha-\sigma}{\sigma}\right)\right\} d\sigma &\leq \max\{2,c\}\int_0^\alpha \min\left\{\frac{1}{\sigma}, \frac{1}{\alpha-\sigma}\right\} d\sigma \\ &= 2\log 2\max\{2,c\},\end{aligned}$$

and conclude. □

Remark 10.4 (further improvements)**.** A similar argument clearly applies also if in the left hand side of (10.8) there appears a term $\int \bar{\mathsf{a}}(\Delta u) f d\mathfrak{m}$, for some linear continuous operator $\bar{\mathsf{a}} \in \mathcal{L}(L^r, L^{s'})$. In particular, in the situation of Example 10.1, we deduce commutator inequalities for $[\mathsf{P}_\alpha, A]f$ requiring no regularity for A.

Inequalities (10.6) and (10.9) are given in terms of the norms in the intersection spaces $L^2 \cap L^r(\mathfrak{m})$ and $L^2 \cap L^s(\mathfrak{m})$. One can check that the proof holds as well for A, $\mathsf{a} \in \mathcal{L}(L^2 \cap L^s(\mathfrak{m}), L^2 + L^{r'}(\mathfrak{m}))$.

Finally, we remark that the arguments above still hold if, in the right hand side in (10.5) or (10.8), v is replaced with v^δ and g with g^γ, for any δ, $\gamma > 0$. As it is intuitively clear, the situation can only improve, and we obtain bounds not depending on δ or γ. ■

10.2.2 The commutator with a Sobolev derivation

In this section, we provide estimates for the commutator between P and a derivation $\boldsymbol{b}$, formally given by

$$[\mathsf{P}_\alpha, \boldsymbol{b}]f = \mathsf{P}_\alpha\left(df(\boldsymbol{b})\right) - d\left(\mathsf{P}_\alpha f\right)(\boldsymbol{b}). \tag{10.10}$$

We notice immediately that some assumption must be introduced, to ensure that the expression above actually represents a function, since $\boldsymbol{b}$ is defined on $\mathscr{A}$ and $\mathsf{P}_\alpha f$ does not necessary belongs to $\mathscr{A}$. Clearly, invariance $\mathsf{P}_\alpha \mathscr{A} \subseteq \mathscr{A}$ is sufficient to define (10.10) but, since below we require anyway $b \in L^q(\mathfrak{m})$ and the validity of the L^r-Γ inequality, we prove that density of $\mathscr{A}$ in $\mathbb{V}^r$ is enough.

More precisely, given a derivation $\boldsymbol{b} \in L^q(\mathfrak{m})$, if $\mathscr{A}$ be dense in $\mathbb{V}^s$, then, by Remark 4.9, $\boldsymbol{b}$ extends uniquely to a linear continuous operator from $\mathbb{V}^s$ into $L^{r'}(\mathfrak{m})$, thus (10.10) is well-defined in $L^{r'}(\mathfrak{m})$, for any $f \in \mathbb{V}^s$.

Our crucial estimate for 10.10 requires a notion of Sobolev regularity for $\boldsymbol{b}$, which corresponds in the smooth setting to some bound on the symmetric part of its derivative, the so-called deformation of $\boldsymbol{b}$.

Definition 10.5 (deformation)**.** *Assume that $\mathscr{A}$ is dense both in $\mathbb{V}^r$ and in $\mathbb{V}^s$, let $\boldsymbol{b}$ be a derivation with $|\boldsymbol{b}| \in L^q(\mathfrak{m})$ and $\operatorname{div} \boldsymbol{b} \in L^q(\mathfrak{m})$. We say that the deformation of $\boldsymbol{b}$ is of type (r,s) if there exists $c \geq 0$ satisfying*

$$\left| \int D^{sym} \boldsymbol{b}(u,f) d\mathfrak{m} \right| \leq c \| \sqrt{\Gamma(u)} \|_r \| \sqrt{\Gamma(f)} \|_s, \tag{10.11}$$

for all $u \in \mathbb{V}^r \cap D^r(\Delta)$ and all $f \in \mathbb{V}^s \cap D^s(\Delta)$, where

$$\int D^{sym} \boldsymbol{b}(u,f) d\mathfrak{m} := -\frac{1}{2} \int \left[df(\boldsymbol{b}) \Delta u + du(\boldsymbol{b}) \Delta f - (\operatorname{div} \boldsymbol{b}) \Gamma(u,f) \right] d\mathfrak{m}. \tag{10.12}$$

and we let $\|D^{sym} \boldsymbol{b}\|_{r,s}$ be the smallest constant c in (10.11).

The density assumption of $\mathscr{A}$ in $\mathbb{V}^r$ and $\mathbb{V}^s$ is again necessary to extend the derivation $\boldsymbol{b}$ to $\mathbb{V}^r$ and $\mathbb{V}^s$, again by Remark 4.9, so that (10.12) is well-defined. Notice that the expression in (10.12) is symmetric with respect to u and f, so it is the role of r and s above. The connection between (10.5), (10.12) and (10.4) can be recovered formally integrating by parts in (10.12), obtaining the identity

$$\int D^{sym} \boldsymbol{b}(u,f) d\mathfrak{m} := -\frac{1}{2} \left[\int u[\Delta, \boldsymbol{b}] f - u (\operatorname{div} \boldsymbol{b}) \Delta f - (\operatorname{div} \boldsymbol{b}) \Gamma(u,f) \right] d\mathfrak{m}. \tag{10.13}$$

In the following remark we show that the notion of deformation is more natural than the commutator between Δ and $\boldsymbol{b}$, at least from the point of view of Riemannian geometry, see also Chapter 11.

Remark 10.6 (deformation in the smooth case)**.** Let $(X, \langle \cdot, \cdot \rangle)$ be a compact Riemannian manifold, let $\mathfrak{m}$ be its associated Riemannian volume and let $\Gamma(u,f) = \langle \nabla u, \nabla f \rangle$. Let $df(\boldsymbol{b}) = \langle b, \nabla f \rangle$ for some smooth vector field b and let Db be the covariant derivative of b. The expression

$$\langle \nabla u, \nabla \langle b, \nabla f \rangle \rangle + \langle \nabla f, \nabla \langle b, \nabla u \rangle \rangle - \langle b, \nabla \langle \nabla f, \nabla u \rangle \rangle = \langle Db \, \nabla u, \nabla f \rangle + \langle Db \, \nabla f, \nabla u \rangle$$

gives exactly twice the symmetric part of the tensor Db, i.e. $2 \langle (D^{sym} b) f, u \rangle$. Integrating over X and then integrating by parts, we obtain twice the expression in (10.12), so that the deformation of a smooth field b is of type (r,s) if $|D^{sym} b| \in L^q(\mathfrak{m})$. ■

We are now in a position to state and prove our main commutator inequality.

Lemma 10.7 (commutator estimate for derivations)**.** *Assume that $\mathscr{A}$ is dense in $\mathbb{V}^p$ and that the L^p-Γ inequality holds, for $p \in \{r,s\}$. Let $\boldsymbol{b}$ be a derivation with $|\boldsymbol{b}|$, $\operatorname{div} \boldsymbol{b} \in L^q$ and deformation of type (r,s).*

Then, for every $\alpha \in (0,1)$, $u \in \mathbb{V}^r \cap D^r(\Delta)$, $f \in \mathbb{V}^s \cap D^s(\Delta)$, it holds

$$\left| \int u [\mathsf{P}_\alpha, \boldsymbol{b}] f d\mathfrak{m} \right| \leq c \left[\|D^{sym} \boldsymbol{b}\|_{r,s} + \|\operatorname{div} \boldsymbol{b}\|_{L^q + L^\infty} \right] \|u\|_{L^2 \cap L^r} \|f\|_{L^2 \cap L^s} \tag{10.14}$$

where c is some constant depending only on c_r^Γ, c_s^Γ in (3.12) *and c_r^Δ and c_s^Δ in* (3.14).

Proof. We introduce the following approximation of $\boldsymbol{b}$ via the action of P, that we denote by B, to stress the fact that it is not a derivation. For every $\alpha > 0$, we let B^α be the operator

$$L^2 \cap L^s(\mathfrak{m}) \ni f \mapsto \mathsf{B}^\alpha(f) := d(f^\alpha)(\boldsymbol{b}), \quad \text{where } f^\alpha := \mathsf{P}_\alpha f,$$

which belongs to $\mathscr{L}(L^2 \cap L^s(\mathfrak{m}), L^{r'}(\mathfrak{m}))$, since $\boldsymbol{b}$ extends to a linear continuous operator mapping $\mathbb{V}^s$ into $L^{r'}(\mathfrak{m})$ and the validity of the L^s-Γ inequality entails that the map $f \mapsto f^\alpha$ is continuous from $L^2 \cap L^s(\mathfrak{m})$ into $\mathbb{V}^s$. Moreover, the semigroup law gives

$$\mathsf{B}^{\delta+\alpha}(f) = \mathsf{B}^\delta(f^\alpha), \quad \text{for every } f \in L^2 \cap L^s(\mathfrak{m}) \text{ and } \delta > 0.$$

We claim that (10.14) is equivalent to the validity of

$$\left| \int u[\mathsf{P}_\alpha, \mathsf{B}^\delta] f d\mathfrak{m} \right| \le c \left[\|D^{sym}\boldsymbol{b}\|_{r,s} + \|\operatorname{div} b\|_{L^q} \right] \|u\|_{L^r \cap L^2} \|f\|_{L^s \cap L^2}, \tag{10.15}$$

for every $\delta > 0$, where $c \ge 0$ is some constant independent of δ.

Indeed, as $\delta \downarrow 0$, the left hand side above converges towards that of (10.14) since it holds

$$\int u^\alpha \mathsf{B}^\delta(f) d\mathfrak{m} = -\int \operatorname{div}(u^\alpha \boldsymbol{b})\, f^\delta d\mathfrak{m} \to -\int \operatorname{div}(u^\alpha \boldsymbol{b})\, f d\mathfrak{m} = \int u^\alpha df(\boldsymbol{b}) d\mathfrak{m},$$

as $f^\delta \to f$ strongly in $L^s(\mathfrak{m})$ and

$$\operatorname{div}(u^\alpha \boldsymbol{b}) = u^\alpha \operatorname{div} \boldsymbol{b} + d(u^\alpha)(\boldsymbol{b}) \in L^{s'}(\mathfrak{m}).$$

The second term in the commutator converges to the expected limit, since $\delta \mapsto \mathsf{B}^{\delta+\alpha}(f)$ is continuous on $[0, \infty)$, at fixed $\alpha > 0$.

As the left hand side in (10.15) is the commutator between P_α and B^δ, the result follows from an application of Lemma 10.3, with $\mathsf{A} = \mathsf{B}^\delta$ and $\mathsf{a} = (\operatorname{div} \boldsymbol{b})\mathsf{P}_\delta$. Notice that the operator norm of B^δ is unbounded as $\delta \downarrow 0$, but this causes no harm as it does not appear in the estimate that we establish.

The key point is to look for a rigorous version of (10.13), relating the deformation of $\boldsymbol{b}$ and the infinitesimal commutator

$$\int u[\Delta, \mathsf{B}^\delta] f d\mathfrak{m} = \int [\Delta u]\, \mathsf{B}^\delta(f) - u\mathsf{B}^\delta(\Delta f) d\mathfrak{m}.$$

By the definition of $\operatorname{div} \boldsymbol{b}$, it holds

$$\int u\mathsf{B}^\delta(\Delta f) d\mathfrak{m} = -\int du(\boldsymbol{b})\Delta f^\delta + u(\operatorname{div} \boldsymbol{b})\Delta f^\delta d\mathfrak{m},$$

and recalling that $\mathsf{B}^\delta(f) = d(f^\delta)(\boldsymbol{b})$, we obtain

$$\int u[\Delta, \mathsf{B}^\delta] f d\mathfrak{m} = \int \left[d(f^\delta)(\boldsymbol{b})\Delta u + du(\boldsymbol{b})\Delta f^\delta - u(\operatorname{div} \boldsymbol{b})\Delta f^\delta \right] d\mathfrak{m}.$$

Adding and subtracting $(\operatorname{div} \boldsymbol{b})\Gamma(u, f^\delta)$, we deduce the identity

$$\int u[\Delta, \mathsf{B}^\delta] f d\mathfrak{m} + \int u(\operatorname{div} \boldsymbol{b})\Delta f^\delta d\mathfrak{m} = -2\int D^{sym}\boldsymbol{b}(u, f^\delta) d\mathfrak{m} - \int (\operatorname{div} \boldsymbol{b})\Gamma(u, f^\delta) d\mathfrak{m}.$$

By the assumption on the deformation of $\boldsymbol{b}$, we conclude that

$$\left|\int u[\Delta, \mathsf{B}^\delta] f d\mathfrak{m} + \int u(\operatorname{div} \boldsymbol{b})\Delta f^\delta d\mathfrak{m}\right| \le \left[2\left\|D^{sym}\boldsymbol{b}\right\|_{r,s} + \left\|\operatorname{div}\boldsymbol{b}\right\|_q\right] \|u\|_{\mathbb{V}^r} \|f^\delta\|_{\mathbb{V}^s}.$$

Finally, in order to apply Lemma 10.3, we are only left with checking that the same inequality holds with u^t in place of u and f^τ in place of f, for t, $\tau \in (0,1)$, but this is straightforward: we only used the assumptions $u \in \mathbb{V}^r \cap D^r(\Delta)$ and $f \in \mathbb{V}^s \cap D^s(\Delta)$, that are stable for the action of the semigroup P. Thus, Lemma 10.3 provides (10.15). □

Remark 10.8 (the divergence free case)**.** If $\operatorname{div}\boldsymbol{b} = 0$, there is no need to add or subtract any term and the result above is a direct consequence of Lemma 10.7. ■

Corollary 10.9 (strong convergence of commutators)**.** *Assume that $\mathscr{A}$ is dense in $\mathbb{V}^p$ and that the L^p-Γ inequality holds, for $p \in \{r,s\}$. Let $\boldsymbol{b}$ be a derivation with $|\boldsymbol{b}|$, $\operatorname{div}\boldsymbol{b} \in L^q$ and deformation of type (r,s).*

Then, for every $\alpha > 0$, the commutator operator

$$\mathbb{V}^s \ni f \quad \mapsto \quad [\mathsf{P}_\alpha, \boldsymbol{b}]f \in L^{r'}(\mathfrak{m})$$

extends uniquely to a linear continuous operator in from $L^2 \cap L^s(\mathfrak{m})$ into $L^2(\mathfrak{m}) + L^{r'}(\mathfrak{m})$. Moreover, for every $f \in L^2 \cap L^s(\mathfrak{m})$, it holds

$$[\mathsf{P}_\alpha, \boldsymbol{b}]f \to 0, \quad \textit{strongly in } L^2(\mathfrak{m}) + L^{r'}(\mathfrak{m}), \textit{ as } \alpha \downarrow 0.$$

Proof. By duality and density, (10.14) entails that the commutator extends uniquely to a linear continuous operator as claimed. Strong convergence as $\alpha \downarrow 0$ is proved arguing first with f^δ in place of f, at fixed $\delta > 0$. This follows because of continuity of $\alpha \mapsto d(f^{\alpha+\delta})(\boldsymbol{b})$, for $\alpha \downarrow 0$. The general case is a consequence of uniform boundedness in $\alpha \in (0,1)$ for the operator norm of $[\mathsf{P}_\alpha, \boldsymbol{b}]$ and density of this class of functions in $L^2 \cap L^s(\mathfrak{m})$. □

10.2.3 The commutator with a Sobolev diffusion operator

In this section we study the commutator between the heat semigroup P and a diffusion operator of the form $a\Delta$, defined by

$$[\mathsf{P}_\alpha, a\Delta]f = \mathsf{P}_\alpha(a\Delta f) - a\Delta(\mathsf{P}_\alpha f).$$

In this case, it is sufficient to assume $\mathscr{A} \subseteq D(\Delta)$ so that the quantity above is well-defined. We are actually interested in the case that $a \in L^q(\mathfrak{m})$, thus the commutator is a linear continuous operator from $D^s(\Delta)$ into $L^{r'}(\mathfrak{m})$. The estimates that we establish require more that $a \in L^q(\mathfrak{m})$, precisely Sobolev regularity for a up to second order, i.e., we assume $a \in \mathbb{V}^q \cap D^q(\Delta)$, together with some bound on the deformation of gradient derivation $\boldsymbol{b}_a$, i.e. $df(\boldsymbol{b}_a) = \Gamma(a,f)$. In what follows, we let

$$H[a] := D^{sym}\boldsymbol{b}_a \tag{10.16}$$

be the Hessian of a.

Remark 10.10. For the only purpose of well-posedness for the FPE associated to the operator $\mathcal{L}f := a\Delta f$, stronger ad-hoc techniques are available in the recent literature, e.g. [Barbu et al., 2011] and [Belaribi and Russo, 2012], which seem to extend readily to the framework of metric measure spaces, and require virtually no regularity of a. The approach that we follow below, based on a refined study of the commutator between P and the multiplication operator, introduces stronger regularity assumptions but entails uniqueness also for diffusion operators of the form $\mathcal{L}f := a\Delta f + df(\boldsymbol{b})$. Finally, similar computations allow for dealing with general diffusions $\mathcal{L}f := a : \nabla^2 f$ (at least in Euclidean spaces, see Chapter 11). ■

Lemma 10.11 (commutator estimate for diffusion operators)**.** *Assume that $\mathscr{A}$ is dense in $\mathbb{V}^p$ and that the L^p-Γ inequality holds, for $p \in \{r, s\}$. Let $a \in \mathbb{V}^q \cap D^q(\Delta)$ and with Hessian of type (r, s).*

Then, for every $\alpha \in (0,1)$, $u \in \mathbb{V}^r \cap D^r(\Delta)$, $f \in \mathbb{V}^s \cap D^s(\Delta)$, it holds

$$\left| \int u[\mathsf{P}_\alpha, a\Delta] f d\mathfrak{m} - \alpha \int u[\Delta, a]\Delta \mathsf{P}_\alpha f d\mathfrak{m} \right| \le c \left[\|H[a]\|_{r,s} + \|\Delta a\|_q \right] \|u\|_{L^2 \cap L^r} \|f\|_{L^2\cap L^s} \tag{10.17}$$

where c is some constant depending only on c_r^Γ, c_s^Γ in (3.12) and c_r^Δ and c_s^Δ in (3.14).

Proof. The proof relies on a refined study of the commutator between P_α and the linear continuous operator $f \mapsto af$. Indeed, since P_α commutes with the Laplacian Δ, it holds

$$\int u[\mathsf{P}_\alpha, a\Delta] f d\mathfrak{m} = \int u[\mathsf{P}_\alpha, a] h d\mathfrak{m},$$

where we denote $h := \Delta f$. Without loss of generality, by replacing f with f^δ for some $\delta > 0$ and then let $\delta \downarrow 0$, we assume also $h \in \mathbb{V}^s \cap D^s(\Delta)$.

As in the proof of Lemma 10.2, we introduce the curve

$$[0, \alpha] \ni \sigma \mapsto F(\sigma) = \int u^\sigma a h^{\alpha - \sigma} d\mathfrak{m},$$

which is $C^1((0,\alpha), \mathbb{R})$, with

$$F'(\sigma) = \int u^\sigma [\Delta, a] h^{\alpha-\sigma} d\mathfrak{m}.$$

After Example 10.1, it holds

$$\begin{aligned} F'(\sigma) &= \int u^\sigma dh^{\alpha-\sigma}(\boldsymbol{b}_a) - du^\sigma(\boldsymbol{b}_a) h^{\alpha-\sigma} d\mathfrak{m} \\ &= \int u^\sigma dh^{\alpha-\sigma}(\boldsymbol{b}_a) + u^\sigma \operatorname{div}(h^{\alpha-\sigma}\boldsymbol{b}_a) d\mathfrak{m} \\ &= 2\int u^\sigma dh^{\alpha-\sigma}(\boldsymbol{b}_a) d\mathfrak{m} + \int u^\sigma (\Delta a) h^{\alpha-\sigma} d\mathfrak{m}. \end{aligned} \tag{10.18}$$

This identity gives at once that $F \in C^2((0,\alpha),\mathbb{R})$, with

$$F''(\sigma) = 2\int u^\sigma [\Delta, \boldsymbol{b}_a] h^{\alpha-\sigma} d\mathfrak{m} + \int u^\sigma [\Delta, (\Delta a)] h^{\alpha-\sigma} d\mathfrak{m}.$$

The main idea is to perform an interpolation up to the second order,

$$F(\alpha) - F(0) - \alpha F'(0) = \int_0^\alpha F''(\sigma)(\alpha - \sigma) d\sigma, \tag{10.19}$$

in place of the fundamental theorem of calculus: the factor $(\alpha - \sigma)$ is useful to compensate the bound on the norm of $h^{\alpha-\sigma} = \Delta f^{\alpha-\sigma}$. Notice that $\alpha F'(0)$ is precisely the second term in the left hand side of (10.17).

As with the case of derivations (see Remark 10.8), our deduction is straightforward in case $\Delta a = 0$, since (10.13) gives

$$\int u^\sigma [\Delta, \boldsymbol{b}_a] h^{\alpha-\sigma} d\mathfrak{m} = -2 \int H[a](u^\sigma, h^{\alpha-\sigma}) d\mathfrak{m}$$

and we estimate

$$\begin{aligned} \left|F''(\sigma)\right| &\leq 4 \left\|H[a]\right\|_{r,s} \|\sqrt{\Gamma(u^\sigma)}\|_r \|\sqrt{\Gamma(h^{\alpha-\sigma})}\|_s \\ &\leq \frac{8 c_r^\Gamma c_s^\Gamma}{\sqrt{\sigma(\alpha-\sigma)}} \left\|H[a]\right\|_{r,s} \|u\|_r \|h^{(\alpha-\sigma)/2}\|_s \\ &\leq \frac{16 c_r^\Gamma c_s^\Gamma c_s^\Delta}{\sqrt{\sigma(\alpha-\sigma)^3}} \left\|H[a]\right\|_{r,s} \|u\|_r \|f\|_s , \end{aligned} \tag{10.20}$$

where we apply both the L^p-Γ inequality 3.11 for $p \in \{r, s\}$ and the L^s-Δ inequality 3.14. Integrating with respect to $\sigma \in (0, \alpha)$, we would conclude 10.17.

To address the general case where $\Delta a \in L^q(\mathfrak{m})$, we argue as in the proof of Lemma 10.2, i.e. we add and subtract suitable quantities. Unfortunately, the proof becomes less straightforward. We consider separately the terms

$$\int_0^\alpha \int u^\sigma [\Delta, \boldsymbol{b}_a] h^{\alpha-\sigma} d\mathfrak{m}\, (\alpha-\sigma) d\sigma \quad \text{and} \quad \int_0^\alpha \int u^\sigma [\Delta, (\Delta a)] h^{\alpha-\sigma} d\mathfrak{m}\, (\alpha-\sigma) d\sigma. \tag{10.21}$$

We focus on the former. By (10.13), it holds, for $\sigma \in (0, \alpha)$,

$$\int H[a](u^\sigma, h^{\alpha-\sigma}) d\mathfrak{m} = -\frac{1}{2} \int u^\sigma [\Delta, \boldsymbol{b}_a] h^{\alpha-\sigma} - u^\sigma (\Delta a) \Delta h^{\alpha-\sigma} - (\Delta a) \Gamma(u^\sigma, h^{\alpha-\sigma}) d\mathfrak{m},$$

thus, in order to reduce to the argument for the case $\Delta a = 0$, it is enough to provide bounds for the quantities

$$\int u^\sigma (\Delta a) \Delta h^{\alpha-\sigma} d\mathfrak{m}, \quad \text{and} \quad \int (\Delta a) \Gamma(u^\sigma, h^{\alpha-\sigma}) d\mathfrak{m}. \tag{10.22}$$

The inequality

$$\left| \int (\Delta a) \Gamma(u^\sigma, h^{\alpha-\sigma}) d\mathfrak{m} \right| \leq \|\Delta a\|_q \|\sqrt{\Gamma(u^\sigma)}\|_r \left\|h^{\alpha-\sigma}\right\|_s$$

allows us to handle the second term in (10.22) exactly as in (10.20), while for the first term in (10.22), we use a second-order analogue of Lemma 10.3. We introduce the quantity

$$\int_0^\alpha \int u^\alpha (\Delta a) \Delta^2 f^{\alpha-\sigma} (\alpha-\sigma) d\mathfrak{m} d\sigma. \tag{10.23}$$

By the Taylor expansion (10.19) with $f^{\alpha-\sigma}$ in place of $F(\sigma)$, we have

$$\int_0^\alpha \Delta^2 f^{\alpha-\sigma} (\alpha-\sigma) d\sigma = f - f^\alpha + \alpha \Delta f^\alpha,$$

entailing the bound

$$\left|\int_0^\alpha \int u^\alpha (\Delta a)\Delta^2 f^{\alpha-\sigma}(\alpha-\sigma) d\mathfrak{m} d\sigma\right| \le (2 + c_s^\Delta) \left\|\Delta a\right\|_q \left\|u\right\|_r \left\|f\right\|_s .$$

Therefore, we are allowed to add and subtract 10.23 in the first term of (10.22), and we are reduced to provide a bound for difference

$$\left|\int (u^\alpha - u^\sigma)(\Delta a)\Delta h^{\alpha-\sigma} d\mathfrak{m}\right|,$$

to be integrated over $\sigma \in (0,\alpha)$, with respect to the measure $(\alpha - \sigma)d\sigma$. By the L^s-Δ inequality, it holds

$$\begin{aligned}\left|\int (u^\alpha - u^\sigma)(\Delta a)\Delta h^{\alpha-\sigma} d\mathfrak{m}\right| &\le \left\|\Delta a\right\|_q \left\|u^\alpha - u^\sigma\right\|_r \left\|\Delta h^{\alpha-\sigma}\right\|_s \\ &\le \frac{2c_s^\Delta}{\alpha-\sigma} \left\|\Delta a\right\|_q \left\|u^\alpha - u^\sigma\right\|_r \left\|\Delta f^{(\alpha-\sigma)/2}\right\|_s\end{aligned}$$

and from this point we conclude identically as in the proof of Lemma 10.3, i.e. by the L^p-Δ inequality and Corollary 3.5. This provides the required bounds for the former term in (10.21).

The latter term in (10.21) is easier to bound, since

$$\int u^\sigma [\Delta, (\Delta a)] h^{\alpha-\sigma} d\mathfrak{m} = \frac{d}{ds} \int u^\sigma (\Delta a) h^{\alpha-\sigma} d\mathfrak{m},$$

and so we can integrate by parts in (10.21),

$$\int_0^\alpha \int u^\sigma [\Delta, (\Delta a)] h^{\alpha-\sigma} d\mathfrak{m}\, (\alpha-\sigma) d\sigma = -\alpha \int u(\Delta a) h^\alpha d\mathfrak{m} + \int_0^\alpha \int u^\sigma (\Delta a) h^{\alpha-\sigma} d\mathfrak{m}.$$

The first addend in the right hand side is uniformly bounded from above by $c_s^\Delta \left\|\Delta a\right\|_q \left\|u\right\|_r \left\|f\right\|_s$, so we are left only with

$$\int_0^\alpha \int u^\sigma (\Delta a) h^{\alpha-\sigma} d\mathfrak{m} = \int_0^\alpha \int u^\sigma (\Delta a) \Delta f^{\alpha-\sigma} d\mathfrak{m},$$

but this can be handled directly as in the proof of Lemma 10.3, i.e. by adding and subtracting

$$\int_0^\alpha \int u^\alpha (\Delta a) \Delta f^{\alpha-\sigma} d\mathfrak{m} = \int u^\alpha (\Delta a)(f^\alpha - f) d\mathfrak{m},$$

and using the L^p-Δ inequality and Corollary 3.5. □

As a consequence of Lemma 10.11 and duality arguments, one define uniquely by extension, the family of continuous operators

$$L^2 \cap L^s(\mathfrak{m}) \supseteq D^s(\Delta) \ni f \mapsto [\mathsf{P}_\alpha, a\Delta] f - \alpha[\Delta, a]\Delta \mathsf{P}_\alpha f \in L^2(\mathfrak{m}) + L^{r'}(\mathfrak{m}),$$

which is moreover uniformly bounded, for $\alpha \in (0,1)$. Since for $f \in D^s(\Delta)$ we obtain strong convergence towards 0 in $L^2(\mathfrak{m}) + L^{r'}(\mathfrak{m})$, by density we also have

$$[\mathsf{P}_\alpha, a\Delta] f - \alpha[\Delta, a]\Delta \mathsf{P}_\alpha f \to 0, \quad \text{as } \alpha \downarrow 0, \text{ for every } f \in L^2 \cap L^s(\mathfrak{m}).$$

Unluckily, it seems impossible to provide uniform bounds for the quantity

$$\alpha \int u[\Delta, a]\Delta \mathsf{P}_\alpha f d\mathfrak{m},$$

and deduce strong convergence for the original commutator. Indeed, by (10.17), the quantity

$$\int u[\Delta, a]\Delta \mathsf{P}_\alpha f d\mathfrak{m} = \int (\Delta u) a \Delta \mathsf{P}_\alpha f d\mathfrak{m} - ua\Delta^2 \mathsf{P}_\alpha f d\mathfrak{m}$$

is of course well defined if $u \in D^r(\Delta)$, and (10.18) shows that

$$\int u[\Delta, a]\Delta \mathsf{P}_\alpha f d\mathfrak{m} = \int u\left[2\, d(\Delta \mathsf{P}_\alpha f)(\boldsymbol{b}_a) + (\Delta a)\Delta \mathsf{P}_\alpha f\right] d\mathfrak{m},$$

leading to the bound

$$\left|\alpha \int u[\Delta, a]\Delta \mathsf{P}_\alpha f d\mathfrak{m}\right| \leq \frac{c}{\sqrt{\alpha}}\left[\|\boldsymbol{b}_a\|_q + \|\Delta a\|_q\right]\|u\|_r \|f\|_s .$$

However, to prove uniqueness, we choose $f = \mathsf{P}_\alpha u$, and this additional symmetry can be crucially exploited, as we do in the following lemma.

Lemma 10.12. *Assume that $\mathscr{A}$ is dense in $\mathbb{V}^p$ and that the L^p-Γ inequality holds, for $p \in \{r, s\}$ and that* (3.13) *holds, for every $f \in \mathbb{V}^p$. Let $a \in \mathbb{V}^q \cap D^q(\Delta)$, with $H[a]$ of type (r, s).*

Then, for every $u \in L^2 \cap L^r \cap L^s(\mathfrak{m})$ and $\alpha \in (0,1)$, it holds

$$\left|\alpha \int u[\Delta, a]\Delta \mathsf{P}_\alpha(\mathsf{P}_\alpha u) d\mathfrak{m}\right| \leq c\left[\|H[a]\|_{r,s} + \|\Delta a\|_q\right]\|u\|^2_{L^2\cap L^r \cap L^s}, \tag{10.24}$$

where c is some constant depending only on c^Γ_r, c^Γ_s in (3.12) *and c^Δ_r and c^Δ_s in* (3.14)*. Moreover, the left hand side in* (10.24) *is infinitesimal as $\alpha \downarrow 0$.*

Proof. First, we obtain an equivalent expression where the semigroup P_α acts on the leftmost u, gaining more symmetry. Indeed, in general, if $u \in \mathbb{V}^r \cap D^r(\Delta)$, $f \in \mathbb{V}^s \cap D^s(\Delta)$, integrating by parts, it holds

$$\begin{aligned}
\int u[\Delta, a]\Delta f^\alpha d\mathfrak{m} &= -2\int du(\boldsymbol{b}_a)\Delta f^\alpha d\mathfrak{m} - \int u(\Delta a)\Delta f^\alpha d\mathfrak{m}\\
&= -2\int (\Delta f)\mathsf{P}_\alpha(du(\boldsymbol{b}_a)) d\mathfrak{m} - \int u(\Delta a)\Delta f^\alpha d\mathfrak{m}\\
&= -2\int (\Delta f)[\mathsf{P}_\alpha, \boldsymbol{b}_a] u d\mathfrak{m} - 2\int (\Delta f) du^\alpha(\boldsymbol{b}_a) d\mathfrak{m} - \int u(\Delta a)\Delta f^\alpha d\mathfrak{m}.
\end{aligned}$$

By the first statement in Corollary 10.9, with the roles of r and s reversed, this last identity extends by continuity to the case $u \in L^2 \cap L^r(\mathfrak{m})$, $f \in D^s(\Delta)$.

We now specialize to the case $f := \mathsf{P}_\alpha u$. As $\alpha \downarrow 0$, it holds

$$\alpha \left|\int (\Delta u^\alpha)[\mathsf{P}_\alpha, \boldsymbol{b}_a] u d\mathfrak{m}\right| \to 0,$$

by the second statement in Corollary 10.9 and boundedness of $\alpha \Delta u^\alpha$ in $L^2 \cap L^s(\mathfrak{m})$. We also have

$$\left|\int u(\Delta a)\Delta f^\alpha d\mathfrak{m}\right| \leq \|\Delta a\|_q \|u\|_r \|\Delta u^{2\alpha}\|_s \to 0,$$

by Proposition 3.4.

In order to handle the last term, our choice of f in terms of u seems crucial. Indeed, we have

$$\int (\Delta u^\alpha) du^\alpha(\boldsymbol{b}_a) d\mathfrak{m} = -\int H[a](u^\alpha, u^\alpha) + \frac{1}{2}\int (\Delta a)\Gamma(u^\alpha) d\mathfrak{m},$$

using the very definition of deformation: this leads to the inequality

$$\left|\int (\Delta u^\alpha) du^\alpha(\boldsymbol{b}_a) d\mathfrak{m}\right| \le \left[\|H[a]\|_{r,s} + \|\Delta a\|_q\right] \|\sqrt{\Gamma(u^\alpha)}\|_r \|\sqrt{\Gamma(u^\alpha)}\|_s,$$

which, by the L^p-Γ inequality, for $p \in \{r, s\}$, entails (10.24) together with convergence towards 0 of the left hand side therein, by Proposition 3.6. □

We conclude by collecting the results proved in this section in the following corollary.

Corollary 10.13 (convergence of commutators, diffusion operators)**.** *Assume that $\mathscr{A}$ is dense in $\mathbb{V}^p$ and that the L^p-Γ inequality holds, for $p \in \{r, s\}$. Let $a \in \mathbb{V}^q \cap D^q(\Delta)$ with $H[a]$ of type (r, s).*

Then, for every $\alpha > 0$, the commutator

$$D^s(\Delta) \ni f \mapsto [\mathsf{P}_\alpha, a\Delta] f \in L^{r'}(\mathfrak{m})$$

extends uniquely to a linear continuous operator from $L^2 \cap L^s$, to $L^2(\mathfrak{m}) + L^{r'}(\mathfrak{m})$. Moreover, for every $u \in L^2 \cap L^r \cap L^s(\mathfrak{m})$, it holds

$$\left|\int u[\mathsf{P}_\alpha, a\Delta](\mathsf{P}_\alpha u) d\mathfrak{m}\right| \to 0, \quad \text{as } \alpha \downarrow 0.$$

Proof. By Lemma 10.11 and the discussion right after it, we deduce

$$\left|\int u[\mathsf{P}_\alpha, a\Delta](\mathsf{P}_\alpha u) d\mathfrak{m} - \alpha\int u[\Delta, a]\Delta \mathsf{P}_\alpha(\mathsf{P}_\alpha u) d\mathfrak{m}\right| \to 0, \quad \text{as } \alpha \downarrow 0.$$

Indeed, it is sufficient to notice that, if a family of uniformly bounded operators $(G_\alpha)_\alpha$ on a Banach space B strongly converges to 0, i.e. for every $u \in B$, $G_\alpha u \to 0$, then if $u_\alpha \to u$ in B, it holds $G_\alpha u_\alpha \to Gf$, by the inequality

$$\|G_\alpha f_\alpha - Gf\|_B \le \|f_\alpha - f\|_B \sup_{\alpha'} \|G_{\alpha'}\| + \|G_\alpha f\|.$$

Finally, the second statement in Lemma 10.12, gives

$$\left|\alpha\int u[\Delta, a]\Delta \mathsf{P}_\alpha(\mathsf{P}_\alpha u) d\mathfrak{m}\right| \to 0, \quad \text{as } \alpha \downarrow 0,$$

and the conclusion follows. □

10.2.4 The commutator with ∂_t

In this section, we study some features that appear in the time-extended framework, as described in Chapter 5. First of all, we notice that the abstract results for the commutators in the sections above still hold when we consider functions, derivations and diffusions, assuming only integrable bounds with respect to $t \in (0,T)$: indeed the bounds established therein are quantitative and one is in a position to apply Lebesgue dominated convergence theorem.

Our next goal is to study the commutator between ∂_t and $\mathsf{R} := \mathsf{P}[\boldsymbol{a}]$, where $\boldsymbol{a} = (\boldsymbol{a}_t)_{t\in(0,T)}$ is some bounded elliptic 2-tensor on $(0,T)\times X$, as in Remark 4.17, formally given by

$$[\mathsf{R}_\alpha, \partial_t] f := \mathsf{R}_\alpha(\partial_t f) - \partial_t(\mathsf{R}_\alpha f). \tag{10.25}$$

The first problem that we address is whether $\partial_t \mathsf{R}_\alpha f$ can be defined as a function, even if f is smooth. This relies on some smoothness assumption for $(0,T) \ni t \mapsto \boldsymbol{a}_t$ and clearly all the results below trivially apply to the case of constant coefficients $\boldsymbol{a}_t(df) = \Gamma(f)$. However, to give an intuition of the regularity that enters in the picture, let us consider the case of $X = \mathbb{R}^n$ with $\boldsymbol{a}_t(df) = a_t |\nabla f|^2$, for some function $a : (0,T)\times\mathbb{R}^n \to [0,\infty)$, so that $\Delta[\boldsymbol{a}]_t f = \operatorname{div}(a_t \nabla f)$. To compute the commutator $[\mathsf{R}_\alpha, \partial_t] f$, by interpolation along the semigroup, we are reduced to the "infinitesimal" commutator, which reads as

$$[\operatorname{div}(a_t\nabla), \partial_t] f = \operatorname{div}(a \nabla \partial_t f) - \partial_t \operatorname{div}(a_t \nabla f) = -\operatorname{div}\left[(\partial_t a)\nabla f\right],$$

since ∂_t commutes with ∇.

In order to address rigorously this issue, we notice first that (10.25) can be defined in a weak form, for u, $f \in C^1_c((0,T); L^2(\mathfrak{m}))$, letting

$$\int u [\mathsf{R}_\alpha, \partial_t] f \, d\widetilde{\mathfrak{m}} := \int \left[u \mathsf{R}_\alpha(\partial_t f) + (\partial_t u)\mathsf{R}_\alpha f\right] d\widetilde{\mathfrak{m}}. \tag{10.26}$$

Then, we formulate Sobolev regularity for the time-dependent derivation $(\boldsymbol{a}_t)_{t\in(0,T)}$ as follows.

Definition 10.14. *Let $\boldsymbol{a} = (\boldsymbol{a}_t)_{t\in(0,T)}$ be bounded 2-tensor, i.e. $|\boldsymbol{a}| \in L^\infty_t(L^\infty_x)$. We say that $\partial_t \boldsymbol{a}$ is of type (r,s) if there exists $c \in L^q(0,T)$ such that, for every $u \in C^1_c((0,T); \mathbb{V}^r)$, $f \in C^1_c((0,T);\mathbb{V}^s)$, the curve $(0,T) \ni t \mapsto \int \boldsymbol{a}_t(u_t, f_t) d\mathfrak{m}$ is weakly differentiable, with*

$$\left| \partial_t \int \boldsymbol{a}_t(u_t, f_t) d\mathfrak{m} - \int \left[\boldsymbol{a}_t(\partial_t u_t, f_t) + \boldsymbol{a}_t(u_t, \partial_t f_t)\right] d\mathfrak{m} \right| \leq c_t \|u_t\|_{\mathbb{V}^r} \|f_t\|_{\mathbb{V}^s}, \mathscr{L}^1\text{-a.e. } t \in (0,T). \tag{10.27}$$

We let $|\partial_t \boldsymbol{a}|_{r,s}$ be the smallest function such that (10.27) holds and $\|\partial_t \boldsymbol{a}\|_{r,s}$ its norm in $L^q(0,T)$.

Recall that we assume $\mathscr{A}$ to be dense in $L^2_t(\mathbb{V})$, thus $\boldsymbol{a}$ above extends to a 2-tensor on it, by Remark 4.16, and the second integral in (10.27) is well-defined.

We introduce translation operators on $(0,T)\times X$, letting $\mathsf{T}_\sigma f(t,x) := f(t+\sigma, x)$ if $t + \sigma \in (0,T)$, and $\mathsf{T}_\sigma f(t,x) := 0$ otherwise. We show that Sobolev regularity for $(0,T) \ni t \mapsto \boldsymbol{a}_t$ entails bounds for suitable difference quotients.

Lemma 10.15. *Let $\boldsymbol{a} = (\boldsymbol{a}_t)_t$ be a bounded 2-tensor, with $\partial_t \boldsymbol{a}$ of type (r,s). Let $u \in L^r_t(\mathbb{V}^r)$, $f \in L^s_t(\mathbb{V}^s)$, with compact support on $(0,T)$, i.e.*

$$u(t) = 0, \quad f(t) = 0 \quad \widetilde{\mathfrak{m}}\text{-a.e. } (t,x) \in X \text{ for } t \text{ or } T-t \text{ small enough.} \tag{10.28}$$

Then, for σ belonging to a suitable neighbourhood of 0 *(depending on the support of u and f only), it holds*

$$\left|\int \boldsymbol{a}(\mathsf{T}_\sigma u, \mathsf{T}_\sigma f)d\widetilde{\mathfrak{m}} - \int \boldsymbol{a}(u,f)d\widetilde{\mathfrak{m}}\right| \le \sigma \left\|\partial_t \boldsymbol{a}\right\|_{r,s} \left\|u\right\|_{L^r_t(\mathbb{V}^r)} \left\|f\right\|_{L^s_t(\mathbb{V}^s)}.$$

Proof. Arguing by density, it is sufficient to assume that $u \in C^1_c((0,T);\mathbb{V}^r)$ and $f \in C^1_c((0,T);\mathbb{V}^s)$. The curves

$$\sigma \mapsto u^\sigma := \mathsf{T}_\sigma u \in L^r_t(\mathbb{V}_r), \quad \sigma \mapsto f^\sigma := \mathsf{T}_\sigma f \in L^s_t(\mathbb{V}_s)$$

are continuously differentiable in a neighbourhood of $\sigma = 0$, with derivatives

$$\frac{d}{d\sigma}u^\sigma = \frac{d}{dt}(u^\sigma), \quad \frac{d}{d\sigma}f^\sigma = \frac{d}{dt}(f^\sigma).$$

Since the map

$$L^r(\mathbb{V}_r) \times L^s_t(\mathbb{V}_s) \ni (u,f) \mapsto \int \boldsymbol{a}(u,f)d\widetilde{\mathfrak{m}}$$

is bilinear and continuous, we deduce that

$$\sigma \mapsto \int \boldsymbol{a}(u^\sigma, f^\sigma)d\widetilde{\mathfrak{m}}$$

is continuously differentiable in a neighbourhood of 0, with

$$\begin{aligned}
\frac{d}{d\sigma}\int \boldsymbol{a}(u^\sigma, f^\sigma)d\widetilde{\mathfrak{m}} &= \int \boldsymbol{a}(\partial_t(u^\sigma), f^\sigma) + \boldsymbol{a}(u^\sigma, \partial_t(f^\sigma))d\widetilde{\mathfrak{m}} \\
&\le \int_0^T \partial_t \int \boldsymbol{a}_t(u^\sigma_t, f^\sigma_t)d\mathfrak{m} + \left\|\partial_t \boldsymbol{a}\right\|_{r,s} \left\|u^\sigma\right\|_{L^r_t(\mathbb{V}^r)} \left\|f^\sigma\right\|_{L^s_t(\mathbb{V}^s)} \\
&\le \left\|\partial_t \boldsymbol{a}\right\|_{r,s} \left\|u\right\|_{L^r_t(\mathbb{V}^r)} \left\|f\right\|_{L^s_t(\mathbb{V}^s)}
\end{aligned}$$

where we use (10.27) and that the continuous representative for $t \mapsto \int \boldsymbol{a}_t(u^\sigma_t, f^\sigma_t)d\mathfrak{m}$ at $t=0$ and $t=T$ is zero, because of (10.28). □

Lemma 10.16 (commutator estimate, bounded elliptic case)**.** *Let $\boldsymbol{a} = (\boldsymbol{a}_t)_t$ be a bounded elliptic form with $\partial_t \boldsymbol{a}$ of type (r,s). For $p \in \{r,s\}$, let the L^p-Γ inequality hold, with respect to the semigroup R, i.e., for some constant c^Γ_p,*

$$\|\sqrt{\Gamma\left(\mathsf{R}_\alpha f\right)}\|_{L^p(\widetilde{\mathfrak{m}})} \le c^\Gamma_p \alpha^{-1/2} \left\|f\right\|_{L^p(\widetilde{\mathfrak{m}})} \quad \textit{for every } f \in L^2 \cap L^p(\widetilde{\mathfrak{m}}),\ \alpha \in (0,1). \tag{10.29}$$

Then, for every $\alpha \in (0,1)$, $u \in C^1_c((0,T);L^2 \cap L^r(\mathfrak{m}))$, $f \in C^1_c((0,T);L^2 \cap L^s(\mathfrak{m}))$, it holds, for the commutator defined by (10.26)*,*

$$\left|\int u[\mathsf{R}_\alpha, \partial_t]f d\mathfrak{m}\right| \le c \left\|\partial_t \boldsymbol{a}\right\|_{r,s} \left\|u\right\|_{L^2\cap L^r(\widetilde{\mathfrak{m}})} \left\|f\right\|_{L^2\cap L^s(\widetilde{\mathfrak{m}})}, \tag{10.30}$$

where c depends on the ellipticity constant of $\boldsymbol{a}$, and c^Γ_p in (10.29)*, for $p \in \{r,s\}$.*

Since $\boldsymbol{a}$ is bounded and elliptic, the validity of (10.29) is equivalent to

$$\|\sqrt{\boldsymbol{a}(\mathsf{P}[\boldsymbol{a}]_\alpha f)}\|_{L^p(\widetilde{\mathfrak{m}})} \le c^\Gamma_p \alpha^{-1/2} \left\|f\right\|_{L^p(\widetilde{\mathfrak{m}})} \quad \text{for every } f \in L^2 \cap L^p(\widetilde{\mathfrak{m}}),\ \alpha \in (0,1).$$

possibly with a different constant c^Γ_p.

Proof. We claim that it is enough to show the following analogue of (10.30), where ∂_t is replaced with $\sigma^{-1}\mathsf{T}_\sigma$, and $\sigma \neq 0$ small enough:

$$\left|\int u[\mathsf{R}_\alpha, \sigma^{-1}\mathsf{T}_\sigma] f d\mathfrak{m}\right| \leq c \left\|\partial_t \boldsymbol{a}\right\|_{r,s} \left\|u\right\|_r \left\|f\right\|_s,$$

where c is some constant depending on the ellipticity of $\boldsymbol{a}$ and c_p^Γ only, for $p \in \{r, s\}$. Once this is holds, we have

$$\int u[\mathsf{R}_\alpha, \sigma^{-1}(\mathsf{T}_\sigma - Id)] f d\mathfrak{m} = \int u[\mathsf{R}_\alpha, \sigma^{-1}\mathsf{T}_\sigma] f d\mathfrak{m},$$

and

$$\int u[\mathsf{R}_\alpha, \sigma^{-1}(\mathsf{T}_\sigma - Id)] f d\mathfrak{m} \to \int u[\mathsf{R}_\alpha, \partial_t] f d\mathfrak{m}, \quad \text{as } \sigma \to 0.$$

Notice that $f \mapsto \sigma^{-1}\mathsf{T}_\sigma f$ is a linear continuous operator mapping $L^2 \cap L^r(\widetilde{\mathfrak{m}})$ into itself, with norm smaller than σ^{-1}. We prove the claim using Lemma 10.7, where we let $\mathsf{P} := \mathsf{R}$ and $\mathsf{A} = \sigma^{-1}\mathsf{T}_\sigma$. Indeed, the infinitesimal commutator (10.4) reads as σ^{-1} times

$$\begin{aligned}
\int u[\Delta[\boldsymbol{a}], \mathsf{T}_\sigma] f d\widetilde{\mathfrak{m}} &= \int (\Delta[\boldsymbol{a}]u) f^\sigma - u^\sigma (\Delta[\boldsymbol{a}] f) d\widetilde{\mathfrak{m}} \\
&= \int (\Delta[\boldsymbol{a}]u) f^\sigma - u^{-\sigma} (\Delta[\boldsymbol{a}] f) d\widetilde{\mathfrak{m}} \\
&= \int \boldsymbol{a}(u^{-\sigma}, f) d\widetilde{\mathfrak{m}} - \int \boldsymbol{a}(u, f^\sigma) d\widetilde{\mathfrak{m}}.
\end{aligned}$$

By Lemma 10.15, with $u^{-\sigma}$ in place of u, we deduce

$$\left|\int u[\Delta[\boldsymbol{a}], \sigma^{-1}\mathsf{T}_\sigma] f d\widetilde{\mathfrak{m}}\right| \leq \left\|\partial_t \boldsymbol{a}\right\|_{r,s} \left\|u^{-\sigma}\right\|_{L^r_t(\mathbb{V}^r)} \left\|f\right\|_{L^s_t(\mathbb{V}^s)} \leq \left\|\partial_t \boldsymbol{a}\right\|_{r,s} \left\|u\right\|_{L^r_t(\mathbb{V}^r)} \left\|f\right\|_{L^s_t(\mathbb{V}^s)},$$

since T is a contraction on $L^r_t(\mathbb{V}^r)$. In order to apply Lemma 10.15, we actually must show that a similar inequality holds with $\mathsf{R}_t u$ and $\mathsf{R}_\tau f$ in place of u and f respectively, for t, $\tau \in (0,1)$. Clearly, this holds, since the only assumptions on u and f that we used are condition (10.28), which is preserved by R_α. □

Useful consequences are summarized in the following

Corollary 10.17 (convergence of commutators, bounded elliptic case)**.** *Let $\boldsymbol{a} = (\boldsymbol{a}_t)$ be a bounded elliptic form, with $\partial_t \boldsymbol{a}$ of type (r, s). For $p \in \{r, s\}$, let the inequality (10.29) hold.*

Then, for every $f \in L^2 \cap L^s(\widetilde{\mathfrak{m}})$, with $\partial_t f \in L^2(\widetilde{\mathfrak{m}}) + L^{r'}(\widetilde{\mathfrak{m}})$, and $\alpha \in (0,1)$, it holds

$$\partial_t \mathsf{R}_\alpha f, \in L^2(\widetilde{\mathfrak{m}}) + L^{r'}(\widetilde{\mathfrak{m}}), \quad \textit{and} \quad [\mathsf{R}_\alpha, \partial_t] f,$$

where the commutator is defined by (10.25). Moreover, $f \mapsto [\mathsf{R}_\alpha, \partial_t] f$ thus defined extends uniquely to a linear continuous operator mapping $L^2 \cap L^s(\widetilde{\mathfrak{m}})$ into $L^2(\widetilde{\mathfrak{m}}) + L^{r'}(\widetilde{\mathfrak{m}})$ and it holds

$$[\mathsf{R}_\alpha, \partial_t] f \to 0, \quad \textit{in duality with } L^2 \cap L^r(\widetilde{\mathfrak{m}}), \textit{ as } \alpha \downarrow 0, \textit{ for every } f \in L^2 \cap L^s(\widetilde{\mathfrak{m}}).$$

Proof. From (10.30), by density and duality we obtain that the weak commutator operator in (10.26) actually induces a linear continuous operator $[\mathsf{R}_\alpha, \partial_t]$ on $L^2 \cap L^s(\widetilde{\mathfrak{m}})$. Moreover, if $\partial_t f \in L^2(\widetilde{\mathfrak{m}}) + L^{r'}(\widetilde{\mathfrak{m}})$, one argues by duality that

$$\partial_t \mathsf{R}_\alpha f = \mathsf{R}_\alpha \partial_t f - [\mathsf{R}_\alpha, \partial_t] f \in L^2(\widetilde{\mathfrak{m}}) + L^{r'}(\widetilde{\mathfrak{m}}).$$

Finally, to deduce the convergence as $\alpha \downarrow 0$, we notice that, for $f \in L^2 \cap L^s(\widetilde{\mathfrak{m}})$, the functions $[\mathsf{R}_\alpha, \partial_t] f$ are uniformly bounded in $L^2(\widetilde{\mathfrak{m}}) + L^{r'}(\widetilde{\mathfrak{m}})$ as $\alpha \downarrow 0$, and any weak limit must be 0 arguing by duality from (10.26). □

Remark 10.18 (trace semigroup at $t = 0$)**.** Another useful consequence of regularity for $t \mapsto \boldsymbol{a}_t$ is existence of a trace semigroup, e.g. for $t = 0$, defined as follows. By Corollary 10.17, the map associating $f \in L^2 \cap L^s(\widetilde{\mathfrak{m}})$ with $\partial_t f \in L^2(\widetilde{\mathfrak{m}}) + L^{r'}(\widetilde{\mathfrak{m}})$ to $(\mathsf{R}_\alpha f)_0$, i.e., the continuous representative at $t = 0$ (as an element of $L^2(\widetilde{\mathfrak{m}}) + L^{r'}(\widetilde{\mathfrak{m}})$) is continuous. We use this map to induce an operator $\mathsf{R}_{0,\alpha}$ from $L^2 \cap L^s(\mathfrak{m})$ to $L^2(\mathfrak{m}) + L^{r'}(\mathfrak{m})$ as follows: given $\overline{f} \in L^2 \cap L^s(\mathfrak{m})$, we let f be any function with $f \in L^2 \cap L^s(\widetilde{\mathfrak{m}})$, $\partial_t f \in L^2(\widetilde{\mathfrak{m}}) + L^{r'}(\widetilde{\mathfrak{m}})$, $f_0 = \overline{f}$, and define $\mathsf{R}_{0,\alpha} \overline{f} := (\mathsf{R}_\alpha f)_0$. It turns out that $\mathsf{R}_{0,\alpha}$ is well-defined and provides a Markov semigroup $(\mathsf{R}_{0,\alpha})_\alpha$. ■

10.3 Uniqueness results

We have all the technical tools to state and prove our uniqueness results for Fokker-Planck equations, following the smoothing scheme from Section 10.1 and relying on the commutator estimates from the previous section.

Let us explicitly remark that we introduce assumptions on the diffusion operator $\mathcal{L}$, e.g., bounds on its divergence, some ellipticity or Sobolev regularity, on $\mathscr{A}$, e.g., density in suitable function spaces, and on the geometry of the space, e.g., the validity of L^p-Γ inequalities. When dealing with explicit examples, one still has to prove whether these conditions met: currently, the largest "abstract" class for which the theory is non-empty, in particular in the deterministic case, is that of RCD metric measure spaces, as investigated in Ambrosio and Trevisan [2014].

By the results in the previous section, we have at disposal three commutator estimates, namely, for Sobolev derivations, diffusion operators and bounded elliptic 2-tensors. Our aim is to deduce correspondingly three uniqueness results, although variants can be devised. As in the previous section, we let throughout $q \in (1, \infty]$, r, $s \in (1, \infty]$ satisfy $q^{-1} + r^{-1} + s^{-1} = 1$.

10.3.1 Back to the approximation scheme

Before we address the concrete situations, we provide a description of the technical points that we must face, in a more accurate way than in Section (10.1). We consider solutions u belonging to some dual Banach space of functions C', e.g. $C^* = L^\infty_t(L^r_x)$. For simplicity, let us also assume that $H = L^2_t(L^2_x) \subseteq C$, so that are in a Gelfand triple setting $C^* \subseteq H^* = H \subseteq C$, and look for the energy inequality entailing bounds in $L^\infty_t(L^2_x)$.

Step 1 (extension of the weak formulation)**.** Intuitively, if $\mathscr{A}$ is too small, we cannot expect uniqueness to hold. to this aim, we extend the validity of the weak formulation for the FPE, from duality with $\mathscr{A}$ to some larger space A.

Step 2 (stability with respect to R)**.** Since our aim is to use $\mathsf{R}_\alpha u$ as a test function, we require

$$\mathsf{R}_\alpha A \subseteq A, \quad \mathsf{R}_\alpha C \subseteq C \quad \text{and} \quad [\mathsf{R}_\alpha, \partial_t + \mathcal{L}] : A \to C.$$

At this point, we deduce that u^α is a solution to

$$\partial_t u^\alpha = \mathcal{L}^* u^\alpha + [\mathsf{R}_\alpha, \partial_t + \mathcal{L}]^* u, \quad u_0^\alpha = \mathsf{R}_{0,\alpha}\overline{u},$$

since for every $f \in A$, it holds

$$\begin{aligned}\int [(\partial_t + \mathcal{L})f]\, u^\alpha d\widetilde{\mathfrak{m}} &= \int u[\mathsf{R}_\alpha, \partial_t + \mathcal{L}]f\, d\widetilde{\mathfrak{m}} + [(\partial_t + \mathcal{L})\mathsf{R}_\alpha f]\, u d\widetilde{\mathfrak{m}} \\ &= \int f[\mathsf{R}_\alpha, \partial_t + \mathcal{L}]^* u d\widetilde{\mathfrak{m}} - \int (\mathsf{R}_\alpha f)_0 \overline{u} d\mathfrak{m} \\ &= \int f[\mathsf{R}_\alpha, \partial_t + \mathcal{L}]^* u d\widetilde{\mathfrak{m}} - \int f_0 \mathsf{R}_{0,\alpha}\overline{u} d\mathfrak{m},\end{aligned}$$

where we use the fact that u is a solution to the FPE associated to $\mathcal{L}$. Notice that the initial condition $u_0 = \overline{u}$ is replaced with $u_0^\alpha = \mathsf{R}_{0,\alpha}\overline{u}$, where $\mathsf{R}_{0,\alpha}$ is the trace semigroup at 0, defined in Remark 10.18.

Step 3 (smoothing action of R)**.** We require that $\mathcal{L}^*\mathsf{R}_\alpha : C^* \mapsto C$, thus the right hand side in (10.2) belongs to C. As a consequence, the equation gives $\partial_t u^\alpha \in C$. Thus, u^α admits an continuous representative $t \mapsto \tilde{u}^\alpha(t) \in H$, with $u_0^\alpha = \mathsf{R}_{0,\alpha}\overline{u}$.

Step 4 (approximate energy inequality)**.** The curve $t \mapsto \int |u_t^\alpha|^2 \, d\mathfrak{m}$ is absolutely continuous, and we bound from above its derivative with

$$\partial_t \int |u_t^\alpha|^2 \, d\mathfrak{m} \le \left\| \operatorname{div} \mathcal{L}^- \right\|_\infty \int |u_t^\alpha|^2 \, d\mathfrak{m} + \int u_t [\mathsf{R}_\alpha, \partial_t + \mathcal{L}_t] u_t^\alpha d\mathfrak{m}, \quad \mathscr{L}^1\text{-a.e. } t \in (0,T).$$

By Gronwall lemma, we obtain the energy inequality

$$\|u^\alpha\|_{L_t^\infty(L_x^2)} \le \exp\left\{ \left\| \operatorname{div} \mathcal{L}^- \right\|_{L_t^1(L_x^\infty)} \right\} \left(\|\mathsf{R}_{0,\alpha}\overline{u}\|_2 + \rho(\alpha) \right),$$

with

$$\rho(\alpha) := \int_0^T \left| \int u_t [\mathsf{R}_\alpha, \partial_t + \mathcal{L}_t] u_t^\alpha d\mathfrak{m} \right| dt.$$

Step 5 (limit as $\alpha \downarrow 0$)**.** We prove that

$$\|u^\alpha\|_{L_t^\infty(L_x^2)} \to \|u^\alpha\|_{L_t^\infty(L_x^2)}, \quad \|\mathsf{R}_{0,\alpha}\overline{u}\|_2 \to \|\overline{u}\|_2, \quad \text{and} \quad \rho(\alpha) \to 0, \quad \text{as } \alpha \downarrow 0.$$

We deduce that

$$\|u\|_{L_t^\infty(L_x^2)} \le \exp\left\{ \left\| \operatorname{div} \mathcal{L}^- \right\|_{L_t^1(L_x^\infty)} \right\} \|\overline{u}\|_2,$$

and in particular uniqueness holds, letting u be the difference between any two solutions, so that it solves the FPE with $\overline{u} = 0$.

10.3.2 Case of Sobolev derivations

In this section, we essentially prove [Ambrosio and Trevisan, 2014, Theorem 5.4], entailing uniqueness for the FPE in the deterministic case, i.e., the continuity equation, when $\mathcal{L} := \boldsymbol{b}$ is a derivation with some bound on its deformation and its divergence.

Theorem 10.19 (uniqueness of solutions, Sobolev derivations)**.** *Let $\mathscr{A}$ be dense in $L^\infty_t(\mathbb{V}^p)$ in the following sense: for every $f \in L^\infty_t(\mathbb{V}^p)$ there exists $(f^k)_k \in \mathscr{A}$ with $f^k \to f$ and $\|f^k_t - f_t\|_{\mathbb{V}^s} \to 0$ weakly-* in $L^\infty(0,T)$, and let the L^p-Γ inequality hold, for $p \in \{r,s\}$.*

Let $\mathcal{L} := \boldsymbol{b} = (\boldsymbol{b}_t)_{t\in(0,T)}$ be a Borel family of derivations, with

$$|\boldsymbol{b}|, \operatorname{div} \boldsymbol{b} \in L^1_t(L^q_x), \quad \|D^{sym}\boldsymbol{b}_t\|_{r,s} \in L^1(0,T) \quad \textit{and} \quad \operatorname{div} \boldsymbol{b}^- \in L^1_t(L^\infty_x).$$

Assume that there exists $(f_n) \subset \mathscr{A}$ as in (9.2)*.*

Then, there exists at most one weakly- continuous solution u in $L^\infty_t(L^2_x \cap L^r_x)$ to the FPE*

$$\partial_t u_t + \operatorname{div}(u_t \boldsymbol{b}_t) = 0, \quad \textit{in } (0,T)\times X, \textit{ with } u_0 = \overline{u},$$

for every initial condition $\overline{u} \in L^2 \cap L^r(\mathfrak{m})$.

Proof. In the scheme described above we let

$$C = L^1(L^2_x + L^{r'}_x), \quad C' = L^\infty_t(L^2_x \cap L^r_x),$$

and $\mathsf{R} := \tilde{\mathsf{P}}$, i.e., we act with P on each fiber $\{t\}\times X$ for $t \in (0,T)$. Clearly, it holds $\mathsf{R}_{0,\alpha} = \mathsf{P}_\alpha$.

Step 1 (extension of the weak formulation)**.** By density of $\mathscr{A}$, the weak formulation extends from duality with $f \in \mathscr{A}$, to duality with $f \in A := W^{1,1}_t(L^2_x + L^s_x) \cap L^\infty(\mathbb{V}^s)$.

Step 2 (stability with respect to R)**.** Since the semigroup is constant with respect to t, the commutator $[\mathsf{P}_\alpha, \partial_t]$ is trivial and, by the L^s-Γ inequality, we deduce $\mathsf{P}_\alpha A \subseteq A$. Moreover, the space $C = L^1_t(L^2_x + L^{r'}_x)$ is stable with respect to the action of $\tilde{\mathsf{P}}$, and $[\mathsf{P}_\alpha, \partial_t + \boldsymbol{b}]f$ belongs to C, by Corollary 10.9.

Step 3 (smoothing action of R)**.** This is a consequence of the validity of the L^r-Γ inequality, as $\mathsf{P}_\alpha u \in L^\infty_t(\mathbb{V}^r)$, thus

$$\mathcal{L}^*\mathsf{P}_\alpha u = (\operatorname{div} u^\alpha \boldsymbol{b}) = (\operatorname{div} \boldsymbol{b})u^\alpha + du^\alpha(\boldsymbol{b}) \in L^1_t(L^2_x + L^{s'}_x).$$

Step 4 (approximate energy inequality)**.** Notice that from the identity above we deduce only $\partial_t u \in L^1_t(L^2_x + L^{s'}_x)$: this forces us to take a small departure from the argument in the previous section, relying instead on the sequence $(f_n) \subset \mathscr{A}$ satisfying (9.2). Starting from $|z|^{1+r/s}$, we let

$$\beta(z) := \begin{cases} 1 + \frac{r+s}{s}(z-1) & \text{if } z > 1 \\ |z|^{1+r/s} & \text{if } |z| \le 1, \\ 1 - \frac{r+s}{s}(z+1) & \text{if } z < -1, \end{cases}$$

so that $g_\beta(z) := z\beta'(z) - \beta(z) \le (r/s)\beta(z)$ for $z \in \mathbb{R}$ (in the points where β' does not exists, we choose the larger among β'_+ and β'_-). Moreover, β has linear growth at infinity, thus Lemma 5.9 and its subsequent remark entail that, for $n \ge 1$, the curve $t \mapsto \int f_n \beta(u^\alpha_t) d\mathfrak{m}$ is absolutely continuous, with

$$\frac{d}{dt}\int f_n\beta(u^\alpha_t)d\mathfrak{m} = \int u[\mathsf{P}_\alpha, \boldsymbol{b}_t]\left(f_n\beta'(u^\alpha)\right)d\mathfrak{m} - \int f_n \operatorname{div}(\beta(u^\alpha_t)\boldsymbol{b}_t) + f_n g_\beta(u^\alpha_t)\operatorname{div} \boldsymbol{b}_t d\mathfrak{m},$$

for $\mathscr{L}^1$-a.e. $t \in (0,T)$. Hence, denoting $d(t) := (r/s)\| \operatorname{div} \boldsymbol{b}_t^- \|_\infty$ we use the inequality $g_\beta \leq (r/s)\beta$ to get

$$\frac{d}{dt}\int f_n \beta(u_t^\alpha) d\mathfrak{m} \leq d(t) \int f_n \beta(u_t^\alpha) d\mathfrak{m} + \int u[\mathsf{P}_\alpha, \boldsymbol{b}_t]\left(f_n \beta'(u_t^\alpha)\right) d\mathfrak{m} + \int \beta(u_t^\alpha) d f_n(\boldsymbol{b}_t) d\mathfrak{m},$$

which, by Gronwall inequality, leads to

$$\sup_{t\in[0,T]} \int f_n \beta(u_t^\alpha) d\mathfrak{m} \leq \exp\left\{\|d\|_{L_t^1}\right\}\left[\int \beta(\mathsf{P}_\alpha \overline{u}) d\mathfrak{m} + \rho(\alpha)\right],$$

with

$$\rho(\alpha) := \int_0^T \left|\int u[\mathsf{P}_\alpha, \boldsymbol{b}_t]\left(f_n \beta'(u_t^\alpha)\right) d\mathfrak{m}\right| + \left|\int \beta(u_t^\alpha) d f_n(\boldsymbol{b}_t) d\mathfrak{m}\right| dt.$$

Step 5 (limit as $\alpha \downarrow 0$)**.** Arguing at fixed $n \geq 1$, we let $\alpha \downarrow 0$, noticing that $u^\alpha \to u$ $\widetilde{\mathfrak{m}}$-pointwise and bounded in $L_t^\infty(L_x^2 \cap L_x^r)$, $\beta(u_t^\alpha) \to \beta(u_t)$ in $L^2 + L^{q'}(\mathfrak{m})$ for a.e. $t \in (0,T)$, and $\beta'(u_t^\alpha) \to \beta'(u)$ in $L^2 \cap L^s(\mathfrak{m})$, with uniform bounds in $t \in (0,T)$.

By Corollary 10.9, we conclude that

$$\int u_t[\mathsf{R}_\alpha, \boldsymbol{b}_t] f_n \beta'(u_t^\alpha) d\mathfrak{m} \to 0, \quad \text{in } L^1(0,T), \text{ as } \alpha \downarrow 0.$$

Letting finally $n \to \infty$, by monotone convergence theorem, we finally deduce

$$\sup_{t\in[0,T]} \int \beta(u_t) d\mathfrak{m} \leq \exp\left\{\|d\|_{L_t^1}\right\} \int \beta(\overline{u}) d\mathfrak{m},$$

that leads to uniqueness. □

10.3.3 Case of Sobolev diffusions

In this section, we prove uniqueness for Fokker-Planck equations with a diffusion operator of the form

$$\mathcal{L} f := a\Delta f + df(\boldsymbol{b}). \tag{10.31}$$

where $a \geq 0$ can be degenerate, but we require some Sobolev regularity up to second order (recall also the notation (10.16) for the Hessian $H[a]$). Notice that the apparently involute assumptions become much clearer in the case $q = \infty$, $r = s = 2$.

Theorem 10.20 (uniqueness of solutions, Sobolev diffusions)**.** *Let $\mathscr{A}$ be dense in $W_t^{1,1}(L_x^2 + L_x^{r'} + L_x^{s'}) \cap L_t^\infty(\mathbb{V}^s \cap D^s(\Delta))$, in the following sense: for every f in such a space, there exists $(f^k)_k \in \mathscr{A}$ with $f^k \to f$ in $W_t^{1,1}(L_x^2 + L_x^{r'} + L_x^{s'})$ and $\left\|f_t^k - f_t\right\|_{\mathbb{V}^s \cap D^s(\Delta)} \to 0$, weakly-$*$ in $L^\infty(0,T)$.*

Let $\boldsymbol{b} = (\boldsymbol{b}_t)_{t\in(0,T)}$ be a Borel family of derivations, with

$$|\boldsymbol{b}|, \operatorname{div} \boldsymbol{b} \in L_t^1(L_x^q) \quad \text{and} \quad \|D^{sym}\boldsymbol{b}_t\|_{r,s} \in L^1(0,T)$$

and let $a \in L_t^1(\mathbb{V}^q \cap D^q(\Delta))$, with $\|H[a_t]\|_{r,s} \in L^1(0,T)$. Define $\mathcal{L}$ by (10.31) and let

$$\operatorname{div} \mathcal{L}^- = (\operatorname{div} \boldsymbol{b} - \Delta a)^- \in L_t^1(L_x^\infty).$$

Then, for every $\overline{u} \in L^2 \cap L^r \cap L^s(\mathfrak{m})$, there exists at most one weakly continuous solution u in $L_t^\infty(L_x^2 \cap L_x^r \cap L_x^s)$ to the FPE

$$\partial_t u = \mathcal{L}^* u, \quad \text{in } (0,T) \times X, \text{ with } u_0 = \overline{u}.$$

Proof. We closely follow the scheme described in Section 10.1, letting

$$C = L^1_t(L^2_x \cap L^r_x \cap L^s_x), \quad C' = L^\infty_t(L^2_x + L^r_x + L^s_x)$$

and the approximation provided by the semigroup $\mathsf{R} := \tilde{\mathsf{P}}$.

Step 1 (extension of the weak formulation)**.** By density of $\mathscr{A}$, we extend the validity of the weak formulation from duality with $f \in \mathscr{A}$, to duality with $f \in A := W^{1,1}_t(L^2_x + L^s_x) \cap L^\infty_t(\mathbb{V}^s \cap D^s(\Delta))$.

Step 2 (stability with respect to R)**.** As in the previous section, the commutator $[\mathsf{P}_\alpha, \partial_t]$ is null and, since the L^s-Γ inequality holds, it holds $\mathsf{P}_\alpha A \subseteq A$. Also the space C is clearly stable with respect to the action of P_α, being a Markov semigroup, while the commutator $[\mathsf{P}_\alpha, \partial_t + \mathcal{L}]f$ belongs to C, by Corollary 10.13.

Step 3 (smoothing action of R)**.** This is a consequence of the L^r-Γ inequality and the L^r-Δinequality: indeed $\mathsf{P}_\alpha u \in L^\infty_t(\mathbb{V}^r \cap D^r(\Delta))$, thus

$$\mathcal{L}^*\mathsf{P}_\alpha u = \operatorname{div} u^\alpha \boldsymbol{b} + \Delta(a u^\alpha)) = (\operatorname{div} \boldsymbol{b}) u^\alpha + du^\alpha(\boldsymbol{b}) + (\Delta a) u^\alpha + a \Delta u^\alpha + 2\Gamma(a, u^\alpha) \in L^1_t(L^{s'}_x),$$

thus we deduce $u \in W^{1,1}_t(L^2_x + L^{r'}_x + L^{s'}_x)$.

Step 4 (approximate energy inequality)**.** Arguing precisely as in Section 10.1, we obtain the energy inequality, entailing $u^\alpha \in L^\infty_t(L^2_x)$, with an error term given by

$$\rho(\alpha) := \int_0^T \left| \int u_t [\mathsf{P}_\alpha, \mathcal{L}_t] u^\alpha_t \, d\mathfrak{m} \right| dt.$$

Step 5 (limit as $\alpha \downarrow 0$)**.** To conclude, we non-trivial point is that $\rho(\alpha) \to 0$. Indeed, by linearity we can split the commutator in two terms, one corresponding to $[\mathsf{P}_\alpha, \boldsymbol{b}_t]$ and the other $[\mathsf{P}_\alpha, a_t \Delta]$. The first one is infinitesimal by Corollary 10.9, while for the second one we rely on last statement of Corollary 10.13.

□

10.3.4 Case of bounded elliptic diffusion operators

In this section, we establish uniqueness for Fokker-Planck equations associated to a diffusion operator in the form

$$\mathcal{L} f := \Delta[\boldsymbol{a}] f + df(\boldsymbol{b}), \tag{10.32}$$

where $\boldsymbol{a}$ is a bounded elliptic form. Recall that in Chapter 9 we show existence as well as uniqueness for solutions $u \in L^2_t(\mathbb{V})$: our aim is to prove here uniqueness in a larger space, assuming only integral bounds. For simplicity of exposition, we restrict ourselves to the case $q = \infty$, $r = s = 2$, which corresponds to that considered in [Figalli, 2008, Theorem 4.3], although one could use the more general commutator estimates from Section 10.2.4 to deal with less regularity for $t \mapsto \boldsymbol{a}_t$, at the price of stronger assumptions on integrability of solutions and the smoothing action for $\mathsf{P}[\boldsymbol{a}]$.

Theorem 10.21 (uniqueness, bounded elliptic case)**.** *Let $\boldsymbol{a}$ be a bounded elliptic 2-tensor, $\boldsymbol{b}$ be a derivation with $|\boldsymbol{b}| \in L^\infty_t(L^\infty_x)$, let $\mathscr{A}$ be dense in $W^{1,2}_t(L^2_x) \cap D(\Delta[\boldsymbol{a}])$ and $\partial_t \boldsymbol{a}$ be of type $(2,2)$.*

Then, for every $\overline{u} \in L^2_x$, there exists at most one weakly continuous solution u in $L^\infty_t(L^2_x)$ to the FPE (10.32), with $u_0 = \overline{u}$.

Proof. We follow the smoothing scheme introduced in Section 10.1, with $C = L^2_t(\mathbb{V}^*)$, proving $u \in L^2_t(\mathbb{V}) =: C'$, so that uniqueness then holds by Theorem 9.1. We consider the approximation provided by the semigroup $\mathsf{R} := \mathsf{P}[\boldsymbol{a}]$.

Step 1 (extension of the weak formulation)**.** By density, we extend the weak formulation to $f \in A := W^{1,2}_t(L^2_x) \cap D(\Delta^{\boldsymbol{a}})$.

Step 2 (stability with respect to R)**.** By Corollary 10.17, it holds $\mathsf{R}_\alpha A \subseteq A$; C is also stable with respect to R_α and finally the commutator

$$[\mathsf{R}_\alpha, \partial_t + \Delta[\boldsymbol{a}] + \boldsymbol{b}]f = [\mathsf{R}_\alpha, \partial_t + \boldsymbol{b}]f \in C,$$

for every $f \in A$.

Step 3 (smoothing action of R)**.** This is a consequence of the smoothing effect of the heat semigroup associated to a Dirichlet form (recall Section 3.1.2).

Step 4 (approximate energy inequality)**.** Our aim is to prove, $u^\alpha \in L^2_t(\mathbb{V})$, uniformly in $\alpha \in (0,1)$: arguing as in Section 8.2, we deduce

$$2\int \boldsymbol{a}(u^\alpha) d\widetilde{\mathfrak{m}} \leq \int (\mathsf{R}_{0,\alpha}\overline{u})^2 d\mathfrak{m} - \int u[\mathsf{R}_\alpha, \partial_t + \boldsymbol{b}]u^\alpha d\widetilde{\mathfrak{m}} + \int u^\alpha du^\alpha(\boldsymbol{b}) d\widetilde{\mathfrak{m}}.$$

We notice that the terms where $\boldsymbol{b}$ appears simplify to

$$\left| \int u du^{2\alpha}(\boldsymbol{b}) d\widetilde{\mathfrak{m}} \right| \leq \|u\,|\boldsymbol{b}|\|_2 \, \|\sqrt{\Gamma\,(u^{2\alpha})}\|_2.$$

Up to some constant $C \geq 0$, depending on the ellipticity of $\boldsymbol{a}$ only, the term $\|\sqrt{\Gamma\,(u^{2\alpha})}\|_2$ is smaller than $\|\sqrt{\boldsymbol{a}(u^{2\alpha})}\|_2$, which is a decreasing function of α, thus we estimate,

$$\left| \int u d(u^{2\alpha})(\boldsymbol{b}) d\widetilde{\mathfrak{m}} \right| \leq C\, \|u\,|\boldsymbol{b}|\|_2 \, \|\sqrt{\Gamma\,(u)}\|_2,$$

where C is some constant depending on the ellipticity of $\boldsymbol{a}$ and $\|\boldsymbol{a}\|_\infty$. To conclude, we split $2xy \leq \varepsilon^{-1}x^2 + \varepsilon y$, for ε small enough, depending on the ellipticity of $\boldsymbol{a}$, to obtain

$$\int \Gamma(u^\alpha) d\widetilde{\mathfrak{m}} \leq C\left[\|\overline{u}\|_2 + C\,\|u\|_2^2\, \||\boldsymbol{b}|\|_\infty^2 + \rho(\alpha)\right]$$

where

$$\rho(\alpha) := \left| \int u[\mathsf{R}_\alpha, \partial_t] u^\alpha d\widetilde{\mathfrak{m}} \right|.$$

Step 5 (limit as $\alpha \downarrow 0$)**.** The only non-trivial term in the limit is $c(\alpha)$, whose convergence to 0 is established in Corollary 10.17 (actually, it is sufficient to show that it is uniformly bounded). □

Let us remark that the argument above holds as long as the product $u\,|\boldsymbol{b}|$ is bounded in $L^t(\mathfrak{m})$: aiming for uniqueness in a smaller space, e.g. $L^\infty_t(L^1_x \cap L^\infty_x)$, we may replace the bound with $|\boldsymbol{b}| \in L^2_t(L^2_x + L^\infty_x)$. Moreover, stronger results can be proved if Sobolev inequalities hold, relying on the energy estimates established in Section 8.3.

Part IV

Examples

Chapter 11

Finite dimensional spaces

In this chapter, on one side we illustrate relevant classes of finite-dimensional spaces for which our theory applies. On the other side, we compare our results on well-posedness of Fokker-Planck equations and martingale problems with some of those available in the current literature. We largely focus on the Euclidean setting, from Section 11.1 to Section 11.2. In Section 11.3, we briefly describe how weighted Riemannian and even sub-Riemannian (Section 11.3) structures also fit in our general framework.

11.1 The Euclidean setting

We show how the theory developed for general metric measure spaces specializes in the Euclidean setting of $\mathbb{R}^d$, providing explicit descriptions of the objects involved.

11.1.1 Dirichlet form setup

We consider a specialization of the framework described in Chapter 3, letting $X = \mathbb{R}^d$, $\mathfrak{m} = \mathscr{L}^d$ be the Lebesgue measure and

$$\mathcal{E}(f) := \int |\nabla f|^2 (x) d\mathscr{L}^d(x), \quad \text{for } f \in W^{1,2}(\mathbb{R}^d), \quad \mathcal{E}(f) = \infty \text{ otherwise},$$

where we recall that $W^{1,2}(\mathbb{R}^d)$ is the usual Sobolev space of functions $f \in L^2(\mathbb{R}^d)$ whose distributional derivative ∇f is represented by an element in $L^2(\mathbb{R}^d; \mathbb{R}^d)$ (here and in what follows, we write $L^p(\mathbb{R}^d)$ in place of $L^p(\mathbb{R}^d, \mathscr{L}^d)$, for $p \in [1, \infty])$. More explicitly, a function $f \in L^2(\mathbb{R}^d)$ belongs to $W^{1,2}(\mathbb{R}^d)$ if and only if, for every $i \in \{1, \ldots, d\}$, there exists $g_i \in L^2(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} f(x) \frac{\partial \varphi}{\partial x^i}(x) dx = -\int_{\mathbb{R}^d} g_i(x) \varphi(x) dx, \quad \text{for every } \varphi \in C_c^\infty(\mathbb{R}^d).$$

We let then $g_i = \partial_i f$ and $\nabla f := (\partial_i f)_{i=1}^d \in L^2(\mathbb{R}^d; \mathbb{R}^d)$. We then introduce the (squared) modulus of the gradient $|\nabla f|^2 := \sum_{i=1}^d (\partial_i f)^2 \in L^1(\mathbb{R}^d)$ and it is easy to prove that $W^{1,2}(\mathbb{R}^d)$, endowed with the norm $\|f\|_{W^{1,2}}^2 := \|f\|_2 + \|\nabla f\|_2^2$ is a Hilbert space, thus, $(\mathcal{E}, W^{1,2}(\mathbb{R}^n))$ defines a closed quadratic form. Equivalently, the lower semicontinuity of $\mathcal{E}$ with respect to convergence in $L^2(\mathbb{R}^d)$ follows from the identity

$$\mathcal{E}(f) := \sup \left\{ \int_{\mathbb{R}^d} f \operatorname{div} \Phi(x) dx \, : \, \Phi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d), \text{ with } \sum_{i=1}^d \int \Phi_i(x)^2 dx \leq 1 \right\},$$

noticing that the functionals in the right hand side above are continuous with respect to convergence in $L^2(\mathbb{R}^d)$. To show the Markov property, i.e. that normal contractions η operate on $\mathcal{E}$ (3.2) thus $\mathcal{E}$ is Dirichlet, we rely on the density of test functions $C_c^\infty(\mathbb{R}^d)$ in $W^{1,2}(\mathbb{R}^d)$, by the usual Meyers-Serrin theorem.

The Laplacian Δ in the theory of Dirichlet forms coincides indeed the usual distributional Laplacian, with $D(\Delta)$ given by the functions $f \in W^{1,2}(\mathbb{R}^d)$, such that the distributional divergence of ∇f belongs to $L^2(\mathbb{R}^d)$, i.e. there exists $g \in L^2(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} g(x)\varphi(x)dx = -\int_{\mathbb{R}^d} \langle \nabla f(x), \nabla \varphi(x)\rangle \, dx, \quad \text{for every } \varphi \in C_c^\infty(\mathbb{R}^d).$$

Notice that, also in this case, we use the density of test functions $C_c^\infty(\mathbb{R}^d)$ in $W^{1,2}(\mathbb{R}^d)$, since the abstract definition of Laplacian would require $\varphi \in W^{1,2}(\mathbb{R}^d)$ above.

The semigroup $(\mathsf{P}_t)_{t\geq 0}$ corresponds to the transition semigroup of a Brownian motion, rescaled by a factor $\sqrt{2}$, for which we have the following representation formula:

$$\mathsf{P}_t f(x) = \int_{\mathbb{R}^d} f(x+\sqrt{2t}y)\frac{e^{-|y|^2/2}}{\sqrt{(2\pi)^n}}dy, \qquad \text{for } x \in \mathbb{R}^d. \tag{11.1}$$

We let in what follows $\rho(y) := e^{-|y|^2/2}/\sqrt{(2\pi)^n}$, for $y \in \mathbb{R}^d$, be the standard Gaussian kernel in $\mathbb{R}^d$.

The carré du champ is given by $\Gamma(f) = |\nabla f|^2$, for $f \in W^{1,2}(\mathbb{R}^d)$; the spaces $\mathbb{V}^p$ and $D(\Delta)$ can be identified respectively as

$$\mathbb{V}^p = W^{1,p} \cap W^{1,2}(\mathbb{R}^d) = \left\{ f \in L^p \cap L^2(\mathbb{R}^d) \,:\, \nabla f \in L^2 \cap L^p(\mathbb{R}^d;\mathbb{R}^d) \right\}$$

$$D^p(\Delta) = \left\{ f \in D(\Delta) \cap L^p(\mathbb{R}^d) \,:\, \Delta f \in L^p(\mathbb{R}^d) \right\},$$

for $p \in [1,\infty]$. Let us notice that, for $p \in (1,\infty)$, it holds $D^p(\Delta) = W^{2,p}(\mathbb{R}^d)$, by the L^p-boundedness of the second order Riesz transform $f \mapsto \nabla^2\Delta^{-1} f$, see e.g. [Gilbarg and Trudinger, 2001].

To prove that the L^p-Γ inequality holds, for $p \in [1,\infty]$ we argue first at fixed $x \in \mathbb{R}^d$ and integrate by parts in (11.1), thus

$$\nabla \mathsf{P}_t f(x) = \frac{1}{\sqrt{2t}}\int_{\mathbb{R}^d} \nabla_y f(x+\sqrt{2t}y)\rho(y)dy = \frac{1}{\sqrt{2t}}\int_{\mathbb{R}^d} f(x+\sqrt{2t}y)y\rho(y)dy. \tag{11.2}$$

Hölder inequality gives, for $p \in (1,\infty]$,

$$|\nabla \mathsf{P}_t f(x)| \leq \frac{1}{\sqrt{2t}}\left[\int_{\mathbb{R}^d} \left|f(x+\sqrt{2t}y)\right|^p \rho(y)dy\right]^{1/p} \left[\int_{\mathbb{R}^d} |y|^{p'} \rho(y)dy\right]^{1/p'},$$

and integration over $x \in \mathbb{R}^d$ entails, by Fubini theorem and translational invariance of $\mathscr{L}^d$, the inequality

$$\|\nabla \mathsf{P}_t f\|_p \leq \frac{c_p^\Gamma}{\sqrt{t}}\|f\|_p \quad \text{for every } t \in (0,\infty), \tag{11.3}$$

with $(c_p^\Gamma)^{p'} = 2^{p/(2p-2)}\int |y|^{p'}\rho(y)dy < \infty$. To prove that the L^1-Γ inequality holds, we argue by duality or directly integrate in (11.2), obtaining

$$\int_{\mathbb{R}^d} |\nabla \mathsf{P}_t f(x)|\,dx \leq \frac{1}{\sqrt{2t}}\int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d}\left|f(x+\sqrt{2t}y)\right|dx\right]|y|\,\rho(y)dy = \frac{c_1^\Gamma}{\sqrt{t}}\|f\|_1,$$

with $c_1^\Gamma = 2^{-1}\int |y|\,\rho(y)dy$.

Similarly, we prove directly the validity of L^p-Δ inequalities, acting with the Laplacian on the Gaussian kernel at fixed $x \in \mathbb{R}^d$,

$$\Delta \mathsf{P}_t f(x) = \frac{1}{2t}\int_{\mathbb{R}^d} \Delta_y f(x+\sqrt{2t}y)\rho(y)dy = \frac{1}{2t}\int_{\mathbb{R}^d} f(x+\sqrt{2t}y)\left(|y|^2 - d\right)\rho(y)dy,$$

so that by Hölder inequality, as above, we obtain

$$\|\Delta \mathsf{P}_t f\|_p \le \frac{c_p^\Delta}{t}\|f\|_p \quad \text{for every } t\in(0,\infty),$$

with $(c_p^\Delta)^{p'} = 2^{-p/(p-1)}\int \big|\,|y|^2 - d\big|^{p'}\rho(y)dy$, for $p\in(1,\infty]$. When $p=1$, we directly integrate and obtain $c_1^\delta = \int \big|\,|y|^2 - d\big|\rho(y)dy$. Recall that the validity of L^p-Δ inequality for the endpoints $p=1$ or $p=\infty$ is not guaranteed by the general results in Chapter 3.

Finally, we notice that similar results can be proved for higher order derivatives, entailing the bound

$$\|\nabla^k \mathsf{P}_t f\|_p \le \frac{c_{p,k}}{t^{k/2}}\|f\|_p \quad \text{for every } t\in(0,\infty), \tag{11.4}$$

where $c_{p,k}$ is some constant depending, besides $p\in[1,\infty]$ and $k\ge 1$, on the dimension d.

11.1.2 Diffusion operators

In the time-independent framework, a natural choice for the algebra $\mathscr{A}$ introduced Section 4.1 is that of test functions $C_c^\infty(\mathbb{R}^d)$.

Vector fields as derivations

For a Borel vector field $b = (b^i)_{i=1}^d \in L^1_{loc}(\mathbb{R}^d;\mathbb{R}^d)$, the associated derivation $\boldsymbol{b}$ is naturally defined by

$$\mathscr{A} \ni f \quad \mapsto \quad df(\boldsymbol{b}) := b\cdot\nabla f = \sum_{i=1}^d b^i \frac{\partial f}{\partial x^i} \in L^1(\mathbb{R}^d).$$

Then, quite obviously, $\operatorname{div}\boldsymbol{b}$ is the usual distributional divergence and, as already noticed in Remark 10.6, the "abstract" deformation $D^{sym}\boldsymbol{b}$ in Definition 10.5, namely

$$\int D^{sym}\boldsymbol{b}(u,f)d\mathfrak{m} := -\frac{1}{2}\int \left[df(\boldsymbol{b})\Delta u + du(\boldsymbol{b})\Delta f - (\operatorname{div}\boldsymbol{b})\Gamma(u,f)\right]d\mathfrak{m},$$

corresponds to the symmetric part of the distributional derivative of $\boldsymbol{b}$, by integration over $\mathbb{R}^d$ of the identity

$$\nabla u\cdot\nabla(b\cdot\nabla f) + \nabla f\cdot\nabla(b\cdot\nabla u) - b\cdot\nabla(\nabla u\cdot\nabla f) = \langle Db\,\nabla f, \nabla u\rangle + \langle Db\,\nabla u, \nabla f\rangle\,.$$

2-tensors

Given a Borel map $a = (a^{i,j})_{i,j=1}^d \in L^1_{loc}(\mathbb{R}^d;\mathbb{R}^{d\times d})$, its associated 2-tensor is given by

$$\mathscr{A}\times\mathscr{A} \ni (f,g) \quad \mapsto \quad \boldsymbol{a}(f,g) := a : (\nabla f\otimes\nabla g) = \sum_{i,j=1}^d a^{i,j}\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial x^j}.$$

Clearly, symmetry and elliptic bounds on $\boldsymbol{a}$ are consequences (actually equivalent) to symmetry and elliptic bounds for the matrix $a(x)$, $\mathfrak{m}$-a.e. $x\in\mathbb{R}^d$.

Diffusion operators

Finally, given a Borel vector field $b \in L^1_{loc}(\mathbb{R}^d;\mathbb{R}^d)$ and a Borel map $a \in L^1_{loc}(\mathbb{R}^d;\mathbb{R}^{d\times d})$ with values in symmetric, non-negative matrices, we introduce the diffusion operator $\mathcal{L} := \mathcal{L}(a,b)$,

$$\mathscr{A} \ni f \quad \mapsto \quad \mathcal{L}f := a : \nabla^2 f + b\cdot\nabla f = \sum_{i,j=1}^d a^{i,j}\frac{\partial^2 f}{\partial x^i \partial x^j} + \sum_{i=1}^d b^i \frac{\partial^i f}{\partial x^i} \in L^1(\mathbb{R}^d). \tag{11.5}$$

Notice that, differently from Part I, we always consider $\mathcal{L}$ as taking values in $\mathscr{L}^d$-equivalence classes; this causes no harm, since at the same time we restrict ourselves to probability measures that are absolutely continuous with respect to $\mathscr{L}^d$.

The diffusion operator $\mathcal{L}$ can written in divergence form whenever the (vector-valued) distributional divergence $\operatorname{div} a$, defined by $(\operatorname{div} a)^i = \sum_{j=1}^d \partial_j a^{i,j}$, for $i \in \{1,\ldots,d\}$ belongs to $L^1_{loc}(\mathbb{R}^d;\mathbb{R}^d)$, so that

$$\mathcal{L}f = \operatorname{div}(a\nabla f) + (b - \operatorname{div} a)\cdot\nabla f, \quad \text{for } f \in \mathscr{A}.$$

Moreover, following Definition 4.21, $\operatorname{div}\mathcal{L}$ reads as the distribution mapping $f \in \mathscr{A}$ into $\int \mathcal{L} f d\mathscr{L}^d$, thus

$$\operatorname{div}\mathcal{L} := \operatorname{div}(b - \operatorname{div} a) = -\sum_{i,j=1}^d \partial_i\partial_j a^{i,j} + \sum_{i=1}^d \partial_i b^i, \quad \text{as a distribution.}$$

11.1.3 FPE's, MP's and flows

In the time-extended setting, we let $\tilde{\mathscr{A}} = C^{1,2}_c((0,T);\mathbb{R}^d)$, and consider Borel families of vector fields $b = (b_t)_t \in L^1_t(L^q(\mathbb{R}^d;\mathbb{R}^d)$, 2-tensors $a = (a_t)_t \in L^1_t(L^q(\mathbb{R}^d;\mathrm{Sym}_+(\mathbb{R}^d))$ and correspondent diffusion operators $\mathcal{L} = (\mathcal{L}_t)_t$ given by $\mathcal{L}_t = \mathcal{L}(a_t,b_t)$, for $q \in [1,\infty]$.

The definitions introduced in Section 6.1 specialize in a straightforward way. As an example, if let $r \in (1,\infty]$, $q \in [r',\infty]$ and $\mathcal{L} = \mathcal{L}(a,b)$ be a time-dependent diffusion, the definition of solution $u \in L^\infty_t(L^r_x)$ to the FPE

$$\partial_t u_t = \mathcal{L}^*_t u_t, \quad \text{on } (0,T)\times\mathbb{R}^d, \tag{11.6}$$

is given in duality with respect to $\tilde{\mathscr{A}}$, i.e., we require

$$\int_0^T \int (\partial_t f(t,x) + \mathcal{L}_t f(x)) u_t(x) dx dt = 0, \quad \text{for every } f \in C^{1,2}_c((0,T);\mathbb{R}^d).$$

Lemma 6.4 gives that every solution to the FPE above admits a weakly-* continuous representative in $L^r(\mathbb{R}^d)$. Moreover, by density, the weak formulation can be extended in duality with $f \in L^\infty_t(W^{2,s}(\mathbb{R}^d))$, where $s \in [1,\infty]$ satisfies $q^{-1} + r^{-1} + s^{-1} = 1$.

One then introduces martingale problems and L^r-regular martingale flows: about the former, we say that $\boldsymbol{\eta} \in \mathscr{P}(C([0,T];\mathbb{R}^d))$ is a L^r-regular solution to the martingale problem associated to $\mathcal{L}$ if, for every $f \in \mathscr{A}$, the process

$$[0,T] \mapsto f\circ e_t - \int_0^T (\partial_t f + \mathcal{L}_s f)\circ e_s ds$$

is a martingale, with respect to the natural filtration on $C([0,T];\mathbb{R}^d)$, and the marginals $\eta_t = (e_t)_\sharp \boldsymbol{\eta}$ are absolutely continuous with respect tot $\mathscr{L}^d$, with densities in $L^\infty_t(L^r(\mathbb{R}^d))$. Martingale flows are defined as selections $(\boldsymbol{\eta}(s,x))_{s,x}$ of probability measures so that, for every $s \in [0,T]$, $\bar{u} \in L^r(\mathbb{R}^d)$, with $\bar{u}\mathscr{L}^d$ probability, the probability measure $\int \boldsymbol{\eta}(s,x)\bar{u}(x)dx$ defines a L^r-regular solution to the martingale problem.

11.1.4 The superposition principle

Not surprisingly, when specialized to the Euclidean framework, the superposition principle gives back the results in Chapter 2, in the special case that $\nu_t = u_t\mathscr{L}^d$ are all absolutely continuous probability densities. Let us sketch how this can be proved. In general, the strategy described in Chapter 7 relies on the choice of a countable set $\mathscr{A}^* \subseteq \mathscr{A}$ and the distance associated to it; a natural choice would be to let $\mathscr{A}^* = \left\{x^i, \ldots x^d\right\}$, but this is not admissible, since we choose to work with compactly supported functions. However, by introducing cut-off functions χ_R, as in Remark 1.3, we may let

$$\mathscr{A}^* = \left\{c_{i,n}x^i(\chi_{n+1} - \chi n) \,:\, i = 1, \ldots, d, \quad n \geq 1\right\}$$

where $c_{i,n}$ are suitable constants, such that the distance associated with $\mathscr{A}^*$ is equivalent the Euclidean distance on $\mathbb{R}^d$.

11.2 Well-posedness results

We are in a position to discuss the specialization of our general existence and uniqueness results for solutions to FPE's and, via superposition principle, MP's and regular martingale flows.

11.2.1 Existence

In the Euclidean setting, the existence results from Chapter 9, providing solutions in $L^\infty_t(L^r_x)$ to FPE's can be strengthened, by means of Proposition 9.5.

Theorem 11.1 (existence of solutions, Euclidean case). *Let $q \in (1,\infty]$ $r \in (1,\infty]$ satisfy $q^{-1} + r^{-1} \leq 1$ and let $\mathcal{L} = \mathcal{L}(a,b)$ be a diffusion operator with coefficients a, $b \in L^1_t(L^{r'}_x)$ and $\operatorname{div}\mathcal{L}^- \in L^1_t(L^\infty_x)$. Then, for every $\bar{u} \in L^r(\mathbb{R}^d)$, there exists a L^r-weakly continuous solution u to the FPE* (11.6)*, which can be built in such a way that*

i) if $\bar{u} \geq 0$, then $u_t \geq 0$, for every $t \in [0,T]$, and

ii) if, for some $p \in [1,\infty]$, $\bar{u} \in L^p(\mathbb{R}^d)$, then $u \in L^\infty_t(L^p(\mathbb{R}^d))$, and

iii) if $\bar{u}$ is a probability density, then u_t is a probability density for every $t \in [0,T]$.

Proof. Indeed, it is sufficient to take convolutions with respect to both variables (t,x) thus providing a sequence for which the criterion quoted above applies. Let us remark that conservation of mass follows from the choice $f_n = \chi_n$ in (9.2), where χ_n is a usual cut-off function, as introduced e.g. in Remark 1.3. $\square$

When compared with existence results available in the literature, such as [DiPerna and Lions, 1989, Proposition II.1] or the first part of [Figalli, 2008, Theorem 4.3], we see that these are fully recovered: actually, we obtain slightly stronger results when compared to the latter case, since we allow for unbounded coefficients. Moreover, in the elliptic case, we exploit the validity of the d-dimensional Sobolev inequality, to strengthen our existence result, obtaining the following

Corollary 11.2 (existence of solutions, elliptic case). *Let $r \in (1,\infty]$, a, $b \in L^1_t(L^{r'}_x)$, $c \in L^1_t(L^d(\mathbb{R}^d;\mathbb{R}^d))$ and, for some $\lambda > 0$, let*

$$a \geq \lambda Id, \quad \operatorname{div}[b - \operatorname{div} a]^- \in L^\infty_t(L^\infty_x + L^{d/2}_x).$$

Then, for every $\bar{u} \in L^2 \cap L^r(\mathbb{R}^d)$, the conclusions of the previous theorem hold for $\mathcal{L} := \mathcal{L}(a, b + c)$, and the solution built belongs to $L^2_t(W^{1,2}(\mathbb{R}^d))$.

11.2.2 Commutator estimates

Commutator estimates lie at the core of our approach to the theory, so in this section we carefully comment on how our computations specialize in the Euclidean setting. At the same time, we show that the Euclidean structure allows for improving what we obtain in the general framework, and provide a comparison with known results in the literature.

The abstract strategy for the commutator estimates developed in Section 10.2 give rise to explicit expressions, using the representation (11.1) for the heat semigroup. We proceed as follows: first, we focus on the case of the commutator between a derivation and the heat semigroup, specializing the results from Section 10.2.2. Then, we consider the case of a diffusion operator whose infinitesimal covariance belongs to the second order Sobolev space $W^{2,q}(\mathbb{R}^d)$, slightly improving the results from Section 10.2.3.

We let throughout q, r, $s \in [1,\infty]$ satisfy $q^{-1} + r^{-1} + s^{-1} = 1$. As already remarked in Section 10.2, the role played by the variable $t \in (0,T)$ is marginal, so that we directly argue in the time-independent setting, and let $\mathscr{A} = C^2_c(\mathbb{R}^d)$.

By standard density results, which are far from being trivial in the metric measure space setting, but in Euclidean spaces follow from straightforward convolution, we may assume all the objects involved to be smooth and compactly supported, as long as we provide a-priori estimates where only on the appropriate Sobolev norms appear.

The commutator with a Sobolev derivation

Given a smooth vector field b and functions u and f, we consider the commutator

$$\int u[\mathsf{P}_\alpha, b\cdot\nabla] f d\mathscr{L}^d = \int u \left[\mathsf{P}_\alpha(b\cdot\nabla f) - b\cdot\nabla \mathsf{P}_\alpha f\right] d\mathscr{L}^d.$$

For simplicity, assume that $\operatorname{div} b = 0$. Then, our Bakry-Émery interpolation argument reads as

$$\int u[\mathsf{P}_\alpha, b\cdot\nabla] f d\mathscr{L}^d = \int_0^\alpha \int u^\sigma [\Delta, b\cdot\nabla] f^{\alpha-\sigma} d\mathscr{L}^d d\sigma$$

and integration by parts provides the identity

$$\int u[\mathsf{P}_\alpha, b\nabla] f d\mathscr{L}^d = -2\int_0^\sigma \int \left\langle (D^{sym} b)\nabla u^\sigma, \nabla f^{\alpha-\sigma}\right\rangle d\mathscr{L}^d d\sigma.$$

At fixed $\sigma \in (0, \alpha)$, we obtain by (11.2) that $\langle (D^{sym} b) \nabla u^{\sigma}, \nabla f^{\alpha - \sigma} \rangle$, evaluated at $x \in \mathbb{R}^d$, coincides with

$$\frac{1}{2\sqrt{\sigma(\alpha-\sigma)}} \int \int \left[D^{sym} b(x) : y \otimes z\right] u(x + \sqrt{2\sigma} y) f(x + \sqrt{2(\alpha-\sigma)} z) \rho(y) \rho(z) dy dz.$$

It is interesting to compare this identity with the classical scheme introduced in [DiPerna and Lions, 1989, Lemma II.1], which relies instead on approximations by convolutions. Indeed, the heat semigroup P_α can be also written as a convolution operator,

$$\mathsf{P}_\alpha f(x) = \int f(x + \sqrt{2\alpha} y) d\rho(y) = \int \tau^y_{\sqrt{2\alpha}} f(x) d\rho(y),$$

i.e. a Gaussian average of translation semigroups $\tau^y_t f(x) := f(x + ty)$. For each translation operator, along the direction $y \in \mathbb{R}^d$, one interpolates

$$\tau^y_\varepsilon (b \cdot \nabla f) - b \cdot \nabla (\tau^y_\varepsilon f) = \int_0^\varepsilon \tau^y_\sigma [y \cdot \nabla, b \cdot \nabla] \tau^y_{\varepsilon - \sigma} f d\sigma = \int_0^\varepsilon \tau^y_\sigma \left[(y \cdot \nabla b) \cdot \nabla \tau^y_{\varepsilon-\sigma} f \right] d\sigma,$$

and from this one is able to bound the commutator, after some manipulations. It looks like that our approach is different, as the symmetric part of the derivative of the vector field b appears as the commutator with respect to the second order operator Δ.

However, as long as $|D^{sym} b| \in L^q(\mathbb{R}^d)$ with $q \in (1, \infty]$, both approaches are equivalent, i.e., they provide comparable estimates. When $q = 1$, as we proved that both the L^∞-Γ and the L^∞-Δ inequalities hold, it is possible to establish the inequality

$$\left| \int u [\mathsf{P}_\alpha, b \cdot \nabla] f d\mathscr{L}^d \right| \leq C \left(\|D^{sym} b\|_1 + \|\operatorname{div} b\|_1 \right) \|u\|_\infty \|f\|_\infty ,$$

allowing for the study of vector fields with $|D^{sym} b| \in L^1(\mathbb{R}^d)$. Let us remark, however that it is not clear whether the case $b \in BV(\mathbb{R}^d; \mathbb{R}^d)$, with $\operatorname{div} b \in L^1(\mathbb{R}^d)$, first settled by Ambrosio [2004] can be studied by means of this technique: certainly, a natural strategy would be to consider anisotropic heat semigroups.

Let us point out, however, that the DiPerna-Lions approach easily allows for localization, simply choosing a compactly supported mollifier, while our setting is intrinsically global. In order to adapt our methods to the local case, such as e.g. that of more general open sets in $\mathbb{R}^d$, one could "localize the Dirichlet form" by considering $X = B_r(0)$ and the form

$$\mathcal{E}_r(f) = \int_{B_r} |\nabla f|^2 \, d\mathscr{L}^d, \quad \text{for } f \in H^1(B_r).$$

Thus, Δ would be the Laplacian with Neumann boundary conditions and $(\mathsf{P}_t)_t$ would be the semigroup correspondent to the Brownian motion, reflected at the boundary $\partial B_r(0)$. Being the ball convex, the validity of L^p-Γ inequalities follows from lower bounds on the Ricci curvature, see e.g. Section 13.1 and also [Ambrosio et al., 2014b, Theorem 6.20].

The commutator with a Sobolev diffusion

When focus on the commutator between the heat semigroup and a diffusion operator. It turns out that the results in Section 10.2.3 extend, from operators of the form $\mathcal{L} f := a \Delta f$, to the general case $\mathcal{L}f := a : \nabla^2 f$, following an identical interpolation strategy. For the sake of clarity, we develop it independently of the results above. Precisely, we let

$$[\mathsf{P}_\alpha, a : \nabla^2]f := \mathsf{P}_\alpha(a : \nabla^2 f) - a : \nabla^2(\mathsf{P}_\alpha f), \quad \text{for } f \in \mathscr{A},$$

and prove the following

Lemma 11.3 (commutator lemma for Sobolev diffusions)**.** *Let $a \in W^{2,q}(\mathbb{R}^d; \mathbb{R}^{d\times d})$. Then, for every $\alpha \in (0,1)$, u, $f \in \mathscr{A}$, it holds*

$$\left| \int u[\mathsf{P}_\alpha, a : \nabla^2] f d\mathscr{L}^d - \alpha \int u[\Delta, a : \nabla^2]\mathsf{P}_\alpha f d\mathscr{L}^d \right| \le c \left\|a\right\|_{W^{2,q}} \left\|u\right\|_{L^2\cap L^r} \left\|f\right\|_{L^2 \cap L^s} \tag{11.7}$$

where c is some constant depending on d only. Moreover, for every $u \in L^2 \cap L^r \cap L^s(\mathfrak{m})$, it holds

$$\left| \int u[\mathsf{P}_\alpha, a : \nabla^2](\mathsf{P}_\alpha u) d\mathscr{L}^d \right| \to 0, \quad \textit{as } \alpha \downarrow 0. \tag{11.8}$$

Proof. In order to prove (11.7), the idea is to write

$$a : \nabla^2 f = \left[a : (\nabla^2 \Delta^{-1})\right] \Delta f = \mathsf{a}\Delta f, \quad \text{for } f \in \mathscr{A},$$

thus obtaining an expression similar to that in Section 10.2.3. In place of the multiplication by some function we have here the linear continuous operator a, from $L^s(\mathbb{R}^d)$ to $L^{s'}(\mathbb{R}^d)$ (due to boundedness of Riesz transforms) but we perform a similar "second order" interpolation along the semigroup, exploiting also the fact that directional derivatives, Laplacians and the heat semigroup commute. To make computations more transparent, we prefer to directly argue on coordinates, thus we fix i, $j \in \{1, \ldots, d\}$ and consider the commutator

$$[\mathsf{P}_\alpha, a^{i,j}\partial^2_{i,j}]f := \mathsf{P}_\alpha(a^{i,j}\partial^2_{i,j} f) - a^{i,j}\partial^2_{i,j}(\mathsf{P}_\alpha f).$$

As in the proof of Lemma 10.2, we introduce the curve

$$[0, \alpha] \ni \sigma \mapsto F(\sigma) = \int u^\sigma a^{i,j}\partial^2_{i,j} f^{\alpha-\sigma} d\mathscr{L}^d,$$

which is $C^1((0,\alpha), \mathbb{R})$, with

$$F'(\sigma) = \int u^\sigma [\Delta, a^{i,j}\partial^2_{i,j}] f^{\alpha-\sigma} d\mathscr{L}^d = \int u^\sigma [\Delta, a^{i,j}]\partial^2_{i,j} f^{\alpha-\sigma} d\mathscr{L}^d,$$

since Δ and partial derivatives commute. We let $h^{\alpha-\sigma} := \partial^2_{i,j} f^{\alpha-\sigma} = (\partial^2_{i,j} f)^{\alpha-\sigma}$, since derivatives and semigroup commute, by the expression (11.1). By Example 10.1, it holds, for the derivation $\boldsymbol{b} = \boldsymbol{b}^{i,j}$ induced by $\nabla a^{i,j}$,

$$\begin{aligned} F'(\sigma) &= \int u^\sigma dh^{\alpha-\sigma}(\boldsymbol{b}) - du^\sigma(\boldsymbol{b}) h^{\alpha-\sigma} d\mathscr{L}^d \\ &= \int u^\sigma dh^{\alpha-\sigma}(\boldsymbol{b}) + u^\sigma \operatorname{div}(h^{\alpha-\sigma}\boldsymbol{b}) d\mathscr{L}^d \\ &= 2\int u^\sigma dh^{\alpha-\sigma}(\boldsymbol{b}) d\mathscr{L}^d + \int u^\sigma (\Delta a^{i,j}) h^{\alpha-\sigma} d\mathscr{L}^d. \end{aligned}$$

This identity gives at once that $F \in C^2((0,\alpha),\mathbb{R})$, with

$$F''(\sigma) = 2\int u^\sigma[\Delta, \boldsymbol{b}]h^{\alpha-\sigma}d\mathscr{L}^d + \int u^\sigma[\Delta, (\Delta a^{i,j})]h^{\alpha-\sigma}d\mathscr{L}^d.$$

We then perform an interpolation up to the second order,

$$F(\alpha) - F(0) - \alpha F'(0) = \int_0^\alpha F''(\sigma)(\alpha-\sigma)d\sigma, \tag{11.9}$$

where the factor $(\alpha-\sigma)$ is useful to compensate the bound on the norm of h^α.

As with the case of derivations, our deductions are straightforward in case $\Delta a^{i,j} = 0$, since (10.13) gives

$$\int u^\sigma[\Delta, \boldsymbol{b}]h^{\alpha-\sigma}d\mathscr{L}^d = -2\int \left\langle (\nabla^2 a^{i,j})\nabla u^\sigma, \nabla h^{\alpha-\sigma}\right\rangle d\mathscr{L}^d$$

and we estimate

$$\begin{aligned} \left|F''(\sigma)\right| &\le 4\left\|\nabla^2 a^{i,j}\right\|_q \|\nabla u^\sigma\|_r \|\nabla h^{\alpha-\sigma}\|_s \\ &\le \frac{4c_r^\Gamma c_{s,3}}{\sqrt{\sigma(\alpha-\sigma)^3}}\left\|\nabla^2 a^{i,j}\right\|_q \|u\|_r \|f\|_s, \end{aligned} \tag{11.10}$$

where we apply both the L^r-Γ inequality (11.3) and inequality (11.4) for $p = s$. Integrating with respect to $\sigma \in (0,\alpha)$, we would conclude 11.7.

To address the general case where $\Delta a^{i,j} \in L^q(\mathfrak{m})$, we argue as in the proof of Lemma 10.2, i.e. we add and subtract suitable quantities. We consider separately the terms

$$\int_0^\alpha \int u^\sigma[\Delta, \boldsymbol{b}]h^{\alpha-\sigma}d\mathscr{L}^d\,(\alpha-\sigma)d\sigma \quad \text{and} \quad \int_0^\alpha \int u^\sigma[\Delta, (\Delta a^{i,j})]h^{\alpha-\sigma}d\mathscr{L}^d\,(\alpha-\sigma)d\sigma. \tag{11.11}$$

We focus on the former, recalling identity (10.13) that gives, for for $\sigma \in (0,\alpha)$, an equivalent expression for $\int \left\langle (\nabla^2 a^{i,j})\nabla u^\sigma, \nabla h^{\alpha-\sigma}\right\rangle d\mathscr{L}^d$, namely

$$-\frac{1}{2}\int u^\sigma[\Delta, \boldsymbol{b}]h^{\alpha-\sigma} - u^\sigma(\Delta a^{i,j})\Delta h^{\alpha-\sigma} - (\Delta a^{i,j})\nabla u^\sigma \cdot \nabla h^{\alpha-\sigma}d\mathscr{L}^d,$$

thus, in order to reduce to the argument for the case $\Delta a^{i,j} = 0$, it is enough to provide bounds for the quantities

$$\int u^\sigma(\Delta a^{i,j})\Delta h^{\alpha-\sigma}d\mathscr{L}^d, \quad \text{and} \quad \int (\Delta a^{i,j})\nabla u^\sigma \cdot \nabla h^{\alpha-\sigma}d\mathscr{L}^d. \tag{11.12}$$

The inequality

$$\left|\int (\Delta a^{i,j})\nabla u^\sigma \cdot \nabla h^{\alpha-\sigma}d\mathscr{L}^d\right| \le \left\|\Delta a^{i,j}\right\|_q \|\nabla u^\sigma\|_r \left\|\nabla h^{\alpha-\sigma}\right\|_s$$

allows us to handle the second term in (11.12) exactly as in (11.10), while for the first term in (10.22), we use a second-order analogue of Lemma 10.3. We introduce the quantity

$$\int_0^\alpha \int u^\alpha(\Delta a^{i,j})R^{i,j}\Delta^2 f^{\alpha-\sigma}(\alpha-\sigma)d\mathscr{L}^d d\sigma, \tag{11.13}$$

where we let $R^{i,j} f := \partial^2_{i,j} \Delta^{-1} f$ be the second-order Riesz transform along the directions i, j. By the Taylor expansion (11.9) with $f^{\alpha-\sigma}$ in place of $F(\sigma)$, we have

$$\int_0^\alpha \Delta^2 f^{\alpha-\sigma}(\alpha-\sigma) d\sigma = f - f^\alpha + \alpha \Delta f^\alpha,$$

entailing the bound

$$\left| \int_0^\alpha \int u^\alpha (\Delta a^{i,j}) R^{i,j} \Delta^2 f^{\alpha-\sigma}(\alpha-\sigma) d\mathscr{L}^d d\sigma \right| \leq \left\| R^{i,j} \right\|_{L^r \to L^r} (2 + c_s^\Delta) \left\| \Delta a^{i,j} \right\|_q \|u\|_r \|f\|_s .$$

Therefore, we are allowed to add and subtract (11.13) in the first term of (11.12), and we are reduced to provide a bound for difference

$$\left| \int (u^\alpha - u^\sigma)(\Delta a^{i,j}) \Delta h^{\alpha-\sigma} d\mathscr{L}^d \right|,$$

to be integrated over $\sigma \in (0, \alpha)$, with respect to the measure $(\alpha - \sigma) d\sigma$. By inequality (11.4) with $k = 2$, it holds

$$\begin{aligned} \left| \int (u^\alpha - u^\sigma)(\Delta a^{i,j}) \Delta h^{\alpha-\sigma} d\mathscr{L}^d \right| &\leq \left\| \Delta a^{i,j} \right\|_q \|u^\alpha - u^\sigma\|_r \|\partial^2_{i,j} \mathsf{P}_{(\alpha-\sigma)/2} (\Delta f^{(\alpha-\sigma)/2})\|_s \\ &\leq \frac{2 c_{2,\sigma}}{\alpha - \sigma} \left\| \Delta a^{i,j} \right\|_q \|u^\alpha - u^\sigma\|_r \|\Delta f^{(\alpha-\sigma)/2}\|_s \end{aligned}$$

and from this point we conclude identically as in the proof of Lemma 10.3, i.e. by the L^p-Δ inequality and Corollary 3.5. This provides the required bounds for the former term in (11.11).

The latter term in (11.11) is easier to bound, since

$$\int u^\sigma [\Delta, (\Delta a^{i,j})] h^{\alpha-\sigma} d\mathscr{L}^d = \frac{d}{ds} \int u^\sigma (\Delta a^{i,j}) \partial^2_{i,j} f^{\alpha-\sigma} d\mathscr{L}^d,$$

and so we can integrate by parts in (11.11),

$$\int_0^\alpha \int u^\sigma [\Delta, (\Delta a^{i,j})] \partial^2_{i,j} f^{\alpha-\sigma} d\mathscr{L}^d \, (\alpha-\sigma) d\sigma = -\alpha \int u (\Delta a^{i,j}) \partial^2_{i,j} f^\alpha d\mathscr{L}^d + \int_0^\alpha \int u^\sigma (\Delta a^{i,j}) \partial^2_{i,j} f^{\alpha-\sigma} d\mathscr{L}^d.$$

The first addend in the right hand side is uniformly bounded from above by $c_{s,2} \left\| \Delta a^{i,j} \right\|_q \|u\|_r \|f\|_s$, so we are left only with

$$\int_0^\alpha \int u^\sigma (\Delta a^{i,j}) \partial^2_{i,j} f^{\alpha-\sigma} d\mathscr{L}^d = \int_0^\alpha \int u^\sigma (\Delta a^{i,j}) R^{i,j} \Delta f^{\alpha-\sigma} d\mathscr{L}^d,$$

but this can be handled directly as in the proof of Lemma 10.3, i.e. by adding and subtracting

$$\int_0^\alpha \int u^\alpha (\Delta a^{i,j}) R^{i,j} \Delta f^{\alpha-\sigma} d\mathscr{L}^d = \int u^\alpha (\Delta a^{i,j}) R^{i,j} (f^\alpha - f) d\mathscr{L}^d,$$

and using the L^p-Δ inequality and Corollary 3.5. Summation upon i, $j \in \{1, \ldots, d\}$ yields (11.7).

Next, we address the validity of (11.8). Arguing at fixed i, $j \in \{1,\ldots,d\}$, we integrate by parts

$$\begin{aligned}\int u[\Delta, a^{i,j}\partial^2_{i,j}]f^\alpha d\mathscr{L}^d &= -2\int du(\boldsymbol{b})\partial^2_{i,j}f^\alpha d\mathscr{L}^d - \int u(\Delta a^{i,j})\partial^2_{i,j}f^\alpha d\mathscr{L}^d\\ &= -2\int (\partial^2_{i,j}f)\mathsf{P}_\alpha(du(\boldsymbol{b}))d\mathscr{L}^d - \int u(\Delta a^{i,j})\partial^2_{i,j}f^\alpha d\mathscr{L}^d\\ &= -2\int (\partial^2_{i,j}f)[\mathsf{P}_\alpha,\boldsymbol{b}]ud\mathscr{L}^d - 2\int (\partial^2_{i,j}f)d(u^\alpha)(\boldsymbol{b})d\mathscr{L}^d - \int u(\Delta a^{i,j})\Delta f^\alpha d\mathscr{L}^d.\end{aligned}$$

By the commutator estimate for Sobolev derivations, this last identity extends by continuity to the case $u \in L^2\cap L^r(\mathbb{R}^d)$, $f \in W^{2,s}(\mathbb{R}^d)$.

We now specialize to the case $f := \mathsf{P}_\alpha u$. As $\alpha \downarrow 0$, it holds

$$\alpha\left|\int (\partial^2_{i,j}u^\alpha)[\mathsf{P}_\alpha,\boldsymbol{b}]ud\mathscr{L}^d\right| \to 0,$$

by the second statement in Corollary 10.9 and boundedness of $\alpha\partial^2_{i,j}u^\alpha$ in $L^2\cap L^s(\mathbb{R}^d)$. We also have

$$\left|\int u(\Delta a^{i,j})\Delta f^\alpha d\mathscr{L}^d\right| \le \left\|\Delta a^{i,j}\right\|_q \|u\|_r \|\partial^2_{i,j}u^{2\alpha}\|_s \to 0.$$

In order to handle the last term, our choice of f in terms of u and the symmetry of a seem crucial. Indeed, we integrate by parts once, obtaining

$$\int (\partial^2_{i,j}u^\alpha)du^\alpha(\boldsymbol{b})d\mathscr{L}^d = -\sum_{k=1}^d\int \partial_i u^\alpha(\partial^2_{j,k}a^{i,j})\partial_k u^\alpha + \partial_i u(\partial_k a^{i,j})\partial^2_{k,j}ud\mathscr{L}^d.$$

The first term, when multiplied by α, is clearly bounded and infinitesimal as $\alpha\downarrow 0$, so we focus on the last one. To show that it is bounded, we recall that a is symmetric and we are summing upon i, $j \in \{1,\ldots d\}$, so that we are reduced to prove that

$$\int \partial_i u^\alpha(\partial_k a^{i,j})\partial^2_{k,j}u^\alpha + \partial_i u^\alpha(\partial_k a^{i,j})\partial^2_{k,i}u^\alpha d\mathscr{L}^d$$

is infinitesimal, when multiplied by α. This symmetric expression can be rewritten

$$\frac{1}{2}\int (\partial_k a^{i,j})\partial_k\left[(\partial_i u^\alpha + \partial_j u^\alpha)^2 - (\partial_i u^\alpha)^2 - (\partial_j u^\alpha)^2\right]d\mathscr{L}^d,$$

from which we integrate by parts, obtaining a bound in terms of $\|a\|_{W^{2,q}}\|\nabla u^\alpha\|^2_{L^2\cap L^r\cap L^s}$, which is sufficient to conclude. □

11.2.3 Uniqueness results

In this section we collect our main uniqueness results for FPE's and flows in the Euclidean setting.

The (possibly) degenerate case

Following the smoothing scheme described in Section 10.3 and relying on the commutator estimate established in the previous section, we obtain the following strengthening of Theorem 10.20, where general diffusion operators are allowed, as long as they satisfy Sobolev bounds on their coefficients.

Theorem 11.4 (uniqueness of solutions, Sobolev diffusions). *Let $q \in (1, \infty]$, r, $s \in (1, \infty)$ satisfy $q^{-1} + r^{-1} + s^{-1} = 1$ and let*

$$(b_t)_{t \in (0,T)} \in L^1_t(W^{1,q}(\mathbb{R}^d; \mathbb{R}^d)), \quad (a_t)_{t \in (0,T)} \in L^1_t(W^{2,q}(\mathbb{R}^d; \mathbb{R}^{d \times d})),$$

define $\mathcal{L} := \mathcal{L}(a, b)$ as in (11.5) and assume that

$$\operatorname{div} \mathcal{L}^- = (\operatorname{div} \boldsymbol{b} - \Delta a)^- \in L^1_t(L^\infty_x).$$

Then, for every $\overline{u} \in L^2 \cap L^r \cap L^s(\mathbb{R}^d)$, there exists at most one weakly continuous solution u in $L^\infty_t(L^2_x \cap L^r_x \cap L^s_x)$ to the FPE

$$\partial_t u = \mathcal{L}^* u, \quad \text{in } (0, T) \times X, \text{ with } u_0 = \overline{u}.$$

Combining this result with the existence result in Theorem 11.1, the validity of the superposition principle for diffusions and the abstract correspondence for well-posedness, we deduce the following uniqueness results for MP's and regular flows:

Theorem 11.5 (well-posedness for martingale problems and flows, Sobolev diffusions). *Let $q \in (1, \infty]$, r, $s \in (1, \infty)$ satisfy $q^{-1} + r^{-1} + s^{-1} = 1$ and let*

$$(b_t)_{t \in (0,T)} \in L^q_t(W^{1,q}(\mathbb{R}^d; \mathbb{R}^d)), \quad (a_t)_{t \in (0,T)} \in L^q_t(W^{2,q}(\mathbb{R}^d; \mathbb{R}^{d \times d})),$$

define $\mathcal{L} := \mathcal{L}(a, b)$ as in (11.5) and assume that

$$\operatorname{div} \mathcal{L}^- = (\operatorname{div} \boldsymbol{b} - \Delta a)^- \in L^1_t(L^\infty_x).$$

Then, for every $\overline{u} \in L^2 \cap L^r \cap L^s(\mathbb{R}^d)$, with $\bar{u}$ probability density, there exists a unique L^r-regular solution u to the martingale problem associated to $\mathcal{L}$, on $C([0, T]; \mathbb{R}^d)$. Moreover, there exists a unique L^∞-regular martingale flow $(\boldsymbol{\eta}(s, x))_{s,x}$ associated to $\mathcal{L}$.

Notice that additional L^q-integrability with respect to $t \in (0, T)$ on the coefficients is introduced order to apply the superposition principle for diffusions.

The uniqueness result above is novel, to the author's knowledge. The literature on well-posedness for degenerate FPE's and related diffusions has been recently growing: we briefly compare our result with those obtained by Le Bris and Lions [2008] and Zhang [2010], although further improvements have been obtained, see e.g. [Röckner and Zhang, 2010], [Fang et al., 2010], [Luo, 2013]. To this aim, let us first point out that the assumptions $a \in W^{2,q}(\mathbb{R}^d)$ entail that the matrix square root σ belongs to $W^{1,q}(\mathbb{R}^d)$, see e.g. [Stroock and Varadhan, 2006, Lemma 3.2.3]. In [Le Bris and Lions, 2008], uniqueness for the FPE is proved in the class of functions $u \in L^\infty_t(L^1 \cap L^\infty_x)$ such that $\sigma \nabla u \in L^2_t(L^2_x)$, provided that b and σ are belong to first-order Sobolev spaces. Our result shows that, assuming slightly stronger differentiability assumptions, we are able to show uniqueness in the whole space $L^\infty_t(L^1 \cap L^\infty_x)$, allowing for the application of the general transfer mechanism between FPE's and martingale flows. The approach by Zhang [2010], which is based on quantitative estimates arguing directly on the flow, originally developed for ODE's by Crippa and De Lellis [2008], is purely "Lagrangian" and leads to uniqueness, provided that σ is (first order) Sobolev differentiable and bounded. In [Röckner and Zhang, 2010], by means of the superposition principle for diffusions from [Figalli, 2008], well-posedness results are transferred to FPE's, at least when solutions are curves of probability measures. Presently, our results impose stronger smoothness assumptions on the coefficients, but do not require their (local) boundedness. However, in view of the stronger superposition principle for diffusions developed in Part I, it is reasonable to expect that one may extend the results from [Röckner and Zhang, 2010] to the case of non-necessarily locally bounded coefficients.

The deterministic case

The deterministic case, where $a = 0$ and $\mathcal{L} = \mathcal{L}(0, b)$ is a time-dependent family of derivations, can be read as a special case of the degenerate case above. In particular, Theorem 11.4 becomes a uniqueness result for continuity equations driven by Sobolev vector fields, akin to [DiPerna and Lions, 1989, Corollary II.1], where uniqueness for transport equations is settled, under similar assumptions.

At the level of flows, after Remark 6.17, we obtain that the unique L^∞-regular martingale flow provided by Theorem 11.5 is actually deterministic, i.e., for every $s \in [0, T]$, $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$, $\boldsymbol{\eta}(s, x)$ is a Dirac measure. Equivalently, for every $s \in [0, T]$, we obtain a map $\mathbb{X}_s : \mathbb{R}^d \times [0, T] \mapsto \mathbb{R}^d$ such that

i) $t \mapsto \mathbb{X}_s(x, t)$ is a $\mathscr{L}^1$-a.e. solution to the ODE $\frac{d}{dt}\gamma(t) = b_t(\gamma(t))$ in $[s, T]$ with $\mathbb{X}_s(s, x) = x$, for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$;

ii) for every $\bar{u} \in L^\infty(\mathbb{R}^d)$, with $\bar{u}\mathfrak{m}$ probability, the probability measure $\mathbb{X}_s(\cdot, t)_\sharp(\bar{u}\mathscr{L}^d)$ is absolutely continuous with respect to $\mathscr{L}^d$, with density uniformly bounded in $L^\infty(\mathbb{R}^d)$.

Moreover, uniqueness entails the semigroup law $\mathbb{X}_s(t, x) = \mathbb{X}_r(t, \mathbb{X}_s(r, x))$, for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$, for every r, s, $t \in [0, T]$ with $s \le r \le t$.

As already observed above (see Remark 1.11), when compared with the formalization of DiPerna-Lions theory introduced by Ambrosio [2004], the first difference is due the presence of the extra parameter $s \in [0, T]$, that seems to play an important role in the case of general diffusions. Another difference, as remarked in the previous section, is that we are currently not able to obtain useful commutator estimates for vector fields $b \in BV(\mathbb{R}^d; \mathbb{R}^d)$.

The elliptic case

In the bounded elliptic case, the uniqueness result provided by Theorem 10.21 roughly corresponds to the uniqueness part of [Figalli, 2008, Theorem 4.3]. Let us remark that the proof of the crucial commutator estimates therein rely on compactness of the embedding $W^{1,2}(\mathbb{R}^d) \to L^2_{loc}(\mathbb{R}^d)$, while in our abstract approach this is not necessary, and this might be useful to prove similar results in infinite dimensional spaces as well. There is however another issue which has to be carefully addressed, that of the density of $\mathscr{A} = C_b^{1,2}((0, T) \times \mathbb{R}^d)$ in the domain of $\Delta[\boldsymbol{a}]f := \operatorname{div}(a\nabla f)$. Assuming that $\operatorname{div} a \in L^\infty_t(L^\infty_x)$, as it is done in Figalli [2008], one obtains that $\mathscr{A}$ is contained in the domain of $\Delta[\boldsymbol{a}]$, but density could not hold: clearly, if some further regularity on the coefficients of a is imposed, (first order Sobolev regularity should be sufficient), one obtains the required density. Another strategy instead would be that of enlarging $\mathscr{A}$ with all the functions belonging to $W^{1,2}_t(\mathbb{R}^d) \cap D(\Delta[\boldsymbol{a}])$, which is a set large enough to entail uniqueness: after all, this is not surprising, since we are considering the operator $\partial_t + \Delta[\boldsymbol{a}]$ as a perturbation of $\Delta[\boldsymbol{a}]$.

11.3 Riemannian and sub-Riemannian spaces

Before we conclude this chapter, we sketch how the abstract arguments provide straightforward extensions of the classical DiPerna-Lions theory to the setting of weighted Riemannian manifolds, and even sub-Riemannian spaces: we consider these as examples showing the flexibility of our techniques, thus we do not enter too deep into details, as this would require specific introductions. Of course, in order to prove strong convergence of commutators and argue well-posedness, one could always reduce to computations in local charts, but these become more cumbersome, compared to the Euclidean case, and here the advantages of our intrinsic (and global) approach become manifest.

Weigthed Riemannian manifolds

Let $(M, \boldsymbol{g})$ be a smooth Riemannian manifold (we refer e.g. to [Bakry et al., 2014, Appendix C] for a brief introduction) and let μ be its associated Riemannian volume measure. Assume that the Ricci curvature tensor $\mathrm{Ric}_{\boldsymbol{g}}$ is pointwise bounded from below, in the sense of quadratic forms, by some constant $K \in \mathbb{R}$. More generally, one can add a "weight" $V : M \to \mathbb{R}$ to the measure, i.e. consider the reference measure $e^{-V}\mu$ and assume that the Bakry-Émery curvature tensor is bounded from below by $K \in \mathbb{R}$, i.e.

$$\mathrm{Ric}_{\boldsymbol{g}} + \mathrm{Hess}(V) \geq K.$$

The form (on smooth compactly supported functions)

$$f \mapsto \mathcal{E}_V(f) = \int_M \boldsymbol{g}(\nabla f, \nabla f) e^{-V} d\mu,$$

is closable and we are in the setup (3.1), the Laplace operator being a (weighted) Laplace-Beltrami on $(M, \boldsymbol{g})$. Once more, the algebra $\mathscr{A}$ of test functions can be chosen to be the space of smooth functions with compact support.

When $V = 0$, Bochner's formula entails that L^p-Γ inequalities (for $p \in (1, \infty)$) holds, since the so-called $\mathsf{BE}(K, \infty)$ curvature condition holds (see Chapter 13) and it is a classical result due to S.-T. Yau that the heat semigroup is conservative. In the case of weighted measures, analogous results can be found in [Bakry, 1994, Proposition 6.2] for the curvature bound and in [Grigor′yan, 1999, Theorem 9.1] for the conservativity of P, relying on a correspondent volume comparison theorem, see e.g. [Wei and Wylie, 2009, Theorem 1.2].

Given a Borel vector field b, i.e. a Borel section of the tangent bundle of M, its associated derivation $\boldsymbol{b}$ acts on smooth functions by

$$f \mapsto df(\boldsymbol{b}) = \boldsymbol{g}(b, \nabla f).$$

The divergence can be given in terms of the μ-distributional divergence of b by

$$\operatorname{div} \boldsymbol{b} = \operatorname{div} b - \boldsymbol{g}(\boldsymbol{b}, \nabla V),$$

while the deformation is the symmetric part of the distributional covariant derivative, see Remark 10.6.

Sub-Riemannian spaces

Sub-Riemannian geometry has very recently become object of study from a point of view close to that of Γ-calculus, see e.g. [Baudoin et al., 2014] [Baudoin and Kim, 2014]. On $\mathbb{R}^d$, let $V = \{v^1, \ldots, v^n\}$ be a finite family of bounded vector fields, with bounded derivatives of all orders, and consider the form

$$\mathcal{E}_V(f) := \int \sum_{i=1}^{n} |\delta_i f|^2 \, d\mathscr{L}^d, \quad \text{for } f \in C^1_c(\mathbb{R}^d),$$

where we introduce the notation $\delta_i f := v^i \cdot \nabla f$. The form is easily seen to be closable and its closure is a Dirichlet form which satisfies (3.1), with

$$\nabla_V f = (\delta_i f)_{i=1}^n, \quad \Gamma_V(f) = |\nabla_V f|^2 = \sum_{i=1}^{n} |\delta_i f|^2 .$$

In particular, the associated domain $\mathbb{V}$ can be described as a suitable Sobolev space along the directions in V), and the Laplacian is given by

$$\Delta_V f := \sum_{i=1}^{n} \operatorname{div}(v^i \delta_i f), \quad \text{for } f \in C^2_c(\mathbb{R}^d),$$

which can be proved to be a dense dense space in $D(\Delta_V)$. If b is a bounded Borel vector field along V, i.e. there exists Borel functions $(b^i)_{i=1}^n$ such that $b = \sum_{i=1}^n b^i v^i$, then our well-posedness results entail existence and uniqueness for bounded solutions the FPE

$$\partial_t u_t = (\Delta_V + b \cdot \nabla_V)^* \, u_t, \quad \text{in } (0,T) \times \mathbb{R}^d, \, u_0 = \bar{u} \in L^1 \cap L^\infty(\mathbb{R}^d).$$

On the other side, the more general case of degenerate diffusions seems to lie presently outside the scope of our theory, in the sense that the conditions provided by Theorem 10.19 appear to be rather restrictive. For simplicity, assume that $\operatorname{div} b = 0$, thus in the commutator estimate we obtain the quantity

$$\int D^{sym}_V b(f,g) d\mathscr{L}^d = \int \langle \nabla_V f, \nabla_V (b \nabla_V g) \rangle + \langle \nabla_V g, \nabla_V (b \nabla_V f) \rangle .$$

An explicit computation shows that the integrand in the right hands side above coincides with the summation upon $i, j \in \{1, \ldots, n\}$ of

$$\begin{aligned} &\delta_i f \, \delta_i (b^j \delta_j g) + \delta_j f \, \delta_j (b^i \delta_i g) \\ =&\delta_i f \left[\delta_i b^j + \delta_j b^i \right] \delta_j g + (\delta_i f) \, b^j \delta_i \delta_j g + (\delta_j g) \, b^i \delta_j \delta_i f \\ =&\delta_i f \left[\delta_i b^j + \delta_j b^i \right] \delta_j g + b^j \delta_j (\delta_i f \delta_i g) + (\delta_i f) \, b^j \left[\delta_i, \delta_j \right] g + (\delta_j g) \, b^i \left[\delta_j, \delta_i \right] f, \end{aligned}$$

and it is not clear how to provide estimates for the last two terms above by means of Γ_V.

Chapter 12

Gaussian spaces

In this chapter, we show that our theory specializes to infinite dimensional spaces as well. We limit the discussion to Gaussian spaces, i.e. Banach spaces endowed with a Gaussian measures, where our results can be compared with those from some recent literature.

Let us remark that (at least) two different theories of analysis on Gaussian spaces can be developed: on one side, the classical Malliavin calculus, see e.g. [Bouleau and Hirsch, 1991] or [Nualart, 2006], on Banach spaces, and on the other side, the case of Gaussian Hilbert spaces, see e.g. [Da Prato, 2014] or [Da Prato, 2004]. Roughly, the main difference between the two approaches is the choice of norm in the tangent space: in the former case, we choose the Cameron-Martin norm, in the latter one, we take the Hilbert norm, by identifying the tangent with the space itself. Despite their distinctive features, e.g., in the latter, Sobolev embeddings are compact, while in the former they are not, it turns out that each of them can be recovered as a special case of our framework. For the sake of brevity, we choose to provide a rather detailed description of the setting (Section 12.1) and well-posedness results (Section 12.2) in the case classical Malliavin calculus only, and briefly describe how the theory specializes to Gaussian Hilbert spaces in Section 12.4.

Moreover, in Section 12.3, we report the main result established in Trevisan [2014a], entailing well-posedness for the continuity equation associated with BV regular vector fields, giving a non-trivial extension to infinite dimensional spaces of the breakthrough by Ambrosio [2004], where the case of Euclidean BV vector fields is settled.

12.1 The Wiener space setting

12.1.1 Dirichlet form setup

In the basic setup of Chapter 3, we let X be a separable Banach space, $\mathfrak{m} = \gamma$ the a Gaussian measure on the Borel sets of X, i.e. for every $x^* \in X^*$, the law of x^* (i.e. the push-forward $x^*_\sharp\gamma$) is a normal law on $\mathbb{R}$. We refer to the already quoted monograph [Bouleau and Hirsch, 1991] for further details. For simplicity, let γ be centred and non-degenerate, i.e. each law $(x^*)_\sharp$ has mean 0 and is absolutely continuous with respect to $\mathscr{L}^1$. We embed $X^* \subseteq L^2(\gamma)$ and moreover the covariance operator $Q : X^* \to X$, given by the Bochner integral

$$Q(x^*) = \int y x^*(y) d\gamma(y), \quad \text{for } x^* \in X^*,$$

defines a linear, continuous and injective operator. Moreover, for every x^*, $y^* \in X^*$, it holds

$$\int_X x^*(x)y^*(x)d\gamma(x) = x^*(Q(y^*)) = y^*(Q(x^*)).$$

For $x = Qx^* \in Q(X^*) \subseteq X$, we define its Cameron-Martin norm by

$$|x|^2_{\mathcal{H}} = x^*(x) = \int_X |x^*|^2 \, d\gamma,$$

and define the Cameron-Martin space $\mathcal{H}$ as the completion of $Q(X^*)$ with respect to $|\cdot|_{\mathcal{H}}$: endowed with the extension of norm it is a Hilbert space and a subspace of X, with $(H, \mathcal{H}) \to (X, \|\cdot\|_X)$ compact. It can be seen that $\mathcal{H}$ is isometric to the closure of X^* in $L^2(\gamma)$, thus an injective map $\mathcal{H} \in h \mapsto \hat{h} \in L^2(\gamma)$ is defined. The space $\mathcal{H}$ plays a crucial role in the definition of Sobolev spaces, because it fully characterizes the directions in X for which the translation along them preserves absolute continuity with respect to γ.

We define the set of smooth cylindrical functions $\mathcal{FC}^\infty_b(X)$ as the set of all functions $f(x)$ representable as $\varphi(x_1^*(x), \ldots, x_n^*(x))$, with $\varphi : \mathbb{R}^n \to \mathbb{R}$ smooth and bounded, $x_i^* \in X^*$ for $i \in \{1, \ldots, n\}$, for some integer $n \geq 1$.

We introduce a notion of gradient on functions $f \in \mathcal{FC}^\infty_b(X)$ letting $\nabla_{\mathcal{H}} f = Qdf$, where df is the Frechét differential of f. With these definitions, for $f = \varphi(x_1^*, \ldots, x_n^*)$, one has

$$\nabla_{\mathcal{H}} f(x) = \sum_{j=1}^n \frac{\partial \varphi}{\partial z_j} Q x_j^* = \sum_{k=1}^\infty \frac{\partial f}{\partial h_k}(x) h_k, \quad \text{where} \quad \frac{\partial f}{\partial h_k}(x) = \lim_{\varepsilon \to 0} \frac{f\left(x + \varepsilon h_k\right) - f\left(x\right)}{\varepsilon}$$

where $(h_k)_k$ is any complete orthonormal system in $\mathcal{H}$. A better description is obtained by choosing a $(h_k)_k$ of the form $h_k = Qe_k^*$, for $(e_k^*) \subseteq X^*$, which can be done by density of QX^* in $\mathcal{H}$. For brevity, we introduce the notation

$$\partial_k f := \frac{\partial f}{\partial h_k} = \left\langle h_k, \nabla_{\mathcal{H}} f \right\rangle.$$

It is well-known, see [Bouleau and Hirsch, 1991], that Sobolev-Malliavin calculus on $(X, \gamma, \mathcal{H})$ fits into the setting (3.1), considering the closure of the quadratic form

$$\mathcal{E}(f) = \int_X |\nabla_{\mathcal{H}} f|^2_{\mathcal{H}} \, d\gamma, \qquad \text{for every } f \in \mathcal{FC}^\infty_b(X).$$

The domain $\mathbb{V}$ coincides with the space $W^{1,2}(X, \gamma)$, defined is the usual Sobolev space of functions $f \in L^2(\gamma)$ with distributional derivative $\nabla_{\mathcal{H}} f \in L^2(\gamma)$, i.e. there exists a (unique) function $g \in L^2(\gamma; \mathcal{H})$ such that, for every $h \in \mathcal{H}$, the Gaussian integration by parts formula holds:

$$\int_X f \left[\frac{d\varphi}{dh} - \varphi \hat{h} \right] \gamma = - \int \left\langle h, g \right\rangle \varphi d\gamma, \quad \text{for every } \varphi \in \mathcal{FC}^\infty_b(X). \tag{12.1}$$

We let $\nabla_{\mathcal{H}} f := g$ and extend the notation $\partial_h f = \langle h, \nabla_{\mathcal{H}} f \rangle$ and $\partial_k f = \langle h_k, \nabla_{\mathcal{H}} f \rangle$ as well. Notice that the notation is consistent with the case of $f \in \mathcal{FC}^\infty_b(X)$. It can be proved that smooth cylindrical functions are dense in $W^{1,p}(X, \gamma)$, for $p \in [1, \infty)$.

The semigroup P is the Ornstein-Uhlenbeck semigroup, given by Mehler's formula

$$\mathsf{P}_t f(x) = \int_X f(e^{-t} x + \sqrt{1 - e^{-2t}} y) d\gamma(y), \quad \text{for } \gamma\text{-a.e. } x \in X.$$

The abstract Laplacian Δ acts on $x^* \in X^*$ by $\Delta x^* = -x^*(Q(x^*)) = -|x^*|_{\mathcal{H}}$, thus on smooth cylindrical functions of the form $f = \varphi(e_1^*, \dots e_n^*)$, the action of Δ is given by

$$\Delta\varphi(e_1^*, \dots, e_n^*) = \sum_{i,j=1}^{n} \partial_{i,j}\varphi(e_1^*, \dots, e_n^*) - \sum_{i=1}^{n} \partial_i\varphi(e_1^*, \dots, e_n^*).$$

The carré du champ is given by $\Gamma(f) = |\nabla_{\mathcal{H}} f|^2$, the spaces $\mathbb{V}^p$ and $D(\Delta)$ are identified respectively as

$$\mathbb{V}^p = W^{1,p} \cap W^{1,2}(X,\gamma) = \left\{ f \in L^p \cap L^2(\gamma) \,:\, \nabla f \in L^p \cap L^2(\gamma) \right\}$$

$$D^p(\Delta) = \{ f \in D(\Delta) \cap L^p(\gamma) \,:\, \Delta f \in L^p(\gamma) \} = W^{2,p}(X,\gamma), \quad \text{for } p \in (1,\infty),$$

where the last identity is a well-known result on Sobolev spaces on Wiener spaces (Meyer's theorem on the boundedness of the second order Riesz transform $\nabla^2\Delta^{-1}$ [Bogachev, 1998, Proposition 5.88]). Higher order Sobolev spaces, such as $W^{2,p}(X,\gamma)$ are defined inductively: for example, $f \in W^{2,p}(X,\gamma)$ if and only if $f \in W^{1,p}(X,\gamma)$ and $\nabla_{\mathcal{H}} f \in W^{1,p}(X,\gamma; H)$. In general, to define Sobolev spaces of maps F taking values in a Hilbert spaces $(E, |\cdot|_E)$, we require that the following extension of the integration by parts formula (12.1) holds, for some $g \in L^p(X,\gamma; H \otimes E)$:

$$\int_X \left\langle F, \left[\frac{d\varphi}{dh} - \varphi\hat{h}\right]\right\rangle_E \gamma = -\int \langle \varphi \otimes h, g\rangle_E \, d\gamma, \quad \text{for every } \varphi \in \mathcal{FC}_b^\infty(X; E).$$

Here and below, we endow the tensor product between two Hilbert spaces with the Hilbert-Schmidt norm.

To prove that L^p-Γ inequality holds, for $p \in (1,\infty]$ we integrate by parts in (11.1), thus

$$\partial_h \mathsf{P}_t f(x) = e^t \int_X \partial_h^y f(e^{-t}x + \sqrt{1-e^{-2t}}y) d\gamma(y) = \frac{e^t}{\sqrt{1-e^{-2t}}} \int_X f(e^{-t}x + \sqrt{1-e^{-2t}}y)\hat{h}(y) d\gamma(y).$$

An application of Hölder inequality gives, for $p \in (1,\infty]$,

$$|\partial_h \mathsf{P}_t f(x)| \le \frac{e^t}{\sqrt{1-e^{-2t}}} \left[\int_X \left| f(e^{-t}x + \sqrt{1-e^{-2t}}y)\right|^p d\gamma(y)\right]^{1/p} \left[\int_X \left|\hat{h}\right|^{p'} d\gamma(y)\right]^{1/p'},$$

passing to the supremum over $h \in \mathcal{H}$, with $|h| \le 1$, gives

$$|\nabla_{\mathcal{H}} \mathsf{P}_t f(x)| \le \frac{e^t C_p^\Gamma}{\sqrt{1-e^{-2t}}} \left[\int_X \left| f(e^{-t}x + \sqrt{1-e^{-2t}}y)\right|^p d\gamma(y)\right]^{1/p},$$

where $(C_p^\Gamma)^{p'} = \int |y|^{p'} \rho(y) dy$, since the law of $\hat{h}$ for $h \in H$ is normal on $\mathbb{R}$, with mean 0 and covariance $|h|^2_{\mathcal{H}}$. Integration over $x \in X$ entails, by Fubini theorem and rotational invariance of product Gaussian measures $\gamma \otimes \gamma$,

$$\|\nabla \mathsf{P}_t f\|_p \le \frac{e^t C_p}{\sqrt{1-e^{-2t}}} \|f\|_p \quad \text{for every } t \in (0,\infty),$$

12.1.2 Derivations and diffusion operators

We let $\mathscr{A} = \mathcal{F}\mathcal{C}_b^\infty(X)$, which is well-known to be dense in every L^p-space and to be stable with respect to the action of P, i.e. (4.2) holds, by Mehler's formula above: in particular we obtain density in $\mathbb{V}^p$ spaces by the results in Section 4.1.

Given a Borel $\mathcal{H}$-valued map field $b = \sum_{i=1}^\infty b^i h^i \in L^1(\gamma; \mathcal{H})$, its associated derivation $\boldsymbol{b}$ is given by

$$\mathscr{A} \ni f \quad \mapsto \quad df(\boldsymbol{b}) := b \cdot \nabla_{\mathcal{H}} f = \sum_{i=1}^\infty b^i \partial_i f.$$

In this setting, $\operatorname{div} \boldsymbol{b}$ is the Gaussian divergence, given by the series

$$\operatorname{div} \boldsymbol{b} = \sum_{i=1}^\infty \partial_i b^i - b^i \hat{h}^i,$$

which defines a distribution, i.e. a linear functional on $\mathscr{A}$.

It is easy to see that $D^{sym}\boldsymbol{b}$ is the symmetric part of the distributional derivative of $\boldsymbol{b}$, given by

$$(D^{sym}\boldsymbol{b}) : h^i \otimes h^j = \frac{1}{2}\left[\partial_i b^j + \partial_j b^i\right], \quad \text{for } i,\, j \geq 1.$$

More generally, given a map a taking values into continuous bilinear forms on $\mathcal{H}$, its associated 2-tensor is defined by

$$\mathscr{A} \times \mathscr{A} \ni (f,g) \quad \mapsto \quad \boldsymbol{a}(f,g) = (a\nabla_{\mathcal{H}} f, \nabla_{\mathcal{H}} g) = \sum_{i,j=1}^\infty a^{i,j}(\partial_i f)(\partial_j g),$$

so symmetry and non-negativity is a consequence (actually, equivalent) of the validity of correspondent properties for $a(x)$, γ-a.e. $x \in X$. A particular case, besides $a = Id$, is that of a being Hilbert-Schmidt valued, but then ellipticity never holds, as Hilbert-Schmidt operators are compact.

Finally, given a Borel vector field $b \in L^1(\gamma; \mathcal{H})$ and a Borel map $a \in L^1(\gamma; \mathcal{H} \otimes \mathcal{H})$ with values in symmetric non-negative Hilbert-Schmidt operators, we introduce the diffusion operator $\mathcal{L} = \mathcal{L}(a, b)$

$$\mathcal{L}f := a : \nabla^2_{\mathcal{H}} f + b\nabla_{\mathcal{H}} f = \sum_{i,j=1}^\infty a^{i,j}\partial_i\partial_j f + \sum_{i=1}^\infty b^i \partial_i f.$$

The integrability assumptions and the choice of $\mathscr{A}$ entail $\mathcal{L}f \in L^1(\mathfrak{m})$, for $f \in \mathscr{A}$. To deal with elliptic operators, we may introduce directly a perturbation of Δ, extending the notation $\mathcal{L}(\sigma, a, b)$ for the diffusion operator

$$\mathcal{L}f := \sigma\Delta f + a : \nabla^2_{\mathcal{H}} f + b \cdot \nabla_{\mathcal{H}} f, \quad \text{for } f \in \mathscr{A},$$

for some non-negative function $\sigma \in L^1(\gamma)$.

Let us notice that for $\mathcal{L}(\sigma, a, b)$ can be written in divergence form when the $\mathcal{H}$-valued distribution $\nabla\sigma + \operatorname{div} a$, given by

$$(\nabla\sigma + \operatorname{div} a)^i = \partial_i \sigma + \sum_{j=1}^\infty \partial_j a^{i,j} - \hat{h}^j a^{i,j}, \quad \text{for } i \geq 1$$

is induced by some vector field in $L^1(\gamma;\mathcal{H})$, so that we rewrite

$$\mathcal{L}f = \operatorname{div}((\sigma Id + a)\nabla_{\mathcal{H}} f) + (b - \operatorname{div} a - \nabla\sigma)\nabla_{\mathcal{H}} f.$$

The distributional divergence of $\mathcal{L}$ is the distribution

$$\operatorname{div}\mathcal{L} = \operatorname{div}[b - \operatorname{div} a] = -\sum_{i,j=1}^{\infty}\left\{\partial_{i,j}a^{i,j} - 2\hat{h}^j\partial_i a^{i,j} + \left[\hat{h}^i\hat{h}^j - \delta_{i,j}\right]a^{i,j}\right\} + \sum_{i=1}^{\infty}\left\{\partial_i b^i - \hat{h}^i b^i\right\}.$$

In the time-extended setting, we naturally let $\mathscr{A} = \mathcal{F}\mathcal{C}_b^\infty((0,T);X)$ we may consider t-dependent Borel families of vector fields $(b_t)_{t\in(0,T)}$, 2-tensors $(a_t)_{t\in(0,T)}$ and associated diffusion operators $(\mathcal{L}_t)_{t\in(0,T)}$.

12.2 Well-posedness results

First, we discuss how the superposition principle specializes in this framework and then focus on existence and uniqueness results for FPE's and associated flows.

12.2.1 The superposition principle

Using the notation introduced in Chapter 7, the choice of $\mathscr{A}^*$ in the Wiener setting causes some issues, since there are at least two eligible distances for which one can hope to lift solutions to FPE's, obtaining continuous processes: the distance induced by the Banach norm on X and the Cameron-Martin (extended) distance $d_{\mathcal{H}}(x,y) := |x-y|_{\mathcal{H}}$, defined as ∞ if $x - y \notin \mathcal{H}$.

In the deterministic case, the picture is rather neat, due to Lemma 7.6. If we let $\mathscr{A}^*$ be the set of smooth cylindrical functions f with $|\nabla_{\mathcal{H}} f| \le 1$, we obtain that any solution to the martingale problem is concentrated on curves that are absolutely continuous with respect to $d_{\mathcal{H}}$. In the general, for diffusions, one cannot expect the curves of the process to be continuous with respect to $d_{\mathcal{H}}$, otherwise, they would always be at finite distance in $\mathcal{H}$ from their initial point: the heat process, which in this case is a Ornstein-Uhlenbeck process, already does not satisfy this property. Indeed, letting $\bar{u} = 1$, one obtains that the trajectory $\gamma(t)$ at time t is almost surely at infinite distance (with respect to $d_{\mathcal{H}}$ from its initial datum $\gamma(0)$:

$$d_{\mathcal{H}}(\gamma(t) - \gamma(0))^2 = \sum_{i=1}^{\infty}\hat{h}_i^2(\gamma(t) - \gamma(0)),$$

which is a series of identically distributed (Gaussian) independent, non-negative and non-null random variables: the law of large numbers gives $d_{\mathcal{H}}(\gamma(t) - \gamma(0))^2 = \infty$ almost surely.

On the other side, continuity of trajectories with respect to the distance induced by the norm on X seem to depend on the Malliavin regularity of $x \mapsto |x|$. Indeed, a natural strategy is to let $\varphi_t := |\gamma(t) - \gamma(0)|$ and argue as in Section 2.2.2, to obtain Hölder regularity for paths. However, this requires some regularity for $\boldsymbol{a}(|\cdot|)$ and $\mathcal{L}|\cdot|$, which may be cause of problems. For example, on the classical Wiener space $X = C([0,S];\mathbb{R})$, where γ is the law of the real-valued Wiener process on $[0,S]$ and $|x| = \sup_{s\in[0,S]}|x(s)|$ is the supremum norm, it is proved in [Trevisan, 2013a] that $x \mapsto |x|$ is Malliavin differentiable only once, and its gradient is a genuine BV vector field, the total variation being singular with respect to

γ. As another example, when X is Hilbert, thus the norm is very smooth, it is easier to provide explicit conditions on the coefficients, ensuring that solutions to the Fokker-Planck equations are lifted to continuous curves on X, see Section 12.4. In conclusion, a weak form of superposition principle, i.e., with respect to a weaker topology, does hold, but we presently lack of sufficient and easy-to-check conditions ensuring continuity of paths with respect to the norm on X, for general norms.

12.2.2 Existence

In the Wiener setting, the existence theorems established in Chapter 9 for weak solutions to FPE's can be strengthened by means of the criterion of Proposition 9.5.

Theorem 12.1 (existence of solutions, Wiener case). *Let $q \in (1,\infty]$, $r \in (1,\infty]$ satisfy $q^{-1}+r^{-1} \le 1$ and let $\mathcal{L} = \mathcal{L}(\sigma, a, b)$ be a diffusion operator, with coefficients*

$$\sigma \in L^1_t(L^{r'}_x), \quad a \in L^1_t(L^{r'}(\gamma; \mathcal{H} \otimes \mathcal{H})), \quad b \in L^1_t(L^{r'}(\gamma; \mathcal{H}))$$

and $\operatorname{div} \mathcal{L}^- \in L^1_t(L^\infty_x)$. Then, for every $\bar u \in L^r(\gamma)$, there exists a L^r-weakly continuous solution u to the FPE (11.6)*, which can be built in such a way that*

i) if $\bar u \ge 0$, then $u_t \ge 0$, for every $t \in [0,T]$, and

ii) if, for some $p \in [1,\infty]$, $\bar u \in L^p(\gamma)$, then $u \in L^\infty_t(L^p(\gamma))$, and

iii) if $\bar u$ is a probability density, then u_t is a probability density for every $t \in [0,T]$.

Proof. Indeed, it is sufficient to consider cylindrical approximations in the form

$$\sigma_N := \mathbb{E}\left[\sigma | e_1^*, \dots e_N^*\right], \quad a_N := \sum_{i,j=1}^N \mathbb{E}\left[a^{i,j} | e_1^*, \dots e_n^*\right] h^i \otimes h^j, \quad b_N := \sum_{i=1}^N \mathbb{E}\left[b^i | e_1^*, \dots e_n^*\right] h^i,$$

and then argue by convolution in the space $(0,T) \times \mathbb{R}^N$, thus providing a sequence for which the criterion quoted above applies (see also the next section for further results on cylindrical approximation). □

In the elliptic case, one can exploit the validity of the logarithmic Sobolev inequality and provide existence as well in case $\operatorname{div}^- \mathcal{L}$ is exponentially integrable, or even when $\mathcal{L}$ is perturbed by adding a vector field c with $|c|^2$ exponentially integrable; for brevity, we omit to write any statement, but see also towards the end of Section 8.3.

12.2.3 Uniqueness

In the deterministic case, it is not difficult to compare our well-posedness results for the continuity equation with those contained in [Ambrosio and Figalli, 2009] and realize that Theorem 10.19 specializes to the uniqueness part of [Ambrosio and Figalli, 2009, Theorem 3.1], with the exception of the case $b \in W^{1,1}(X, \gamma; \mathcal{H})$. As in the Euclidean setting, this seems reachable, with some extra effort, but it is not clear whether the case $b \in BV(X, \gamma; \mathcal{H})$ can been settled by means of this technique. In the next section, we precisely address this issue.

In the general, possibly degenerate, case, Theorem 10.20 provides uniqueness for FPE's associated to diffusion operators $\mathcal{L}(\sigma, a, b)$ where $a = 0$: it seems possible to argue similarly as in Lemma 11.3 and prove well-posedness whenever a belongs to some suitable second-order Sobolev space.

12.3 Uniqueness for BV vector fields

In this section, we take a brief detour from the specialization of the abstract framework and sketch instead the argument originally developed in [Trevisan, 2014a], where we refine in a non-trivial way the strategy by Ambrosio [2004] to show uniqueness for the continuity equation associated to BV vector fields. We briefly introduce all these notions, referring to [Ambrosio et al., 2010] for more details: the space $BV(X,\gamma;\mathcal{H})$ consists of the maps $b \in L\log^{1/2} L\,(X,\gamma;\mathcal{H})$, such that there exists some $\mathcal{H}\otimes\mathcal{H}$-valued measure Du on X, with finite total variation, for which it holds

$$\int_X \left\langle \left[\frac{d\varphi}{dh} - \varphi\hat{h}\right], dDu \right\rangle = -\int \langle \varphi \otimes h, g\rangle_E \, d\gamma, \quad \text{for every } \varphi \in \mathcal{FC}_b^\infty(X;E).$$

In [Ambrosio et al., 2010, Theorem 4.1] one proves the following alternative characterization: $b \in BV(X,\gamma;\mathcal{H})$ if and only if there exists some sequence $(b_n)_{n\geq 1}$ of smooth cylindrical fields such that, as $n \to \infty$, $\|b_n - b\|_1 \to 0$ and $\|\nabla b_n\|_1$ is uniformly bounded (and the smallest bound among all the sequences gives $|Db|(X)$). Actually, the result stated therein refers to the scalar case, but the argument extends easily to vector fields.

Theorem 12.2 (uniqueness for $L_t^\infty(L_x^\infty)$-solutions)**.** *Let $p > 1$, $b \in L_t^1(BV \cap L^p(X,\gamma;\mathcal{H}))$, with $\operatorname{div} b \in L_t^1(L_x^1)$ and let $\bar{u} \in L^\infty(\gamma)$.*

Then, there exists a unique weakly- continuous solution $u = (u_t)_{t\in[0,T]} \in L_t^\infty(L_x^\infty)$ to the continuity equation*

$$\partial_t u_t + \operatorname{div}(b_t u_t) = 0, \quad \text{in } (0,T)\times X, \quad \text{with } u_0 = \bar{u}.$$

Proof. The key idea is to introduce a two parameter family of mollified solutions $u_\rho^\alpha := \mathsf{T}_\alpha^\rho u$, and provide a refined estimate on the commutator, showing that, for every smooth cylindrical function φ, it holds

$$\limsup_{\varepsilon\downarrow 0} \int_X \left|\beta'(u_\rho^\alpha)[\mathsf{T}_\alpha^\rho, b\nabla]\varphi\right| d\gamma \leq \left\|\beta'\right\|_\infty \int_X |\varphi|\, \Lambda_\rho d\,|Db|\,, \tag{12.2}$$

where $|Db|$ is the total variation measure associated to the BV field, and Λ_ρ is some explicit density, depending on the choice of the mollifier. This entails that the distribution

$$\partial_t \beta(u_t) + \operatorname{div}(b\beta(u_t)) - \left[\beta(u_t) - \beta'(u_t)\,u_t\right]\operatorname{div} b = \sigma_t$$

is actually a measure (the so-called defect measure), but the same expression shows that does not depend on ρ. To conclude, it is sufficient to show that

$$|\sigma| \leq \bigwedge_\rho \Lambda_\rho\,|Db| = 0, \tag{12.3}$$

so that, letting $\beta(z) = |z|$ (better, some suitable approximation of it) and integrating, uniqueness is settled. □

The optimization step (12.3) is based on a Wiener space analogue of an argument due to Alberti, see [Ambrosio, 2008, Lemma 35], and we skip its proof, referring to [Trevisan, 2014a] for a detailed exposition. Before we address the proof of the crucial step (12.2), we compare our technique with the original approach by Ambrosio [2004]. In the Euclidean setting, one argues by means of two different commutator estimates: anisotropic estimates, which are rather good in the regions where the measure-derivative Db is mostly singular with respect to the Lebesgue measure, and isotropic estimates, which are useful instead in the regions where the derivative is mostly absolutely continuous. Then, an optimization procedure on the choice of approximations gives the result. In the Wiener setting, a direct implementation of this method fails, because of error terms depending on the dimension of the space. The novelty consists of establishing a refined anisotropic estimate, namely (12.2), which is well-behaved at every point and, after an optimization procedure, turns out to be sufficient to conclude.

12.3.1 Cylindrical approximations

We establish two propositions instrumental to the proof of (12.2). The first one is a slight generalization of an argument appearing in the proof of [Ambrosio and Figalli, 2009, Proposition 3.5].

We use the notation introduced in Section 12.1, in particular we let $(h_k)_k \subseteq \mathcal{H}$ be a complete orthonormal system of the form $h_k = Qe_k^*$, for $e_k^* \in X^*$, we let π_N be the projection on the linear span of $h_1, \ldots, h_N$ and $\mathcal{F}_N$ be the σ-algebra generated by $e_1^*, \ldots e_N^*$, for $N \geq 1$.

Moreover, the map $x \mapsto (\pi_N(x), x - \pi_N(x))$ induces decompositions $X = Im\pi_N \oplus Ker\pi_N$ and $\mathcal{H} = Im\pi_N \oplus Im\pi_N^\perp$. Recall that we tacitly identify $Im\pi_N = \mathbb{R}^N$ via $h_i \mapsto e_i$. The same map induces a decomposition $\gamma = \gamma_N \otimes \gamma_N^\perp$, where γ_N is the standard N-dimensional normal law on $\mathbb{R}^N$ and $\gamma_N^\perp$ is a non-degenerate Gaussian measure on $Ker\pi_N$, with Cameron-Martin space given by $Im\pi_N^\perp$.

Proposition 12.3. *Let $b \in BV(\gamma; \mathcal{H})$ with $\operatorname{div} b \in L^1(\gamma)$ and define*

$$b_N := \mathbb{E}\left[\pi_N b \,|\, \mathcal{F}_N\right], \quad \text{for } N \geq 1. \tag{12.4}$$

Then, b_N is a cylindrical BV vector field, with

$$Db_N = \left[(\pi_N)_\sharp (\pi_N Db\, \pi_N)\right] \otimes \gamma_N^\perp \quad \text{and} \quad \operatorname{div} b_N = \mathbb{E}\left[\operatorname{div} b \,|\, \mathcal{F}_N\right]. \tag{12.5}$$

Moreover, it holds

$$\lim_{N\to\infty} \|b_N - b\|_1 + \|\operatorname{div} b_N - \operatorname{div} b\|_1 = 0.$$

Proof. The second statement follows from the second identity in (12.5), and by density of cylindrical fields and uniform boundedness of the operators involved. Thus, it is enough to focus on (12.5): since we argue at fixed $N \geq 1$, we let, with a slight abuse of notation, $\pi := \pi_N$ and $\mathbb{E}[\cdot] := \mathbb{E}[\cdot \,|\, \mathcal{F}_N]$.

Notice that $b_N \in L^1(\gamma; \mathcal{D})$ because projection and conditional expectation are contractions. The thesis follows from rather algebraic identities, arguing in duality with functions $\varphi \in \mathcal{F}\mathcal{C}_b^\infty(X)$, using symmetry of π and $\mathbb{E}$, and the commutation relation

$$\pi\mathbb{E}\left[\nabla\varphi\right] = \mathbb{E}\left[\pi\nabla\varphi\right] = \nabla\mathbb{E}\left[\varphi\right].$$

It holds

$$\begin{aligned}\int \varphi \operatorname{div} b_N d\gamma &= -\int \langle \nabla\varphi, b_N\rangle\, d\gamma = -\int \langle \pi\mathbb{E}\left[\nabla\varphi\right], b\rangle\, d\gamma \\ &= -\int \langle \nabla\mathbb{E}\left[\varphi\right], b\rangle\, d\gamma = \int \varphi\, \mathbb{E}\left[\operatorname{div} b\right] d\gamma,\end{aligned}$$

which entails the second identity in (12.5).

Analogous identities hold in duality with $\mathcal{H}\otimes\mathcal{H}$ smooth maps, acting with $\pi\otimes I$ in place of π, where I denotes the identity operator. From this, one deduces first the identity $D(\pi b) = (\pi\otimes I)(Db)$ and then conclude that, for every smooth cylindrical $\mathcal{H}\otimes\mathcal{H}$ map, Φ, it holds

$$\int \langle \operatorname{div}\Phi, b_N\rangle = \int \langle \mathbb{E}\left[\operatorname{div}\Phi\right], \pi b\rangle = \int \langle (\pi\otimes\pi)\boldsymbol{E}\left[\Phi\right], dDb\rangle \,,$$

which gives the thesis. □

The second result is actually purely measure-theoretical: its proof is based on disintegration of measures and Jensen's inequality. Notice that it proves and generalizes the inequality

$$|Db_N|(X) \leq |Db|(X)\,,$$

where we let b_N be as in (12.4).

Proposition 12.4. *Let $b\in BV(\gamma;\mathcal{H})$ and let b_N be as in* (12.4). *Assume that*

$$f : X\times(\mathcal{H}\otimes\mathcal{H}) \to [0,\infty]$$

is Borel, positively homogeneous and convex in the second variable, i.e. at every $x\in X$, $f(x,\cdot)$ is convex on $\mathcal{H}\otimes\mathcal{H}$ and positively homogeneous.

Then, it holds

$$\int f\left(\pi_N, \frac{Db_N}{|Db_N|}\right) d\,|Db_N| \leq \int f\left(\pi_N, \pi_N\frac{Db}{|Db|}\pi_N\right) d\,|Db|\,. \tag{12.6}$$

Proof. Again, we make implicit the dependence upon $N\geq 1$ by writing $\pi := \pi_N$. We let also $\mu := \pi Db\pi$, $\nu = Db_N$ and $\rho = \gamma_N^{\perp}$, the first identity in (12.5) simlply reads as $\nu = (\pi_\sharp\mu)\otimes\rho$, and this factorization is actually all the information we need from the case we are considering.

Indeed, the total variation and the polar decomposition of ν factorize as

$$|\nu|(dx,dy) = |\pi_\sharp\mu|(dx)\otimes\rho(dy) \quad\text{and}\quad \frac{\nu}{|\nu|}(x,y) = \frac{\pi_\sharp\mu}{|\pi_\sharp\mu|}(x)\,,$$

thus, the left hand side in (12.6) reads as

$$\int f\left(x, \frac{\nu}{|\nu|}(x,y)\right) d\,|\nu|(x,y) = \int f\left(x, \frac{\pi_\sharp\mu}{|\pi_\sharp\mu|}(x)\right)|\pi_\sharp\mu|(dx)\,.$$

Since $|\pi_\sharp\mu| \leq \pi_\sharp|\mu|$, it holds

$$\frac{\pi_\sharp\mu}{\pi_\sharp|\mu|} = \frac{\pi_\sharp\mu}{|\pi_\sharp\mu|}\frac{|\pi_\sharp\mu|}{\pi_\sharp|\mu|},$$

and by positive homogeneity of f we obtain

$$\int f\left(x, \frac{\pi_\sharp\mu}{|\pi_\sharp\mu|}(x)\right)|\pi_\sharp\mu|(dx) = \int f\left(x, \frac{\pi_\sharp\mu}{\pi_\sharp|\mu|}(x)\right)\pi_\sharp|\mu|(dx)\,.$$

We next disintegrate $|\mu|$ with respect to π and apply Jensen's inequality. More precisely, since X is a Banach space, there exists a probability kernel $(N_x)_{x\in X}$ such that, for every bounded Borel function $g : X\times X\mapsto\mathbb{R}$ it holds

$$\int g(\pi(z),z)\,d\,|\mu|(z) = \int \pi_\sharp|\mu|(dx)\int g(x,y)\,N_x(dy)\,.$$

Moreover, if we let $\sigma |\mu| = \mu$ be its polar decomposition, using $g(z) = h(\pi(z))\sigma(z)$ above, we obtain

$$\frac{\pi_\sharp \mu}{\pi_\sharp |\mu|}(x) = \int \sigma(y) N_x(dy), \quad \pi_\sharp |\mu|\text{-a.e. } x \in X.$$

By Jensen's inequality, it holds

$$f\left(x, \frac{\pi_\sharp \mu}{\pi_\sharp |\mu|}(x)\right) \leq \int f(x, \sigma(y)) N_x(dy), \quad \pi_\sharp |\mu|\text{-a.e. } x \in X,$$

and integrating with respect to $\pi_\sharp |\mu|$, the right hand side above gives

$$\int f(\pi(x), \sigma(x)) \, d|\mu|(x) = \int f\left(\pi(z), \frac{\mu}{|\mu|}(z)\right) d|\mu|(z).$$

To conclude, we integrate also with respect to $\rho(dy)$ and use again the positive homogeneity of f, together with the identities

$$\frac{\pi Db \pi}{|\pi Db \pi|} \frac{|\pi Db \pi|}{|Db|} = \frac{\pi Db \pi}{|Db|} = \pi \frac{Db}{|Db|} \pi.$$

$\square$

12.3.2 Proof of the refined anisotropic estimate (12.8)

For simplicity of notation, we address the time-independent case only, and omit to write the variable $t \in (0, T)$ in all what follows. We first describe our family of mollifiers. For any smooth cylindrical function ρ, such that $\rho\gamma$ is a probability measure, we introduce a modified Ornstein-Uhlenbeck operator, letting

$$\mathsf{T}_\alpha^\rho \varphi(x) = \int \varphi(x_\alpha) \rho(y) \gamma(dy),$$

where we write, here and in what follows,

$$x_\alpha = e^{-\alpha} x + \sqrt{1 - e^{-2\alpha}} y, \quad y_\alpha = -\sqrt{1 - e^{-2\alpha}} x + e^{-\alpha} y.$$

Our aim is to establish (12.2), with

$$\Lambda_\rho(x) = \int_X \left| \operatorname{div}_y \left(\hat{M}_x(y) \rho(y) \right) \right| d\gamma(y), \tag{12.7}$$

where $M_x := (Db / |Db|)(x)$ is the polar decomposition of Db with respect to its total variation measure, and $\hat{M}$ denotes the vector field associated to M, i.e. roughly, the linear operator associated to the Hilbert-Schmidt operator M, see [Trevisan, 2014a] for more details. The precise expression in (12.7) is crucial for the optimization step, which we do not address here: let us however recall that the strategy is to choose ρ to approximate an invariant measure for the exponential flow associated to $\hat{M}_x$, at $|Db|$-a.e. $x \in X$, letting e.g.

$$\rho_T := \frac{1}{T} \int_0^T \exp(t \hat{M}_x)_\sharp \gamma, \quad \text{in the limit } T \to \infty.$$

The technical problem is then to check that these objects are well defined, but this can be addressed with minor difficulties, see Trevisan [2014a].

To prove (12.8) with (12.7), we argue as follows. First, we let $\alpha > 0$ be fixed and let b to be a cylindrical smooth vector field, and obtain an estimate for $[\mathsf{T}_\alpha^\rho, b \cdot \nabla]$ in terms of Db. Then, still at fixed $\alpha > 0$, we extend its validity to general BV vector fields. Finally, we let $\alpha \to 0$ and conclude.

For simplicity, but without any loss of generality, we assume that $\|\beta'\|_\infty \leq 1$ and we omit to write ρ.

Step 1 (fixed $\alpha > 0$ and cylindrical smooth b)**.** We aim at obtaining estimate where three terms appear, two of them being negligible as $\alpha \to 0$, and the third leading to the result, in the limit. Since Sobolev and BV spaces are well-behaved with respect to linear push-forwards, without any loss of generality, we argue in same finite-dimensional Gaussian space $(\mathbb{R}^N, \gamma_N)$.

Performing some integration by parts and change of measures, one obtains the inequality

$$\int |[\mathsf{T}_\alpha^\rho, b \cdot \nabla]\varphi| \, d\gamma \leq \int |\varphi| (x_\alpha) \left| \operatorname{div}_y \left(\frac{b(x) - e^\alpha b(x_\alpha)}{C_\alpha} \rho(y) \right) \right| dxdy, \tag{12.8}$$

where we write dx and dy, here and in what follows, for the Gaussian measure γ_N on $\mathbb{R}^N$ (not Lebesgue measure).

We add subtract $b(x_\alpha)$ in the difference and we split

$$\int |\varphi| (x_\alpha) \left\{ \frac{e^\alpha - 1}{C_\alpha} \left| \operatorname{div}_y (b(x_\alpha) \rho(y)) \right| + \left| \operatorname{div}_y \left(\frac{b(x_\alpha) - b(x)}{C_\alpha} \rho(y) \right) \right| \right\} dxdy.$$

The first term in the sum above gives the an error term which is smaller than

$$\sqrt{\alpha} \, \|\varphi\|_\infty \left[\|b\|_1 \|\nabla \rho\|_\infty + \|\operatorname{div} b\|_1 \|\rho\|_\infty \right], \tag{12.9}$$

noticing that $C_\alpha \leq C\sqrt{\alpha}$, for $\alpha \in (0, 1]$, and some absolute constant C.

We focus then on what is left, namely the expression

$$\int |\varphi| (x_\alpha) \left| \operatorname{div}_y \left(\frac{b(x_\alpha) - b(x)}{C_\alpha} \rho(y) \right) \right| dxdy. \tag{12.10}$$

We interpolate as follows:

$$b(x_\alpha) - b(x) = \int_0^\alpha \frac{d}{ds} b(x_s) \, ds = \int_0^\alpha Db(x_s) \, y_s \frac{ds}{C_s},$$

using the identity $\frac{d}{ds} x_s = y_s / C_s$. For brevity, we introduced the notation

$$⨍_0^\alpha f(s) = \frac{1}{C_\alpha} \int_0^\alpha f(s) \frac{ds}{C_s},$$

justified by the fact that, as $\alpha \to 0$,

$$\frac{1}{C_\alpha} \int_0^\alpha \frac{ds}{C_s} \to 1. \tag{12.11}$$

Exchanging divergence and integration, we obtain

$$\begin{aligned}
\operatorname{div}_y\left(\frac{b(x_\alpha)-b(x)}{C_\alpha}\rho(y)\right) &= \fint_0^\alpha \operatorname{div}_y\left(Db(x_s)\,y_s\rho(y)\right) \\
&= \fint_0^\alpha \left[\operatorname{div}_y\left(Db(x_s)\,y_s\right)\rho(y) + \langle Db(x_s)\,y_s, \nabla\rho(y)\rangle\right].
\end{aligned} \tag{12.12}$$

Let us consider the first term in the sum above: write

$$v(x,y) := Db(x)\,y = \partial_y b(x),$$

and for $s \in (0,\alpha)$, an explicit computation gives

$$\operatorname{div}_y\left(Db(x_s)\,y_s\right) = \sqrt{1-e^{-2s}}\left[\operatorname{div}_x(v)\right](x_s,y_s) + e^{-s}\left[\operatorname{div}_y(v)\right](x_s,y_s) \tag{12.13}$$

Since the term $\operatorname{div}_x(v)$ involves further spatial derivatives of b, the following identity, which can be obtained by inspection in coordinates, plays a crucial role:

$$\operatorname{div}_x(v)(x_s,y_s) = C_s\frac{d}{ds}\operatorname{div} b(x_s) + \hat{b}(x_s,y_s),$$

where we used the notation $\hat{b}$ for the map

$$(x,y) \mapsto \hat{b}(x,y) = \sum_{i=1}^N x^i b^i(y).$$

This allows us to integrate by parts,

$$\fint_0^\alpha \sqrt{1-e^{-2s}}\operatorname{div}_x(v)(x_s,y_s) = \left[e^{-\alpha}\operatorname{div} b(x_\alpha) - \fint_0^\alpha \operatorname{div} b(x_s)\,e^{-s}\right] + \fint_0^\alpha \sqrt{1-e^{-2s}}\hat{b}(x_s,y_s),$$

since $\frac{d}{ds}\sqrt{1-e^{-2s}} = e^{-s}/C_s$.

Thanks to these computations we separate from (12.10) another error term, smaller than

$$\|\varphi\|_\infty \|\rho\|_\infty \left[\int \left|e^{-\alpha}\operatorname{div} b(x_\alpha) - \fint_0^\alpha \operatorname{div} b(x_s)\,e^{-s}\right| dxdy + \frac{\alpha}{2C_\alpha}\|b\|_1\right].$$

The integrand above is a linear expression in $\operatorname{div} b$, which reminds of some averaged Ornstein-Uhlenbeck operator. By rotational invariance of Gaussian measures and by (12.11) above, its L^1 norm is bounded by some absolute constant, uniformly in $\alpha \in (0,1]$. By density of smooth functions in L^1, it defines therefore a family of continuous operators G_α and we estimate

$$\|\varphi\|_\infty \|\rho\|_\infty \left[\|\mathsf{G}_\alpha(\operatorname{div} b)\|_1 + \frac{\alpha}{C_\alpha}\|b\|_1\right]. \tag{12.14}$$

The following expression contains precisely what remains to be estimated from (12.10), i.e., the second term in the second line of (12.12) and the second term in the right hand side of 12.13,

$$\int |\varphi|(x_\alpha) \fint_0^\alpha \left|e^{-s}\left[\operatorname{div}_y(v)\right](x_s,y_s)\,\rho(y) + \langle v(x_s,y_s), \nabla\rho(y)\rangle\right| dxdy.$$

Once we exchange integration and perform a change of variables mapping x_α to $x_{\alpha-s}$, y_α to $y_{\alpha-s}$ we rewrite this expression in a way that easily easily extends to the BV case, i.e.,

$$\int f\left(x, \frac{Db}{|Db|}(x)\right)|Db|(dx), \tag{12.15}$$

where

$$f(x, M) = \fint_0^\alpha \int |\varphi|(x_{\alpha-s})\left|e^{-s}\operatorname{div}_y(My)\rho(y^s) + \langle My, (\nabla\rho)(y^s)\rangle\right| dy.$$

Step 2 (fixed $\alpha > 0$ and BV vector field b)**.** The expression (12.8) is smaller than the sum of three terms, namely (12.9), (12.14) and (12.15). We extend the validity of this fact first to cylindrical BV fields, and then to the general case.

Under the assumption that b is cylindrical, we already showed that everything is reduced to a computation in $\mathbb{R}^N$, thus it is possible to find smooth cylindrical fields $(b_n)_n$ such that, as $n \to \infty$,

$$\|b_n - b\|_1 \to 0, \quad \|\operatorname{div} b_n - \operatorname{div} b\|_1 \to 0, \quad |Db_n|(X) \to |Db|(X)$$

and the sequence $(Db_n)_n$ weakly-* converges to Db, choosing e.g. a sequence extracted from the usual Ornstein-Uhlenbeck semigroup for infinitesimal times provides such a sequence, see [Ambrosio et al., 2010]. The left hand side in (12.8), together with the first and second error terms (12.9), (12.14) pass to the limit with respect to this convergence. The only trouble might be caused by (12.15), and at this point we apply Reshetnyak continuity theorem, see e.g. [Ambrosio et al., 2000, Theorem 2.39].

We now extend the estimate to handle general BV fields, letting $(b_N)_{N\geq 1}$ as in (12.4). Again, (12.8), together with the first and second error terms (12.9), (12.14), pass to the limit essentially because of Proposition 12.3. To handle the term (12.15), we prove first that for every N sufficiently large, so that both φ and ρ are N-cylindrical, it holds

$$\int f\left(x, \frac{Db_N}{|Db_N|}(x)\right) d|Db_N|(x) \leq \int f\left(x, \frac{Db}{|Db|}(x)\right) d|Db|(x).$$

This follows from Proposition 12.4. Indeed, by direct inspection, the left hand side above coincides with

$$\int f_N\left(\pi_N(x), \frac{Db_N}{|Db_N|}(x)\right) d|Db_N|(x),$$

where

$$f_N(x, M) = \fint_0^\alpha \int |\varphi|(x_{\alpha-s})\left|e^{-s}\operatorname{div}_y(My)\rho(y^s) + \langle My, (\nabla\rho)(y^s)\rangle\right| d\gamma_N(y),$$

which is positively homogeneous and convex in the second variable. We get therefore

$$\int f\left(x, \frac{Db_N}{|Db_N|}(x)\right) d|Db_N|(x) \leq \int f_N\left(x, \pi_N\frac{Db}{|Db|}(x)\pi_N\right) d|Db|(x).$$

We finally recognize that

$$\pi_N M \pi_N(y) = \mathbb{E}\left[\pi_N \hat{M} \mid \pi_N = y\right],$$

and so, again by Proposition 12.3, applied this time to $\hat{M}$, we obtain

$$\operatorname{div}_y\left(\pi_N M \pi_N(y)\right) = \mathbb{E}\left[\operatorname{div}_y \hat{M} \mid \pi_N = y\right].$$

Combining these identities in the expression for f_N and recalling that φ and ρ are N-cylindrical we conclude, since the conditional expectation is a contraction in $L^1(\gamma)$.

Step 3 (limit as $\alpha \downarrow 0$)**.** The first error term (12.9) is clearly infinitesimal, but also the term (12.14), since $\|\mathsf{G}_\alpha(\operatorname{div} b)\|_1 \to 0$ by uniform boundedness of G_α and the fact that convergence to 0 holds for smooth cylindrical vector fields.

Finally, the term (12.15) converges to

$$\int |\varphi|(x)\left[\int \left|\operatorname{div}_y\left(\hat{M}_x(y)\right)\rho(y) + \left\langle \hat{M}_x(y), \nabla\rho(y)\right\rangle\right| d\gamma(y)\right] d|Db|(x).$$

Indeed, the integrands converge because φ and ρ are cylindrical smooth and, for $p \in]1,\infty[$, we have the bound, for some constant c_p depending on p only,

$$f(x, M) \leq c_p \|\varphi\|_\infty \left(\|\rho\|_p + \|\nabla\rho\|_p\right)|M|.$$

and $|M| \leq 1$, as assured by the polar decomposition theorem.

12.4 Gaussian Hilbert spaces

In this section we further specialize the setting above, letting $X = H$ be a separable Hilbert space, with norm $|\cdot|$, thus γ is a Gaussian centered and nondegenerate measure in H. For the sake of brevity, we do not provide a complete discussion of all the results, but focus on two aspects: the superposition principle and uniqueness results for FPE's.

12.4.1 Basic setup and diffusion operators

By identifying $H = H^*$ via the Riesz isomorphism induced by the norm, the covariance operator $Q : H \to H$ is a symmetric positive trace class operator, thus compact. In this setting the Cameron-Martin space reads as $\mathcal{H} = Q^{1/2}H$, with the norm $|h|_{\mathcal{H}} = \left|Q^{-1/2}h\right|$.

We let $(e_i) \subset H$ be an orthonormal basis of H consisting of eigenvectors of Q, with eigenvalues (λ_i), i.e. $Qe_i = \lambda_i e_i$ for every $i \geq 1$: in this setting, we define the class of smooth cylindrical functions $\mathcal{FC}_b^\infty(H)$ as those functions $f : X \to \mathbb{R}$ of the form $f(x) = \varphi(\langle e_i, x\rangle, \ldots \langle e_n, x\rangle)$, with $\varphi : \mathbb{R}^n \to \mathbb{R}$ smooth and bounded. Given $f \in \mathcal{FC}_b^\infty(H)$, from its Fréchet derivative $df : H \to H^*$ we introduce $\nabla f : H \to H$ via $H = H^*$, in coordinates:

$$\nabla f(x) = \sum_i \partial_i f(x) e_i, \quad \text{where } \partial_i f(x) = \lim_{\varepsilon \to 0} \frac{f(x + \varepsilon e_i) - f(x)}{\varepsilon}.$$

To recover the abstract setting of the previous section, notice that family $h_i = \lambda_i^{1/2} e_i$ is an orthonormal basis of $\mathcal{H}$ and that $\partial/\partial h_i = \lambda_i^{-1/2}\partial_i$, thus it holds $Q\nabla f = \nabla_{\mathcal{H}} f$.

For $\alpha \in \mathbb{R}$, we introduce the form

$$\mathcal{E}^\alpha(f) = \int_X \left|Q^{(1-\alpha)/2}\nabla f\right|^2 d\gamma, \qquad f \in \mathcal{FC}_b^\infty(H),$$

which is closable: its domain is the space $W^{1,2}_\alpha(H,\gamma)$, see [Da Prato, 2004, Chapters 1 and 2] for more details. Evidently, we recover (3.1), with $\Gamma(f) = \sum_i \lambda_i^{1-\alpha}\left|\partial_i f\right|^2$. Notice that the associate distance is the one induced by the norm $|Q^{(\alpha-1)/2}x|$, which is extended if and only if $\alpha < 1$.

The associated semigroup can be still be seen as the transition semigroup of an infinite dimensional Ornstein-Uhlenbeck process, with Laplacian Δ_α given by

$$\Delta_\alpha f(x) = \operatorname{Tr}\left[Q^{1-\alpha}D^2 f(x)\right] - \left\langle x, Q^{-\alpha}\nabla f(x)\right\rangle, \quad \text{for } f \in \mathcal{FC}_b^\infty(H).$$

It can be shown that the abstract curvature bound $\mathsf{BE}_2(1,\infty)$ holds [Da Prato, 2004, Proposition 2.60], entailing the validity of the L^p-Γ inequality (see the next Chapter for a more detailed discussion). We let $\mathscr{A} = \mathcal{FC}_b^\infty(H)$, which is dense in every $L^p(\mathfrak{m})$ space and satisfies (4.2), thus obtaining density results in $\mathbb{V}^p$ ($p \in [1,\infty)$) by Remark 4.4 and Proposition 4.2

For $\alpha = 0$, we recover the abstract Wiener space setting discussed above, while for $\alpha = 1$ we obtain the "genuine" Gaussian-Hilbert setting.

Given $b : H \to H$, $b = \sum_i b_i e_i$ Borel, we consider the map

$$\mathcal{FC}_b^\infty(H) \ni f \quad \mapsto \quad df(\boldsymbol{b}) := \langle b, \nabla f\rangle_H = \sum_i b_i \partial_i f.$$

If $\left|Q^{(\alpha-1)/2}b\right| \in L^q(H,\gamma)$ for some $q \in [1,\infty]$, then $\boldsymbol{b}$ is a well-defined derivation, with $|\boldsymbol{b}| \leq \left|Q^{(\alpha-1)/2}b\right|$.

The Cameron-Martin theorem entails an integration by parts formula [Da Prato, 2004, Theorem 1.4 and Lemma 1.5] that reads in our notation as

$$\operatorname{div}\boldsymbol{e}_i(x) = -\frac{\langle e_i, x\rangle}{\lambda_i}, \quad \text{where } df(\boldsymbol{e}_i) = \partial_i f.$$

On smooth "cylindrical" fields $b = \sum_i^n b_i e_i$, this gives

$$\operatorname{div}\boldsymbol{b}(x) = \sum_i \partial_i b_i(x) - \frac{\langle e_i, x\rangle}{\lambda_i} b_i,$$

where the series reduces to a finite sum. Notice that the expression does not depend on α but only on γ, in agreement with the notion of divergence as dual to derivation.

Notice also that vector fields do not need to take values in $\mathcal{H}$ and actually our results hold even for some classes of fields not taking values in $\mathcal{H}$ (although their divergence must still satisfy some bounds).

Arguing on smooth cylindrical functions, we see that

$$\int D^{sym}\boldsymbol{b}(u,f)d\gamma = \int \sum_{i,j}\frac{1}{2}\left[\left(\frac{\lambda_i}{\lambda_j}\right)^{(1-\alpha)/2}\partial_i b_j + \left(\frac{\lambda_j}{\lambda_i}\right)^{(1-\alpha)/2}\partial_j b_i\right]\left(\lambda_i^{(1-\alpha)/2}\partial_i u\right)\left(\lambda_j^{(1-\alpha)/2}\partial_j f\right)d\gamma, \tag{12.16}$$

thus our bound on $D^{sym}\boldsymbol{b}$ is implied by an L^q bound of the Hilbert-Schmidt norm of the expression is square brackets above.

2-tensor s are be defined by maps on H taking values into bilinear functionals on H and we let $a^{i,j} := a(e_i, e_j)$. A case of particular interest is that of a being non-negative and of trace class, i.e. it holds $\sum_{i\geq 1} a^{i,i} < \infty$.

Similarly as in the general case, we let $\mathcal{L}(\sigma, a, b)$ be the diffusion operator

$$\mathcal{L}f := \sigma\Delta_1 f + a : \nabla^2 f + \langle b, \nabla f\rangle_H, \quad \text{for } f \in \mathcal{F}C_b^\infty(H).$$

12.4.2 The superposition principle

As we already remarked, in the case of infinite dimensional diffusions, it is not completely clear whether the superposition principle lifts solutions to FPE's to continuous processes on H, as the general arguments provides processes with values on a larger space, namely, the completion of H with respect to a weaker distance. It is interesting to provide sufficient conditions on the diffusion operator so that the process has continuous paths, with respect to the natural norm on H, and in the Hilbert setting these are somewhat explicit.

Indeed, following the Kolmogorov-type argument from Section 2.2.2, the crucial quantity to consider is $\mathcal{L}(|\cdot|_H^2)$, which we aim at expressing in terms of a, and b. By linearity, we focus on two separate cases. When $\mathcal{L} = a : \nabla^2 f$, we have for every $i \geq 1$, $\mathcal{L}(\langle e_i, \cdot\rangle^2) = a^{i,i}$, thus if a takes values in the space of (bounded and) trace class operators, and belongs to $L^q(\gamma)$, for some $q \in (1, \infty]$, one deduces that the superposition principle for metric measure spaces lifts solutions belonging to $L^\infty_t(L^r_x)$, where $r \in (1, \infty]$, $q^{-1} + r^{-1} \leq 1$, to solutions to the martingale problem which have Hölder continuous paths in H. A similar fact holds when $\mathcal{L} = b \cdot \nabla f$, whenever b belongs to $L^q(\gamma)$, for some $q > 1$.

12.4.3 Uniqueness results

As a consequence of the specializations above, Theorem 10.19 provides uniqueness for solutions to the continuity equation associated to vector fields b with $\operatorname{div} b$, $|D^{sym} b| \in L^1_t(L^q_x)$. Comparing our setting with that in [Da Prato et al., 2014] we recognize Theorem 2.3 therein as a consequence our uniqueness result.

We end this section considering a field b taking values outside $\mathcal{H}$, to which our theory applies (although well-posedness was already shown in [Mayer-Wolf and Zakai, 2005]). Assume that that each eigenvalue of Q admits a two-dimensional eigenspace thus, slightly changing the notation, we write $(e_i, \tilde{e}_i)$ for an orthonormal basis of H consisting of eigenvectors of Q. We let

$$b = \sum_{i=1}^{\infty} \lambda_i^{1/2} \left[(\operatorname{div} \tilde{\boldsymbol{e}}_i) e_i - (\operatorname{div} \boldsymbol{e}_i) \tilde{e}_i \right], \quad \text{thus} \quad \int \left| Q^{(\alpha-1)/2} b \right|^2 d\gamma = \sum_{i=1}^{\infty} \lambda_i^{\alpha}.$$

The series above converges if $\alpha = 1$, and it does not if $\alpha = 0$. Since $(\operatorname{div} \boldsymbol{e}_i, \operatorname{div} \tilde{\boldsymbol{e}}_i)_i$ are independent, Kolmogorov's 0-1 law entails that b is well defined as an H-valued map, but $b(x) \notin \mathcal{H}$ for γ-a.e. $x \in H$. The derivation $\boldsymbol{b}$ is therefore well-defined if $\alpha = 1$, and $|\boldsymbol{b}| \in L^2(\mathfrak{m})$. From its structure and (12.16), both its divergence and its deformation are seen to be identically 0, thus our results apply.

Chapter 13

Metric measure spaces with curvature bounds

In this chapter, which follows closely [Ambrosio and Trevisan, 2014, §6 and §9F], we describe how the abstract theory specializes, in a non-trivial way, to suitable classes of metric measure spaces, namely those enjoying an infinitesimal Riemannian structure and uniform lower bounds on the Ricci curvature, called $\mathsf{RCD}(K, \infty)$ spaces, recently introduced in Ambrosio et al. [2014b].

In this (and more general) frameworks, some duality appears, because these spaces can be studied both from a "Eulerian" point of view, emphasizing the role of functions and their calculus, and from a "Lagrangian" one, focusing instead on measures and their distances, measured by means via optimal transport. It turns out that in the case infinitesimal Riemannian spaces, the former is connected to Γ-calculus and Bakry-Émery curvature condition $\mathsf{BE}(\mathsf{K}, \infty)$, whose development dates back to [Bakry and Émery, 1984] and [Bakry, 1985])see the already quoted [Bakry et al., 2014] for a more recent exposition). The latter point of view developed more recently, starting from the class of $\mathsf{CD}(K, \infty)$ metric measure spaces, introduced and deeply studied in [Lott and Villani, 2009], [Sturm, 2006a], [Sturm, 2006b]. Subsequently, connections with the theory of Dirichlet forms and gave rise to a series of works, [Ambrosio et al., 2012], [Savaré, 2014] and [Ambrosio et al., 2014c]. For a brief introduction to the setting and its notation, we refer to Sections 4.1 and 4.2 in Savaré [2014], and in particular to Theorem 4.1 therein, which collects non-trivial equivalences among different conditions.

We proceed as follows. In Section 13.1 we introduce the "Eulerian" curvature condition, showing that it entails many of the assumptions that we imposed in the previous parts for non-triviality of our results on Fokker-Planck equations: in particular, the validity of L^p-Γ inequalities and existence of derivations with some bound on their deformation. In Section 13.2, we specialize on the "Lagrangian" viewpoint, showing that, in the case of continuity equations, L^∞-regular solutions to the deterministic martingale problem provide instances of so-called 2-plans. For simplicity, we manly focus on the deterministic case: the case of diffusion processes can be clearly reached by the techniques developed throughout all this thesis, and will be object of future investigations.

13.1 $\mathsf{BE}(K,\infty)$ spaces

In this section we add to the basic setting (3.1) a suitable curvature condition, and see the implication of this assumption on the structural conditions made in the previous parts of the thesis.

In the sequel K denotes a generic but fixed real number, and I_K denotes the real function

$$\mathrm{I}_K(t) := \int_0^t \mathrm{e}^{Kr}dr = \begin{cases} \frac{1}{K}(\mathrm{e}^{Kt}-1) & \text{if } K \neq 0, \\ t & \text{if } K = 0. \end{cases}$$

Definition 13.1 (Bakry-Émery conditions)**.** *We say that* $\mathsf{BE}_2(K,\infty)$ *holds if*

$$\Gamma\left(\mathsf{P}_t f\right) \leq \mathrm{e}^{-2Kt}\,\mathsf{P}_t\big(\Gamma\left(f\right)\big) \quad \mathfrak{m}\text{-}a.e.\ in\ X,\ for\ every\ f \in \mathbb{V},\ t \geq 0. \tag{13.1}$$

We say that $\mathsf{BE}_1(K,\infty)$ *holds if*

$$\sqrt{\Gamma\left(\mathsf{P}_t f\right)} \leq \mathrm{e}^{-Kt}\,\mathsf{P}_t\big(\sqrt{\Gamma\left(f\right)}\big) \quad \mathfrak{m}\text{-}a.e.\ in\ X,\ for\ every\ f \in \mathbb{V},\ t \geq 0.$$

We stated both the curvature conditions for the sake of completeness only, but we remark that $\mathsf{BE}_2(K,\infty)$ is sufficient for many of the results we are interested in this section. Obviously, $\mathsf{BE}_1(K,\infty)$ implies $\mathsf{BE}_2(K,\infty)$; the converse, first proved by Bakry [1985], has been recently extended to a nonsmooth setting in [Savaré, 2014, Corollary 3.5]) under the assumption that $\mathcal{E}$ is *quasi-regular*. The quasi-regularity property has many equivalent characterizations, a transparent one is for instance in terms of the existence of a sequence of compact sets $F_k \subset X$ such that

$$\bigcup_k \{f \in \mathbb{V} :\ f = 0\ \mathfrak{m}\text{-a.e. in } X \setminus F_k\}$$

is dense in $\mathbb{V}$.

The validity of the following inequality is actually equivalent to $\mathsf{BE}_2(K,\infty)$, see for instance [Ambrosio et al., 2012, Corollary 2.3] for a proof.

Proposition 13.2 (Reverse Poincaré inequalities)**.** *If* $\mathsf{BE}_2(K,\infty)$ *holds, then*

$$2\mathrm{I}_{2K}(t)\Gamma\left(\mathsf{P}_t f\right) \leq \mathsf{P}_t f^2 - \big(\mathsf{P}_t f\big)^2 \qquad \mathfrak{m}\text{-}a.e.\ in\ X, \tag{13.2}$$

for all $t > 0$, $f \in L^2(\mathfrak{m})$.

Corollary 13.3 (L^p-Γ inequalities)**.** *If* $\mathsf{BE}_2(K,\infty)$ *holds, then* L^p-Γ *inequalities hold for* $p \in [2,\infty]$.

Proof. The validity of L^p-Γ inequalities for $p \in [2,\infty]$ is obtained integrating (13.2),

$$(2\mathrm{I}_{2K}(t))^{p/2}\int \Gamma\left(\mathsf{P}_t f\right)^{p/2} d\mathfrak{m} \leq \int (\mathsf{P}_t f^2)^{p/2} d\mathfrak{m} \leq \int f^p d\mathfrak{m}$$

and using $2\mathrm{I}_{2K}(t)^{-1} = O(t^{-1})$ as $t \downarrow 0$. □

Another consequence of $\mathsf{BE}_2(K,\infty)$ is the following higher integrability of $\Gamma(f)$, recently proved in [Ambrosio et al., 2014c, Thm. 3.1] assuming higher integrability of f and Δf.

Theorem 13.4 (Gradient interpolation)**.** *Assume that* $\mathsf{BE}_2(K,\infty)$ *holds and let* $\lambda \geq K^-$, $f \in L^2 \cap L^\infty(\mathfrak{m})$*. If* $p \in \{2,\infty\}$ *and* $\Delta f \in L^p(\mathfrak{m})$*, then* $\Gamma(f) \in L^p(\mathfrak{m})$ *and*

$$\big\|\Gamma(f)\|_p \leq c\|f\|_\infty \|\Delta f + \lambda f\|_p$$

for a universal constant c *(i.e. independent of* λ*,* K*,* X*,* $\mathfrak{m}$*).*

Finally, we will need two more consequences of the $\mathsf{BE}_2(K,\infty)$ condition, proved under the quasi-regularity assumption in Savaré [2014]: the first one, first proved in [Savaré, 2014, Lemma 3.2] and then slightly improved in [Ambrosio et al., 2014c, Thm. 5.5], is the implication

$$f \in \mathbb{V}, \quad \Delta f \in L^4(\mathfrak{m}) \quad \Longrightarrow \quad \Gamma(f) \in \mathbb{V}. \tag{13.3}$$

In particular, this implication provides L^4 integrability of $\sqrt{\Gamma(f)}$, consistently with the integrability of the Laplacian. Moreover, it will be particularly useful the quantitative estimate, first proved in [Savaré, 2014, Thm. 3.4] and then slightly improved in [Ambrosio et al., 2014c, Corollary 5.7]:

$$\Gamma(\Gamma(f)) \leq 4\gamma_{2,K}[f]\,\Gamma(f) \quad \mathfrak{m}\text{-a.e. in } X, \text{ whenever } f \in \mathbb{V},\ \Delta f \in L^4(\mathfrak{m}). \tag{13.4}$$

The function $\gamma_{2,K}[f]$ in (13.4) is nonnegative, it satisfies the L^1 estimate

$$\int \gamma_{2,K}[f]d\mathfrak{m} \leq \int_X \big((\Delta f)^2 - K\Gamma(f)\big)d\mathfrak{m} \tag{13.5}$$

and it can be represented as the density w.r.t. $\mathfrak{m}$ of the nonnegative (and possibly singular w.r.t. $\mathfrak{m}$) measure defined by

$$\mathbb{V} \ni \varphi \mapsto \int_X -\frac{1}{2}\Gamma(\Gamma(f),\varphi) + \Delta f\,\Gamma(f,\varphi) + \big((\Delta f)^2 - K\Gamma(f)\big)\varphi d\mathfrak{m}.$$

The nonnegativity of this measure is one of the equivalent formulations of $\mathsf{BE}_2(K,\infty)$, see [Savaré, 2014, §3] for a more detailed discussion.

13.1.1 Choice of the algebra $\mathscr{A}$

We first prove that the following "minimal" choice for the algebra $\mathscr{A}$ provides (4.1) and optimal density conditions.

Proposition 13.5. *Under assumption* $\mathsf{BE}_2(K,\infty)$*, the algebra*

$$\mathscr{A}_1 := \left\{ f \in \bigcap_{1\leq p\leq\infty} L^p(\mathfrak{m}) : \ f \in \mathbb{V},\ \sqrt{\Gamma(f)} \in \bigcap_{1\leq p\leq\infty} L^p(\mathfrak{m}) \right\}$$

satisfies (4.1) *and it is dense in every space* $\mathbb{V}^p$*, for* $p \in [1,\infty)$*.*

Proof. Since (4.1) is obviously satisfied by the chain rule, we need only to show density of $\mathscr{A}_1$. First, we consider the algebra $\mathcal{A} = \mathbb{V} \cap \mathbb{V}^\infty$, which satisfies the weak "Feller" condition (4.2) because of (13.1). Moreover, for $p \in [2,\infty)$, the validity of the L^p-Γ inequality entail that $\mathcal{A}$ is dense in $L^2 \cap L^p$, and taking the $L^{p/2}$ norm in (13.1) gives that (3.13) holds. By Remark 4.4 we conclude that $\mathcal{A}$ is dense in $\mathbb{V}^p$, for every $p \in [2,\infty)$.

To establish density of $\mathscr{A}_1$ in $\mathbb{V}^p$ for $p \in [1,\infty)$, we argue as in Proposition 4.2. □

Retaining the density condition and the algebra property, one can also consider classes smaller than $\mathscr{A}_1$, including for instance bounds in $L^p(\mathfrak{m})$ for the Laplacian.

13.1.2 Conservation of mass

In this section we see prove that the curvature condition, together with the conservativity condition $\mathsf{P}_t^\infty 1 = 1$ for all $t > 0$ (recall that $\mathsf{P}_t^\infty : L^\infty(\mathfrak{m}) \to L^\infty(\mathfrak{m})$ is the dual semigroup in (3.1.2)), imply the existence of a sequence $(f_n) \subset \mathscr{A}_1$ as in (9.2), at least if $\mathcal{L}$ is of the form $a\Delta + \boldsymbol{b}$, where a, $|\boldsymbol{b}| \in L^q(\mathfrak{m})$, for $q \in (1,\infty]$. Notice that the conservativity is loosely related to a mass conservation property, for the continuity equation with derivation induced by the logarithmic derivative of the density; therefore, even though sufficient conditions adapted to the prescribed derivation $\boldsymbol{b}$ could be considered as well, it is natural to consider the conservativity of P in connection with (9.2).

Proposition 13.6. *If $\mathsf{BE}_2(K,\infty)$ holds and P is conservative, then there exist $(f_n) \subset \mathscr{A}_1$ satisfying* (9.2).

Proof. Let $(g_n) \subset L^1 \cap L^\infty(\mathfrak{m})$ be a non-decreasing sequence of functions (whose existence is ensured by the σ-finiteness assumption on $\mathfrak{m}$) with

$$0 \le g_n \le 1 \text{ for every } n \ge 1 \text{ and } \lim_{n\to\infty} g_n = 1, \ \mathfrak{m}\text{-a.e. in } X.$$

These conditions imply in particular that $g_n \to 1$ weakly* in $L^\infty(\mathfrak{m})$.

Let $h_n = \int_0^1 \mathsf{P}_s g_n ds = \int_0^1 \mathsf{P}_s^\infty g_n ds$ and define $f_n := \mathsf{P}_1 h_n = \mathsf{P}_1^\infty h_n$. By linearity and continuity of P^∞ we obtain that $f_n \to \mathsf{P}_1^\infty 1 = 1$ weakly* in $L^\infty(\mathfrak{m})$. In addition, expanding the squares, it is easily seen that

$$\lim_{n\to\infty} \int (1-f_n)^2 v d\mathfrak{m} = 0 \qquad \forall v \in L^1(\mathfrak{m}).$$

Hence, by a diagonal argument we can assume (possibly extracting a subsequence) that $f_n \to 1$ $\mathfrak{m}$-a.e. in X.

Since $h_n \le 1$, the reverse Poincaré inequality (13.2) entails

$$\Gamma(f_n) \le \frac{\mathsf{P}_1 h_n^2 - (f_n)^2}{2\mathrm{I}_{2K}(1)} \le \frac{1-(f_n)^2}{2\mathrm{I}_{2K}(1)}, \quad \mathfrak{m}\text{-a.e. in } X.$$

Taking the square roots of both sides and using the a.e. convergence of f_n we obtain, thanks to dominated convergence, that $\sqrt{\Gamma(f_n)}$ weakly* converge to 0 in $L^\infty(\mathfrak{m})$.

Finally, we discuss the regularity of f_n. Since

$$\Delta f_n = \int_1^2 \Delta \mathsf{P}_s g_n ds = \mathsf{P}_2 g_n - \mathsf{P}_1 g_n \in L^\infty(\mathfrak{m})$$

we can use Theorem 13.4 to obtain $\sqrt{\Gamma(f_n)} \in L^\infty(\mathfrak{m})$. In order to obtain integrability of the gradient for powers between 1 and 2 we can replace f_n by $k_n := \Phi_1(f_n)/\Phi_1(1)$, with $\Phi_1 : \mathbb{R} \to \mathbb{R}$ as introduced in Proposition 4.1. □

13.1.3 Derivations associated to gradients and their deformation

In this section, we study more in detail the class of "gradient" derivations $\boldsymbol{b}_V$ in (4.5). More generally, we analyze the regularity of the derivation $f \mapsto \omega\Gamma(f,V)$ associated to sufficiently regular V and ω in $\mathbb{V}$.

For $p \in (1,\infty]$, let us denote

$$D_{L^p}(\Delta) := \big\{ f \in \mathbb{V} \cap L^p(\mathfrak{m}) :\ \Delta f \in L^p(\mathfrak{m}) \big\}.$$

Thanks to the implication (13.3), $D_{L^4}(\Delta) \subset \mathbb{V}_4$ and the Hessian

$$(f,g) \mapsto H[V](f,g) := \frac{1}{2}\left[\Gamma(f,\Gamma(V,g)) + \Gamma(g,\Gamma(V,f)) - \Gamma(V,\Gamma(f,g))\right] \in L^1(\mathfrak{m}),$$

is well defined on $D_{L^4}(\Delta) \times D_{L^4}(\Delta)$. Notice that the expression is symmetric in (f,g), that $(V,f,g) \mapsto H[V](f,g)$ is multilinear, and that

$$H[V](f, g_1 g_2) = H[V](f,g_1)g_2 + g_1 H[V](f,g_2).$$

By [Savaré, 2014, Thm. 3.4], we have the estimate

$$|H[V](f,g)| \le \sqrt{\gamma_{2,K}[V]}\sqrt{\Gamma(f)}\sqrt{\Gamma(g)}, \quad \mathfrak{m}\text{-a.e. in } X, \tag{13.6}$$

for every $f,\, g \in D_{L^4}(\Delta)$.

Theorem 13.7. *If $\mathsf{BE}_2(K,\infty)$ holds and $\mathcal{E}$ is quasi-regular, then for all $V \in D(\Delta)$, $\omega \in \mathbb{V} \cap L^\infty(\mathfrak{m})$ with $\sqrt{\Gamma(\omega)} \in L^\infty(\mathfrak{m})$ and $c \in \mathbb{R}$, the derivation $\boldsymbol{b} = (\omega + c)\boldsymbol{b}_V$ has deformation of type $(4,4)$ according to Definition 10.5 with $q = 2$, and it satisfies*

$$\|D^{sym}\boldsymbol{b}\|_{4,4} \le \|\omega + c\|_\infty \left\|(\Delta V)^2 - K\Gamma(V)\right\|_1 + \left\|\sqrt{\Gamma(\omega)}\right\|_\infty \left\|\sqrt{\Gamma(V)}\right\|_2. \tag{13.7}$$

Proof. Assume first that $V \in D_{L^4}(\Delta)$. Let $f,\, g \in D_{L^4}(\Delta)$. After integrating by parts the Laplacians of f and g, the very definition of $D^{sym}\boldsymbol{b}$ gives

$$\int D^{sym}\boldsymbol{b}(f,g)d\mathfrak{m} = \int (\omega + c)H[V](f,g) + \frac{1}{2}\left[\Gamma(\omega,f)\Gamma(V,g) + \Gamma(\omega,g)\Gamma(V,f)\right]d\mathfrak{m}.$$

By Hölder inequality, we can use (13.6) to estimate $\left|\int D^{sym}\boldsymbol{b}(f,g)d\mathfrak{m}\right|$ from above with

$$\left[\|\omega\|_\infty \left\|\sqrt{\gamma_{2,K}[V]}\right\|_2 + \left\|\sqrt{\Gamma(\omega)}\right\|_\infty \left\|\sqrt{\Gamma(V)}\right\|_2\right] \left\|\sqrt{\Gamma(f)}\right\|_4 \left\|\sqrt{\Gamma(g)}\right\|_4.$$

Thus, by definition of $\|D^{sym}\boldsymbol{b}\|_{4,4}$, (13.7) follows, taking also (13.5) into account. To pass to the general case $V \in D(\Delta)$, it is sufficient to approximate V with $V_n \in D_{L^4}(\Delta)$ in such a way that $V_n \to V$ in $\mathbb{V}$ and $\Delta V_n \to \Delta V$ in $L^2(\mathfrak{m})$ and notice that $\int D^{sym}\boldsymbol{b}_n(f,g)d\mathfrak{m}$ converge to $\int D^{sym}\boldsymbol{b}(f,g)d\mathfrak{m}$ directly from (10.12). The existence of such an approximating sequence is obtained arguing as in [Ambrosio et al., 2014c, Lemma 4.2], i.e. given $f \in D(\Delta)$, we let $h = f - \Delta f \in L^2(\mathfrak{m})$,

$$h_n := \max\{\min\{h,n\}, -n\} \in L^2 \cap L^\infty(\mathfrak{m})$$

and define f_n as the unique (weak) solution to $f_n - \Delta f_n = h_n$. The maximum principle for Δ (or equivalently the fact that the resolvent operator $R_1 = (I - \Delta)^{-1}$ is Markov) gives $f_n \in L^2 \cap L^\infty(\mathfrak{m})$, thus $\Delta f_n \in L^2 \cap L^\infty(\mathfrak{m})$ and by L^2-continuity of R_1, as $n \to \infty$, both h_n and f_n converge, respectively towards h and f. By difference, also Δf_n converge towards Δf in $L^2(\mathfrak{m})$ and this gives also easily convergence of f_n to f in $\mathbb{V}$. □

13.2 $\mathsf{RCD}(K,\infty)$ metric measure spaces

The class of $\mathsf{CD}(K,\infty)$ consists of complete metric measure spaces such that the Shannon relative entropy w.r.t. $\mathfrak{m}$ is K-convex along Wasserstein geodesics, see Villani [2009] for a full account of the theory and its geometric and functional implications. The class of $\mathsf{RCD}(K,\infty)$ metric measure spaces was first introduced in Ambrosio et al. [2014b], from a metric perspective, as class of spaces smaller than that of $\mathsf{CD}(K,\infty)$ metric measure spaces. The additional requirement, in this class of spaces, is that the so-called Cheeger energy is quadratic; with this axiom, Finsler geometries are ruled out and stronger structural (and stability) properties can be established.

We will use the notation $W^{1,2}(X,d,\mathfrak{m})$ for the Sobolev space, Ch for the Cheeger energy arising from the relaxation in $L^2(X,\mathfrak{m})$ of the local Lipschitz constant

$$|\mathrm{D}f|(x) := \limsup_{y\to x}\frac{|f(y)-f(x)|}{d(y,x)}$$

of $L^2(\mathfrak{m})$ and Lipschitz maps $f : X \to \mathbb{R}$.

To introduce $\mathsf{RCD}(K,\infty)$ spaces we restrict the discussion to metric measure spaces $(X,d,\mathfrak{m})$ satisfying the following three conditions:

(a) (X,d) is a complete and separable length space;

(b) $\mathfrak{m}$ is a nonnegative Borel measure with $\mathrm{supp}(\mathfrak{m}) = X$, satisfying

$$\mathfrak{m}(B_r(x)) \le c\,\mathrm{e}^{Ar^2} \tag{13.8}$$

for suitable constants $c \ge 0$, $A \ge 0$;

(c) $(X,\mathsf{d},\mathfrak{m})$ is infinitesimally Hilbertian according to the terminology introduced in Gigli [2012], i.e., the Cheeger energy Ch is a quadratic form.

As explained in Ambrosio et al. [2014b], Ambrosio et al. [2012], the quadratic form Ch canonically induces a strongly regular Dirichlet $\mathcal{E}$ form in (X,τ) (where τ is the topology induced by the distance d), as well as a carré du champ $\Gamma : D(\mathcal{E})\times D(\mathcal{E}) \to L^1(\mathfrak{m})$. Thus, we recover the basic setting of (3.1) and we can identify $W^{1,2}(X,d,\mathfrak{m})$ with $\mathbb{V}$. In addition, P is conservative because of (13.8) and the definition of Ch provides the approximation property

$$\exists\ f_n \in \mathrm{Lip}(X)\cap L^2(\mathfrak{m}) \text{ with } f_n \to f \text{ in } L^2(\mathfrak{m}) \text{ and } |\mathrm{D}f_n| \to \sqrt{\Gamma(f)} \text{ in } L^2(\mathfrak{m}) \tag{13.9}$$

for all $f \in \mathbb{V}$.

The above discussions justify the following definition of $\mathsf{RCD}(K,\infty)$. It is not the original one given in Ambrosio et al. [2014b], but it is more appropriate for our purposes; the equivalence of the two definitions is given in Ambrosio et al. [2012].

Definition 13.8 ($\mathsf{RCD}(K,\infty)$ metric measure spaces)**.** *We say that* $(X,d,\mathfrak{m})$*, satisfying* (a), (b), (c) *above, is an* $\mathsf{RCD}(K,\infty)$ *space if:*

(a) the Dirichlet form associated to the Cheeger energy of $(X,d,\mathfrak{m})$ *satisfies* $\mathsf{BE}_2(K,\infty)$ *according to Definition 13.1;*

(b) any $f \in W^{1,2}(X,d,\mathfrak{m})\cap L^\infty(\mathfrak{m})$ *with* $\left\|\Gamma(f)\right\|_\infty \le 1$ *has a* 1*-Lipschitz representative.*

From [Ambrosio et al., 2014b, Lemma 6.7] we obtain that $\mathcal{E}$ is *quasi-regular*. We set throughout $\mathscr{A}$ be the class of Lipschitz functions with bounded support. It is easily seen that $\mathscr{A}$ is dense in $\mathbb{V}$.

Since both (X,d) and $\mathbb{V}$ are separable, it is not difficult to exhibit a countable family $\mathscr{A}^* \subset \mathscr{A}$ such that (7.4) is satisfied: let $(x_h) \subset X$ be dense, and set $f_{h,k} := (d(x_h,\cdot)-k)^- \in \mathscr{A}$ for h, $k \in \mathbb{N}$; then, define

$$\mathscr{B} := \bigcup_{h,\,k=0}^{\infty} \{f_{h,k}\} \cup \bigcup_{h=0}^{\infty} \{g_h\},$$

with $(g_h) \subset \mathscr{A}$ dense in $\mathbb{V}$. Then, defining $\mathscr{A}^* = \{f \in \mathscr{B} : \ \|\Gamma(f)\|_\infty \le 1\} \subset \mathscr{A}$, since $\mathbb{R}\mathscr{A}^* = \mathscr{B}$ we obtain (7.4). Regarding (7.6), it is not clear in general whether one has to enlarge the topology in order to ensure its validity. Nevertheless, the distance d_∞ defined by (7.5) coincides with d: $d_{\mathscr{A}^*} \le d$ is obvious, while $d \le d_{\mathscr{A}^*}$ follows from taking $f = f_{h,k}$ in (7.5), with x_h arbitrarily close to x and k larger than $d(x,y)$: thus by Lemma 7.6, we still obtain, at least in the deterministic case, probability measures on continuous curves with respect to the original distance d.

We discuss now the fine regularity properties of functions in $\mathbb{V}$, recalling some results developed in Ambrosio et al. [2014a]. We start with the notion of 2-plan.

Definition 13.9 (2-plans)**.** *We say that a positive finite measure $\boldsymbol{\eta}$ in $\mathscr{P}(C([0,T];X))$ is a* 2*-plan if $\boldsymbol{\eta}$ is concentrated on $AC([0,T];(X,d))$ and the following two properties hold:*

(a) $\int \int_0^T |\dot\eta|^2(t)dtd\boldsymbol{\eta}(\eta) < \infty$*;*

(b) there exists $C \in [0,\infty)$ satisfying $(e_t)_\# \boldsymbol{\eta} \le C\mathfrak{m}$ for all $t \in [0,T]$.

Accordingly, we say that $V : X \to \mathbb{R}$ is $W^{1,2}$ *along* 2*-almost every curve* if, for all $s \le t$ and all 2-plans $\boldsymbol{\eta}$, the family of inequalities

$$\int |V(\eta(s)) - V(\eta(t))| d\boldsymbol{\eta}(\eta) \le \int\int_s^t g(\eta(r))|\dot\eta(r)|drd\boldsymbol{\eta}(\eta), \quad \text{for all } s,\ t \in [0,T) \text{ with } s \le t \tag{13.10}$$

holds for some $g \in L^2(\mathfrak{m})$. Since Lipschitz functions with bounded support are dense in $\mathbb{V}$, a density argument [Ambrosio et al., 2014a, Theorem 5.14] based on (13.9) provides the following result:

Proposition 13.10. *Any $V \in \mathbb{V}$ is $W^{1,2}$ along* 2*-almost every curve. In addition,* (13.10) *holds with $g = \sqrt{\Gamma(V)}$.*

Actually, a much finer result could be established (see [Ambrosio et al., 2014a, §5]), namely the existence of a representative $\tilde V$ of V in the $L^2(\mathfrak{m})$ equivalence class, with the property that $\tilde V \circ \eta$ is absolutely continuous $\boldsymbol{\pi}$-a.e. η for any 2-plan $\boldsymbol{\pi}$, with $|(\tilde V \circ \eta)'| \le \sqrt{\Gamma(V)}|\dot\eta|$ a.e. in $(0,T)$. However, we shall not need this fact in the sequel. Here we notice only that since $\chi_B\boldsymbol{\eta}$ is a 2-plan for any Borel set $B \subset C([0,T];X)$, it follows from (13.10) with $g = \sqrt{\Gamma(V)}$ that

$$|V(\eta(s)) - V(\eta(t))| \le \int\int_s^t \sqrt{\Gamma(V)}(\eta(r))|\dot\eta(r)|dr, \qquad \text{for } \boldsymbol{\eta}\text{-a.e. } \eta \tag{13.11}$$

for all s, $t \in [0,T)$ with $s \le t$.

Now, we would like to relate these known facts to solutions to the ODE $\dot\eta = \boldsymbol{b}_t(\eta)$. The first connection between 2-plans and probability measures concentrated on solutions to the ODE is provided by the following proposition.

Proposition 13.11. *Let $\boldsymbol{b}=(\boldsymbol{b}_t)$ be a Borel family of derivations with $|\boldsymbol{b}|\in L^1_t(L^2)$ and let $u\in L^\infty_t(L^\infty_x)$. Let $\boldsymbol{\eta}$ be concentrated on solutions to the ODE $\dot\eta=\boldsymbol{b}_t(\eta)$, with $(e_t)_\#\boldsymbol{\eta}=u_t\mathfrak{m}$ for all $t\in(0,T)$. Then $\boldsymbol{\eta}$ is a 2-plan.*

Proof. The fact that $\boldsymbol{\eta}$ has bounded marginals follows from the assumption $u\in L^\infty_t(L^\infty_x)$. By identification $d=d_{\mathscr{A}^*}$, $\boldsymbol{\eta}$ is concentrated on $AC([0,T];(X,d))$, with $|\dot\eta|(t)\le|\boldsymbol{b}_t|(\eta(t))$, $\mathscr{L}^1$-a.e. in $(0,T)$ for $\boldsymbol{\eta}$-a.e. η. Thus,

$$\int\int_0^T|\dot\eta|^2(t)dtd\boldsymbol{\eta}(\eta)\le\int_0^T\int|\boldsymbol{b}_t|^2u_td\mathfrak{m}dt<\infty.$$

□

We now focus on the case of a "gradient" and time-independent derivation $\boldsymbol{b}_V$ associated to $V\in\mathbb{V}$. Recall that in this case $|\boldsymbol{b}_V|^2=\Gamma(V)$ $\mathfrak{m}$-a.e. in X.

Theorem 13.12. *Let $V\in D(\Delta)$ with $\Delta V^-\in L^\infty(\mathfrak{m})$. Then, there exist weakly continuous solutions (in $[0,T)$, in duality with $\mathscr{A}$) $u\in L^\infty_t(L^1_x\cap L^\infty_x)$ to the continuity equation, for any initial condition $\bar u\in L^1\cap L^\infty(\mathfrak{m})$. In addition, if $\boldsymbol{\eta}$ is given by Theorem 7.3 (namely $\boldsymbol{\eta}$ is concentrated on solutions to the ODE $\dot\eta=\boldsymbol{b}_V(\eta)$ and $(e_t)_\#\boldsymbol{\eta}=u_t$ for all $t\in(0,T)$), then:*

(a) $\boldsymbol{\eta}$ is concentrated on curves η satisfying $|\dot\eta|(t)=\Gamma(V)^{1/2}(\eta(t))$ for a.e. $t\in(0,T)$;

(b) for all s, $t\in[0,T)$ with $s\le t$, there holds

$$V\circ\eta(t)-V\circ\eta(s)=\int_s^t\Gamma(V)(\eta(r))dr,\quad \textit{for } \boldsymbol{\eta}\textit{-a.e. } \eta.$$

Proof. The proof of the first statement follows immediately by Proposition 9.5, approximating V via truncations and the smoothing action of P. Since

$$\int_s^t\int\Gamma(V,f)u_rd\mathfrak{m}dr=\int fu_t-\int fu_s\quad\text{for all } s,\ t\in[0,T) \text{ with } s\le t$$

for all $f\in\mathscr{A}$, we can use the density of $\mathscr{A}$ in $\mathbb{V}$ and a simple limiting procedure to obtain

$$\int_s^t\int\Gamma(V)u_rd\mathfrak{m}dr=\int Vu_t-\int Vu_s\quad\text{for all } s,\ t\in[0,T) \text{ with } s\le t. \tag{13.12}$$

If $\boldsymbol{\eta}$ is as in the statement of the theorem, since $\boldsymbol{\eta}$ is a 2-plan we can combine Proposition 13.10 and the inequality $|\dot\eta|\le|\boldsymbol{b}_V|(\eta)$ stated in Lemma 7.6 to get

$$\int V(\eta(t))-V(\eta(s))d\boldsymbol{\eta}(\eta)\le\int\int_s^t\Gamma(V)^{1/2}(\eta(r))\,|\dot\eta|(r)drd\boldsymbol{\eta}(\eta)\le\int\int_s^t\Gamma(V)(\eta(r))d\boldsymbol{\eta}(\eta),$$

for all s, $t\in[0,T)$ with $s\le t$. Since $(e_r)_\#\boldsymbol{\eta}=u_r\mathfrak{m}$ for all $r\in[0,T)$, it follows that

$$\int Vu_t-\int Vu_s=\int V(\eta(t))-V(\eta(s))d\boldsymbol{\eta}(\eta)\le\int_s^t\Gamma(V)u_rd\mathfrak{m}dr. \tag{13.13}$$

Combining (13.12) and (13.13) it follows that all the intermediate inequalities we integrated w.r.t. $\boldsymbol{\eta}$ are actually identities, so that for $\boldsymbol{\eta}$-a.e. η it must be $|\dot\eta|=\sqrt{\Gamma(V)}\circ\eta$ a.e. in $(0,T)$ and equality holds in (13.11). □

In particular, one could prove that $\boldsymbol{\eta}$ is a 2-plan representing the 2-weak gradient of V, according to [Gigli, 2012, Def. 3.7], where a weaker asymptotic energy dissipation inequality was required at $t = 0$. Our global energy dissipation is stronger, but it requires additional bounds on the Laplacian.

We can also prove uniqueness for the continuity equation, considering just for simplicity still the autonomous version.

Theorem 13.13. *Let $V \in D(\Delta)$ with $\Delta V^- \in L^\infty(\mathfrak{m})$. Then the continuity equation induced by $\boldsymbol{b}_V$ has existence and uniqueness in $L^\infty_t(L^1_x \cap L^\infty_x)$ for any initial condition $\bar{u} \in L^1 \cap L^\infty(\mathfrak{m})$.*

Proof. We already discussed existence in Theorem 13.12. For uniqueness, we want to apply Theorem 10.19 with $q = 2$ and $r = s = 4$ (which provides uniqueness in the larger class $L^2 \cap L^4(\mathfrak{m})$). In order to do this we need only to know that (9.2) holds (this follows by conservativity of P and $\mathsf{BE}_2(K,\infty)$), that L^4-Γ inequalities hold in $\mathsf{RCD}(K,\infty)$ spaces (this follows by $\mathsf{BE}_2(K,\infty)$ thanks to Corollary 13.3) and that the deformation of $\boldsymbol{b}_V$ is of type $(4,4)$ (this follows by Theorem 13.7). $\square$

Bibliography

L. Ambrosio and D. Trevisan. Well posedness of Lagrangian flows and continuity equations in metric measure spaces. *Analysis and PDE*, 7(5):1179–1234, 2014. doi: 10.2140/apde.2014.7.1179.

L. Ambrosio, N. Gigli, and G. Savaré. Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds. *ArXiv e-prints*, September 2012.

Luigi Ambrosio. Transport equation and Cauchy problem for *BV* vector fields. *Invent. Math.*, 158(2):227–260, 2004. ISSN 0020-9910. doi: 10.1007/s00222-004-0367-2.

Luigi Ambrosio. Transport equation and Cauchy problem for non-smooth vector fields. In *Calculus of variations and nonlinear partial differential equations*, volume 1927 of *Lecture Notes in Math.*, pages 1–41. Springer, Berlin, 2008. doi: 10.1007/978-3-540-75914-0_1.

Luigi Ambrosio and Gianluca Crippa. Existence, uniqueness, stability and differentiability properties of the flow associated to weakly differentiable vector fields. In *Transport equations and multi-D hyperbolic conservation laws*, volume 5 of *Lect. Notes Unione Mat. Ital.*, pages 3–57. Springer, Berlin, 2008. doi: 10.1007/978-3-540-76781-7_1.

Luigi Ambrosio and Alessio Figalli. On flows associated to Sobolev vector fields in Wiener spaces: an approach à la DiPerna-Lions. *J. Funct. Anal.*, 256(1):179–214, 2009. ISSN 0022-1236. doi: 10.1016/j.jfa.2008.05.007.

Luigi Ambrosio and Bernd Kirchheim. Currents in metric spaces. *Acta Math.*, 185(1):1–80, 2000. ISSN 0001-5962. doi: 10.1007/BF02392711.

Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000. ISBN 0-19-850245-1.

Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008. ISBN 978-3-7643-8721-1.

Luigi Ambrosio, Michele Miranda, Jr., Stefania Maniglia, and Diego Pallara. BV functions in abstract Wiener spaces. *J. Funct. Anal.*, 258(3):785–813, 2010. ISSN 0022-1236. doi: 10.1016/j.jfa.2009.09.008.

Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *Invent. Math.*, 195(2): 289–391, 2014a. ISSN 0020-9910. doi: 10.1007/s00222-013-0456-1.

Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *Duke Math. J.*, 163(7):1405–1490, 2014b. ISSN 0012-7094. doi: 10.1215/00127094-2681605.

Luigi Ambrosio, Andrea Mondino, and Giuseppe Savaré. On the Bakry–émery Condition, the Gradient Estimates and the Local-to-Global Property of $\mathsf{RCD}^*(K,N)$ Metric Measure Spaces. *The Journal of Geometric Analysis*, pages 1–33, 2014c. ISSN 1050-6926. doi: 10.1007/s12220-014-9537-7.

D. Bakry. Transformations de Riesz pour les semi-groupes symétriques. II. Étude sous la condition $\Gamma_2 \geq 0$. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 145–174. Springer, Berlin, 1985. doi: 10.1007/BFb0075844.

Dominique Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory (Saint-Flour, 1992)*, volume 1581 of *Lecture Notes in Math.*, pages 1–114. Springer, Berlin, 1994. doi: 10.1007/BFb0073872.

Dominique Bakry and Michel Émery. Hypercontractivité de semi-groupes de diffusion. *C. R. Acad. Sci. Paris Sér. I Math.*, 299(15):775–778, 1984. ISSN 0249-6291.

Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014. ISBN 978-3-319-00226-2; 978-3-319-00227-9. doi: 10.1007/978-3-319-00227-9.

Viorel Barbu, Michael Röckner, and Francesco Russo. Probabilistic representation for solutions of an irregular porous media type equation: the degenerate case. *Probab. Theory Related Fields*, 151(1-2):1–43, 2011. ISSN 0178-8051. doi: 10.1007/s00440-010-0291-x.

Fabrice Baudoin and Bumsik Kim. Sobolev, Poincaré, and isoperimetric inequalities for subelliptic diffusion operators satisfying a generalized curvature dimension inequality. *Rev. Mat. Iberoam.*, 30(1):109–131, 2014. ISSN 0213-2230. doi: 10.4171/RMI/771.

Fabrice Baudoin, Michel Bonnefont, and Nicola Garofalo. A sub-Riemannian curvature-dimension inequality, volume doubling property and the Poincaré inequality. *Math. Ann.*, 358(3-4):833–860, 2014. ISSN 0025-5831. doi: 10.1007/s00208-013-0961-y.

Nadia Belaribi and Francesco Russo. Uniqueness for Fokker-Planck equations with measurable coefficients and applications to the fast diffusion equation. *Electron. J. Probab.*, 17:no. 84, 28, 2012. ISSN 1083-6489. doi: 10.1214/EJP.v17-2349.

Vladimir Bogachev, Giuseppe Da Prato, and Michael Rökner. On weak parabolic equations for probability measures. *Dokl. Akad. Nauk*, 386(3):295–299, 2002. ISSN 0869-5652.

Vladimir Bogachev, Giuseppe Da Prato, and Michael Röckner. Uniqueness for solutions of Fokker-Planck equations on infinite dimensional spaces. *Comm. Partial Differential Equations*, 36(6):925–939, 2011. ISSN 0360-5302.

Vladimir I. Bogachev. *Gaussian measures*, volume 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1998. ISBN 0-8218-1054-5. doi: 10.1090/surv/062.

Nicolas Bouleau and Francis Hirsch. *Dirichlet forms and analysis on Wiener space*, volume 14 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1991. ISBN 3-11-012919-1. doi: 10.1515/9783110858389.

I. Capuzzo Dolcetta and B. Perthame. On some analogy between different approaches to first order PDE's with nonsmooth coefficients. *Adv. Math. Sci. Appl.*, 6(2):689–703, 1996. ISSN 1343-4373.

Zhen-Qing Chen and Masatoshi Fukushima. *Symmetric Markov processes, time change, and boundary theory*, volume 35 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2012. ISBN 978-0-691-13605-9.

Gianluca Crippa and Camillo De Lellis. Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.*, 616:15–46, 2008. ISSN 0075-4102. doi: 10.1515/CRELLE.2008.016.

G. Da Prato, F. Flandoli, E. Priola, and M. Röckner. Strong uniqueness for stochastic evolution equations in Hilbert spaces perturbed by a bounded measurable drift. *Ann. Probab.*, 41(5):3306–3344, 2013. ISSN 0091-1798. doi: 10.1214/12-AOP763.

Giuseppe Da Prato. *Kolmogorov equations for stochastic PDEs*. Advanced Courses in Mathematics. CRM Barcelona. Birkhäuser Verlag, Basel, 2004. ISBN 3-7643-7216-8. doi: 10.1007/978-3-0348-7909-5.

Giuseppe Da Prato. *Introduction to stochastic analysis and Malliavin calculus*, volume 13 of *Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)]*. Edizioni della Normale, Pisa, third edition, 2014. ISBN 978-88-7642-497-7; 88-7642-497-7; 978-88-7642-499-1. doi: 10.1007/978-88-7642-499-1.

Giuseppe Da Prato, Franco Flandoli, and Michael Röckner. Uniqueness for continuity equations in Hilbert spaces with weakly differentiable drift. *Stoch. Partial Differ. Equ. Anal. Comput.*, 2(2):121–145, 2014. ISSN 2194-0401. doi: 10.1007/s40072-014-0031-9.

GianMaria Dall'Ara and Dario Trevisan. Uncertainty Inequalities on Groups and Homogeneous Spaces via Isoperimetric Inequalities. *The Journal of Geometric Analysis*, pages 1–22, 2014. ISSN 1050-6926. doi: 10.1007/s12220-014-9512-3.

Eleonora Di Nezza, Giampiero Palatucci, and Enrico Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573, 2012. ISSN 0007-4497. doi: 10.1016/j.bulsci.2011.12.004.

R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511–547, 1989. ISSN 0020-9910. doi: 10.1007/BF01393835.

Andreas Eberle. *Uniqueness and non-uniqueness of semigroups generated by singular diffusion operators*, volume 1718 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1999. ISBN 3-540-66628-1.

Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. ISBN 0-471-08186-8. doi: 10.1002/9780470316658. Characterization and convergence.

Shizan Fang, Dejun Luo, and Anton Thalmaier. Stochastic differential equations with coefficients in Sobolev spaces. *J. Funct. Anal.*, 259(5):1129–1168, 2010. ISSN 0022-1236. doi: 10.1016/j.jfa.2010.02.014.

Alessio Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.*, 254(1):109–153, 2008. ISSN 0022-1236. doi: 10.1016/j.jfa.2007.09.020.

Masatoshi Fukushima, Yoichi Oshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, extended edition, 2011. ISBN 978-3-11-021808-4.

N. Gigli. On the differential structure of metric measure spaces and applications. *ArXiv e-prints*, May 2012.

N. Gigli. Nonsmooth differential geometry - An approach tailored for spaces with Ricci curvature bounded from below. *ArXiv e-prints*, July 2014.

David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. ISBN 3-540-41160-7. Reprint of the 1998 edition.

Alexander Grigor′yan. Analytic and geometric background of recurrence and non-explosion of the Brownian motion on Riemannian manifolds. *Bull. Amer. Math. Soc. (N.S.)*, 36(2): 135–249, 1999. ISSN 0273-0979. doi: 10.1090/S0273-0979-99-00776-4.

Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998. ISSN 0036-1410. doi: 10.1137/S0036141096303359.

Tosio Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. ISBN 3-540-58661-X. Reprint of the 1980 edition.

Alexander V. Kolesnikov and Michael Röckner. On continuity equations in infinite dimensions with non-Gaussian reference measure. *J. Funct. Anal.*, 266(7):4490–4537, 2014. ISSN 0022-1236. doi: 10.1016/j.jfa.2014.01.010.

N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, 131(2):154–196, 2005. ISSN 0178-8051. doi: 10.1007/s00440-004-0361-z.

N.V. Krylov. On Kolmogorov's equations for finite dimensional diffusions. In Giueppe Da Prato, editor, *Stochastic PDE's and Kolmogorov Equations in Infinite Dimensions*, volume 1715 of *Lecture Notes in Mathematics*, pages 1–63. Springer Berlin Heidelberg, 1999. ISBN 978-3-540-66545-8. doi: 10.1007/BFb0092417.

Thomas G. Kurtz. The Yamada-Watanabe-Engelbert theorem for general stochastic equations and inequalities. *Electron. J. Probab.*, 12:951–965, 2007. ISSN 1083-6489. doi: 10.1214/EJP.v12-431.

Thomas G. Kurtz and Richard H. Stockbridge. Existence of Markov controls and characterization of optimal Markov controls. *SIAM J. Control Optim.*, 36(2):609–653 (electronic), 1998. ISSN 0363-0129. doi: 10.1137/S0363012995295516.

C. Le Bris and P.-L. Lions. Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients. *Comm. Partial Differential Equations*, 33(7-9):1272–1317, 2008. ISSN 0360-5302. doi: 10.1080/03605300801970952.

Stefano Lisini. Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces. *ESAIM Control Optim. Calc. Var.*, 15(3):712–740, 2009. ISSN 1292-8119.

John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009. ISSN 0003-486X. doi: 10.4007/annals.2009.169.903.

De Jun Luo. Fokker-Planck type equations with Sobolev diffusion coefficients and BV drift coefficients. *Acta Math. Sin. (Engl. Ser.)*, 29(2):303–314, 2013. ISSN 1439-8516. doi: 10.1007/s10114-012-1302-x.

Zhi Ming Ma and Michael Röckner. *Introduction to the theory of (nonsymmetric) Dirichlet forms*. Universitext. Springer-Verlag, Berlin, 1992. ISBN 3-540-55848-9. doi: 10.1007/978-3-642-77739-4.

E. Mayer-Wolf and M. Zakai. The divergence of Banach space valued random variables on Wiener space. *Probab. Theory Related Fields*, 132(2):291–320, 2005. ISSN 0178-8051. doi: 10.1007/s00440-004-0397-0.

Luca Natile, Mark A. Peletier, and Giuseppe Savaré. Contraction of general transportation costs along solutions to Fokker-Planck equations with monotone drifts. *J. Math. Pures Appl. (9)*, 95(1):18–35, 2011. ISSN 0021-7824.

David Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006. ISBN 978-3-540-28328-7; 3-540-28328-5.

Mark A. Peletier, D. R. Michiel Renger, and Marco Veneroni. Variational formulation of the Fokker-Planck equation with decay: a particle approach. *Commun. Contemp. Math.*, 15 (5):1350017, 43, 2013. ISSN 0219-1997.

M. Pratelli and D. Trevisan. Functions of bounded variation on the classical Wiener space and an extended Ocone-Karatzas formula. *Stochastic Process. Appl.*, 122(6):2383–2399, 2012. ISSN 0304-4149. doi: 10.1016/j.spa.2012.03.010.

Michael Röckner and Xicheng Zhang. Weak uniqueness of Fokker-Planck equations with degenerate and bounded coefficients. *C. R. Math. Acad. Sci. Paris*, 348(7-8):435–438, 2010. ISSN 1631-073X. doi: 10.1016/j.crma.2010.01.001.

Giuseppe Savaré. Self-improvement of the Bakry-Émery condition and Wasserstein contraction of the heat flow in $\mathrm{RCD}(K,\infty)$ metric measure spaces. *Discrete Contin. Dyn. Syst.*, 34(4):1641–1661, 2014. ISSN 1078-0947. doi: 10.3934/dcds.2014.34.1641.

Ichiro Shigekawa. Schrödinger operators on the Wiener space. *Commun. Stoch. Anal.*, 1(1): 1–17, 2007. ISSN 0973-9599.

R. E. Showalter. *Monotone operators in Banach space and nonlinear partial differential equations*, volume 49 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997. ISBN 0-8218-0500-2.

S. K. Smirnov. Decomposition of solenoidal vector charges into elementary solenoids, and the structure of normal one-dimensional flows. *Algebra i Analiz*, 5(4):206–238, 1993. ISSN 0234-0852.

Wilhelm Stannat. The theory of generalized Dirichlet forms and its applications in analysis and stochastics. *Mem. Amer. Math. Soc.*, 142(678):viii+101, 1999. ISSN 0065-9266. doi: 10.1090/memo/0678.

Elias M. Stein. *Topics in harmonic analysis related to the Littlewood-Paley theory.* Annals of Mathematics Studies, No. 63. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1970.

Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. ISBN 978-3-540-28998-2; 3-540-28998-4. Reprint of the 1997 edition.

Karl-Theodor Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1): 65–131, 2006a. ISSN 0001-5962. doi: 10.1007/s11511-006-0002-8.

Karl-Theodor Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1): 133–177, 2006b. ISSN 0001-5962. doi: 10.1007/s11511-006-0003-7.

D. Trevisan. Lagrangian flows driven by BV fields in Wiener spaces. *To appear in Probability Theory and Related Fields*, 2014a. doi: 10.1007/s00440-014-0589-1.

D. Trevisan. A short proof of Stein's universal multiplier theorem. *To appear in Séminaire de Probabilités XLVI*, 2014b.

Dario Trevisan. BV-regularity for the Malliavin derivative of the maximum of a Wiener process. *Electron. Commun. Probab.*, 18:no. 29, 9, 2013a. ISSN 1083-589X. doi: 10.1214/ECP.v18-2314.

Dario Trevisan. Zero noise limits using local times. *Electron. Commun. Probab.*, 18:no. 31, 7, 2013b. ISSN 1083-589X. doi: 10.1214/ECP.v18-2587.

N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon. *Analysis and geometry on groups*, volume 100 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1992. ISBN 0-521-35382-3.

A. Ju. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb. (N.S.)*, 111(153)(3):434–452, 480, 1980. ISSN 0368-8666.

Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. ISBN 978-3-540-71049-3. doi: 10.1007/978-3-540-71050-9. Old and new.

Nik Weaver. Lipschitz algebras and derivations. II. Exterior differentiation. *J. Funct. Anal.*, 178(1):64–112, 2000. ISSN 0022-1236. doi: 10.1006/jfan.2000.3637.

Guofang Wei and Will Wylie. Comparison geometry for the Bakry-Emery Ricci tensor. *J. Differential Geom.*, 83(2):377–405, 2009. ISSN 0022-040X.

Toshio Yamada and Shinzo Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11:155–167, 1971. ISSN 0023-608X.

Kōsaku Yosida. *Functional analysis*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. ISBN 3-540-58654-7. Reprint of the sixth (1980) edition.

Xicheng Zhang. Stochastic flows of SDEs with irregular coefficients and stochastic transport equations. *Bull. Sci. Math.*, 134(4):340–378, 2010. ISSN 0007-4497. doi: 10.1016/j.bulsci.2009.12.004.