DE GRUYTER

Claude Le Bris, Pierre-Louis Lions

PARABOLIC EQUATIONS WITH IRREGULAR DATA AND RELATED ISSUES

APPLICATIONS TO STOCHASTIC DIFFERENTIAL EQUATIONS

SERIES IN APPLIED AND NUMERICAL MATHEMATICS 4

Claude Le Bris, Pierre-Louis Lions

Parabolic Equations with Irregular Data and Related Issues

De Gruyter Series in Applied and Numerical Mathematics

Volume 4

Claude Le Bris, Pierre-Louis Lions

Parabolic Equations with Irregular Data and Related Issues

Applications to Stochastic Differential Equations

DE GRUYTER

Mathematics Subject Classification 2010
35Kxx, 60Gxx, 60Jxx

Authors
Prof. Dr. Claude Le Bris
Ecole des Ponts ParisTech and INRIA
Centre d'Enseignement
et de Recherche en Mathématiques
et Calcul Scientifique
6/8 Avenue Blaise Pascal
77455 Marne-La-Vallée Cedex 2
France
claude.le-bris@enpc.fr

Prof. Dr. Pierre-Louis Lions
Collège de France
11, place Marcelin Berthelot
75231 Paris Cedex 05
France
and CEREMADE
Université Paris-Dauphine
Place du Maréchal De Lattre De Tassigny
75775 Paris Cedex 16
France
lions@ceremade.dauphine.fr

ISBN 978-3-11-063313-9
e-ISBN (PDF) 978-3-11-063550-8
e-ISBN (EPUB) 978-3-11-063314-6
ISSN 2512-1820

Library of Congress Control Number: 2019934287

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Typesetting: Dimler & Albroscheit, Müncheberg
Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Contents

Introduction

We study here the existence and the uniqueness of solutions to a large class of parabolic type equations with irregular coefficients and/or initial conditions

$$\begin{cases} \partial_t u - b_i \partial_i u - a_{ij} \partial_{ij}^2 u = 0, \\ u|_{t=0} = u_0. \end{cases} \tag{1}$$

In (1), the solution u is a real-valued function of time t and space $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $1 \leq d < +\infty$. The partial derivative with respect to x_i is denoted by ∂_i, that with respect to t by ∂_t. The data are the coefficients $\mathbf{a} = [a_{ij}]_{1 \leq i,j \leq d}$, $\mathbf{b} = (b_1, \ldots, b_d)$, and the initial condition u_0. We use in (1), and unless otherwise stated throughout these notes, Einstein's convention of summation over repeated indices, thus, for instance, $\mathbf{b} . \nabla u = b_i \partial_i u$. The equation is, for the time being, formally posed on $\mathbb{R}^d$ and for $t \in [0, T]$. Our major assumption on the matrix $\mathbf{a}$ will be made precise shortly. We will also consider at length the variant of equation (1) where $a_{ij} \partial_{ij}^2 u$ is replaced by $\partial_i (a_{ij} \partial_j u)$ (the so-called equation in *divergence form*) and, more incidentally, the variant with $\partial_{ij}^2 (a_{ij} u)$.

The purpose of the present contribution is to understand the issues of existence and uniqueness of the solution to equation (1) (and its variants) with *minimal* assumptions of regularity on the coefficients **a** and **b**, and on the initial condition u_0.

We mention at once that the approach described here does not only establish existence and uniqueness of the solution, but also, and this is an important feature, the continuous dependence of the solution u, in the suitable functional space, upon the initial condition u_0. Even though we will not make this precise hereafter, the techniques of proof additionally allow to obtain the uniqueness of the limit of regularized solutions, and the stability of the solution with respect to perturbations of the coefficients **a** and **b**.

As will be extensively discussed below, the question of the well-posedness of (1) is intimately related to the question of the *deformation of the Lebesgue measure* by the flow of the underlying ordinary differential equation $\dot{X} = \mathbf{b}(X)$ or stochastic differential equation $d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)\, dt + \sigma(\mathbf{X}_t)\, d\mathbf{W}_t$ when $\mathbf{a} = \frac{1}{2} \sigma \sigma^t$ (or similar differential equations for equations with coefficients expressed in terms of **a** and **b**). In the case of *transport equations* (that is, $\mathbf{a} = 0$ in (1)), this question has been originally addressed, in the regular setting, by Liouville himself. It has been substantially studied, for Sobolev regular **b**, by the second author in collaboration with R. J. DiPerna in the work [41], and later, in [5] and subsequent works, by L. Ambrosio and his collaborators, for **b** with bounded variations. In the case of genuinely parabolic equations, that is, $\mathbf{a} \not\equiv 0$ in (1), *and* non-regular coefficients or initial condition, this question has not attracted much attention. This is the purpose of the present contribution to mathematically clarify all these issues. Related issues on the underlying stochastic differential equations, and on the elliptic variant of equation (1), will also be addressed.

https://doi.org/10.1515/9783110635508-203

Main mathematical assumptions. Before we proceed, we now make slightly more precise the above setting.

Our main, and actually in a sense *only* assumption is that the second-order coefficients a_{ij} form a *nonnegative symmetric* matrix. We emphasize that one essential feature of the present work is that this coefficient matrix is not always assumed uniformly elliptic. Both cases, uniformly elliptic and possibly degenerate, will be considered. In particular, our setting covers the setting $a \equiv 0$ of transport equations with an irregular coefficient b or initial condition u_0, a setting originally studied in [41]. Our arguments and considerations are therefore strongly connected to those of [41]. Besides the assumption "**a** *nonnegative symmetric*", we will make, for simplicity of exposition, the following two classes of simplification:

- *Coefficients independent of time*. We will assume throughout these notes that the coefficients functions a_{ij}, b_i, $1 \leq i, j \leq d$, are functions of the space variable x only and therefore that they do not depend on time t. Unless otherwise stated, our arguments carry over to the case of time-dependent coefficients a_{ij} and b_i, provided we assume their spatial regularity or integrability suitably depends on the time variable, typically in an L^1 manner (although it may occasionally happen that it should be in an L^2 or even an L^∞ manner, but these technicalities are not central to our discussion). We will however not go in this direction. Likewise, a non-zero right-hand side f, depending in a suitable manner of possibly both space and time, can be inserted in (1). We typically think of f being L^1 in time and valued in the space L^p, for some $1 \leq p \leq +\infty$, when the initial condition u_0 is L^p for the same p. Again we leave this extension aside.
- *Periodic boundary conditions*. In order to avoid all technicalities related to boundary conditions, we will assume that (1) is set on a bounded domain with periodic boundary conditions and that all the functions we manipulate are correspondingly periodic. The case when the equation is set on the entire space $\mathbb{R}^d$ requires additional assumptions regarding the growth at infinity of all the coefficients functions. One typically assumes for (1) that $\frac{\mathbf{a}}{1+|x|^2}$ and $\frac{\mathbf{b}}{1+|x|}$ belong to $L^1 + L^\infty$. This growth control is then used in a Gronwall-type argument to estimate the tails of the integrals. This extension of the results to the case of $\mathbb{R}^d$ is tedious, but does not bring any substantial additional mathematical difficulty. Arguments of that type have been conducted in [41, 64, 65]. We refer the reader not familiar with such issues to those contributions. Likewise, bounded domains with specific (other than periodic) boundary conditions might require some additional work. We shall skip these technicalities here (except for some specific remarks spread along our text).

Motivation. There are of course many practically relevant contexts where equations of the form (1) arise, with possibly irregular data. For most models of Physics, it is not rare that the data do not have all the necessary regularity for (1) to be considered in a classical setting, say with all derivatives up to second order making classical sense. There are also many (and actually many more) cases where no information is available beforehand on the regularity of these data. Having a theory of well-posedness of (1) that accommodates "sufficiently irregular" data is therefore of major practical importance. Related to this ubiquity of this parabolic-type equation in modelling is the following general mathematical observation. An equation such as (1) preserves positiveness. This claim holds true at least formally and this may be rigorously shown under suitable assumptions. Given the absence of zeroth-order term, the equation admits constant functions as particular solutions. Additionally, for time-independent coefficients such as those we consider here, it is expected that the solutions to (1) generate a solution semi-group (one may solve from 0 to t, use the instantaneous value at time t as an initial condition, solve again from 0 to s, in order to obtain the solution from 0 to $t + s$). From semi-group theory (works by W. Feller, E. Hille, K. Yosida, etc.), we actually know as a general property the fact that if, in a suitable functional setting, we consider a linear semi-group that preserves positiveness and preserves constants, then this semi-group corresponds to an evolution equation of the form (1), at least up to nonlocal terms (which we omit throughout these notes). Additionally, we evidently need to mention the links between the theory for (1) and various other mathematical theories: the theory of linear transport equations (when the coefficient a vanishes in (1)), the theory of parabolic partial differential equations (1) with *regular* coefficients, and the theories of the underlying differential equations, either ordinary (in the case of transport equations), or stochastic (in the case $a_{ij} = \frac{1}{2}\sigma_{ik}\sigma_{jk}$ for a certain dispersion matrix σ_{ik}). For all these theories, the consideration of non-regular data is likewise a crucial issue. In addition, and for instance for the latter topic – ordinary or stochastic differential equations –, we would also like to note that, even in the case of coefficients sufficiently regular in order to uniquely define the solution to the differential equation itself, the question of less regular coefficients readily arises when considering linear tangent flows (that is, derivatives with respect to initial conditions), an obviously equally important mathematical issue.

Plan of our contribution. Our contribution is articulated as follows.

To begin with, we recall in Chapter 1 two issues related to the questions we examine: the theory for transport equations in Section 1.1 and the connection with the theory of stochastic differential equations in Section 1.2. We also briefly recall in Section 1.3 some existing results on the specific question we address in this contribution for (1). Although the reader familiar with all these somewhat elementary and now classical considerations may skip the section and directly proceed to the subsequent sections of our work, we would like to recommend at least a cursory reading of Chapter 1. We have indeed deliberately presented some more elaborate and general considerations and also mentioned some open mathematical questions. Note also that we outline the techniques and results in Section 1.3.3.

The new results begin with Chapter 2 which we devote to the study of equation (1) when the nonnegative symmetric second-order operator **a** has *constant* coefficients. In that case, the existing results on linear transport equation can be readily applied, but we wish to do better and completely analyze the specificities of the situation in the presence of the second-order operator. The first part of Chapter 2, namely Section 2.1, considers the case of a bounded initial condition. The main result is stated in Theorem 1. In short, the whole section is devoted to establishing this result along with corollaries, variants and extensions of this result. A particular attention is paid to a delicate estimate in L^1, in Section 2.1.3. The consequences of Theorem 1 on the theory of stochastic differential equations are specifically examined in Section 2.2. Then, using a specific renormalization procedure well adapted to parabolic equations, the results of Section 2.1 are adapted in Section 2.3 to an initial condition that is only in L^p, $1 \leq p < +\infty$. Miscellaneous remarks, comments and extensions are collected in Section 2.4.

With Chapter 3, we proceed to cases when the second-order coefficient matrix $\mathbf{a}$ in (1) is not constant. The chapter is devoted to the cases when the second-order operator is in divergence form. Chapter 4, will then, in particular, address equations in non-divergence form. Chapter 3 begins with an outline of the results obtained in [65] on that setting, and describes the main technical tools developed (essentially concerning the regularization procedure). The results of [65] are recalled, and improved, first in the case of possibly degenerate, and next positive definite second-order terms. These are respectively the contents of Sections 3.1 and 3.2. After some remarks, in Section 3.3, and a detailed new discussion of the L^1 estimate in Section 3.4, further extensions are exposed in Section 3.5, with a particular focus on the link between the questions we examine and the theory of hypo-ellipticity.

Chapter 4 contains various comments and extensions. It mentions some mathematical questions left open when these notes are being written. As announced, Section 4.1 addresses equations with varying coefficients $\mathbf{a}$ that do not correspond to equations in divergence form. The consequences on the theory of well-posedness for stochastic differential equations are discussed in Section 4.2. Section 4.3 briefly mentions further extensions, in particular towards the possible adaptation of the techniques and the results of these notes to *nonlinear* conservation laws and related equations.

We emphasize that the new results presented in these notes may be seen as significant extensions of some earlier results of ours essentially contained in [64, 65] and briefly recalled in Section 1.3. In [64], the second-order term in (1) was taken equal to the Laplacian (or, equivalently, to a constant elliptic operator) and the initial condition assumed L^∞. More stringent conditions on the field $\mathbf{b}$ than those considered in the present contribution were assumed therein. In [65], second-order terms more general than the Laplacian were considered but again the setting was not so general as it is here (for what concerns the possible degeneracy of the second-order operator, the regularity of the coefficient $\mathbf{b}$ and that of the initial condition u_0). We also mention that some partial results were obtained in related works such as [40, 70] where some specific parabolic-type equations were addressed. Some arguments we develop here are reminiscent from arguments of those works. The present notes do not only collect the results spread in all those previous works but significantly extend them. They also offer a unified, systematic, self-contained view of the problem. They clarify the link with related issues and point out several unsolved mathematical questions. Notice that, in order to keep these *lecture notes* (we emphasize that terminology!) as pedagogic as possible, we have chosen, in several instances, to only *sketch* the proofs. We deliberately spare to the reader some technicalities, which are certainly required to make the exposition entirely rigorous, which would be needed in a *research article*, but which, on the other hand, would make the exposition unnecessarily tedious. We also provide some background on the classical functional analysis and analysis of partial, ordinary and stochastic differential equations, so as to make these *lecture notes* as self-contained as possible. Notice also that the bibliography at the end of these notes contains more references on the topic than those explicitly cited within the text.

In short, these notes, written up jointly by the two authors, lay out the background on the various issues and present the recent results obtained by the second author. They are an expanded version of the lectures [72] delivered at Collège de France during the academic year 2012–13.

Acknowledgements. The authors would like to thank a number of colleagues for helpful discussions and for useful remarks on a preliminary version of these notes: Xavier Blanc, François Bouchut, Pierre Cardaliaguet, Nicolas Champagnat, Pierre-Emmanuel Jabin, Tony Lelièvre, Francis Nier, Mathias Rousset, Panagiotis Souganidis. They would also like to thank Jean-Michel Coron for his editorial work and the two anonymous referees for their constructive comments. The final stage of the writing of these notes has been completed ultimately when the authors were visiting the University of Chicago. The hospitality of this institution is gratefully acknowledged.

1 General context

As announced above, we recall in this chapter some useful elements of the theory of transport equations, and of ordinary and stochastic differential equations, along with some classical (and sometimes less classical) results known to date on parabolic-type equations.

1.1 Transport equations and ordinary differential equations

Our first set of remarks concerns the case where the second-order term vanishes, that is, $\mathbf{a} \equiv 0$ in (1). In that setting, the purpose of our mathematical endeavor, namely studying the well-posedness of the equation for non-regular coefficients and initial conditions is a now well established topic.

Some general considerations. It is well known that the linear transport equation

$$\partial_t u - b_i \partial_i u = 0 \tag{1.1.1}$$

is formally associated to the ordinary differential equation

$$\begin{cases} \dot{X} = \mathbf{b}(X), \\ X(t=0) = x, \end{cases} \tag{1.1.2}$$

and that the solution u to (1.1.1) is expected to read

$$u(t,x) = u_0(X(t,x)), \tag{1.1.3}$$

where $X(t,x)$ denotes the solution to (1.1.2). The connection between the partial differential equation (1.1.1) and the differential equation (1.1.2) is indeed ensured by the *method of characteristics*, also called *method of lines*, that is, formula (1.1.3). Showing the formal correspondence is elementary and giving a rigorous meaning to the manipulations performed is likewise easy in the presence of all the necessary regularity assumptions. One observes that, fixing $t > 0$ and next picking $s \in [0,t]$,

$$\frac{\partial}{\partial s} u(t-s, X(s,x)) = -\frac{\partial u}{\partial t}(t-s, X(s,x)) + \dot{X}(s)\nabla u(t-s, X(s,x)) = 0, \tag{1.1.4}$$

because of (1.1.1)–(1.1.2), which, integrated from $s = 0$ to $s = t$, proves (1.1.3). When one knows how to solve (1.1.2), one may thus solve (1.1.1) using (1.1.3). Conversely, solving (1.1.1) for *all* initial conditions u_0 allows to solve it in particular for the coordinate fields and thus to entirely know the solution $X(t,x)$ to (1.1.2) again using (1.1.3). Note that, in line with our introductory remarks of the previous section, we consider for simplicity a time-independent coefficient b in the linear transport equation (1.1.1). The adaptation of the above correspondence in the time-dependent case is immediate and well known.

https://doi.org/10.1515/9783110635508-001

The Cauchy–Lipschitz theory of ordinary differential equations allows to make rigorous the above formal discussion. For $\mathbf{b} \in W^{1,\infty}$, that is, $\mathbf{b}$ Lipschitz continuous, it is well known that one may uniquely define a solution to the dynamical system (1.1.2). Remark that, since $\mathbf{b}$ is valued in a d-dimensional space, we should more appropriately write $\mathbf{b} \in (W^{1,\infty})^d$ but we will make this slight abuse of notation throughout these notes. The well-posedness of the dynamical system in turn allows to solve (1.1.1) for $u_0 \in L^p$, $1 \leq p \leq +\infty$, using the above correspondence. The solution is continuous in time and valued in L^p for $1 \leq p < +\infty$ (and bounded in time with values in L^∞_x when u_0 is bounded). Put differently, there exists a semi-group flow solution to (1.1.1). The fact that such a regularity of $\mathbf{b}$ allows to adequately define a unique solution stems from the fact that, then, the dynamical system is well behaved. We indeed have the following a priori L^p estimate on the solution u to (1.1.1):

$$\int |u(x,t)|^p \, dx = \int |u_0(X(t,x))|^p \, dx = \int |u_0(x)|^p \, J^{-1} \, dx, \tag{1.1.5}$$

where the evolution of the *Jacobian* $J = |\det \frac{\partial X(t,x)}{\partial x}|$ is ruled by the equation

$$\frac{d}{dt}\left(\det \frac{\partial X(t,x)}{\partial x}\right) = \operatorname{div} \mathbf{b}(X(t,x)) \det \frac{\partial X(t,x)}{\partial x}. \tag{1.1.6}$$

The Lebesgue measure is thus transported by the flow $\operatorname{div} \mathbf{b}$. When $\operatorname{div} \mathbf{b}$ is bounded, the evolution (expansion or contraction) of the Lebesgue measure is thus controlled, hence the well-posedness of the dynamical system, and consequently that of the transport equation. Note that, more precisely, (1.1.6) gives an *exponential* deformation

$$e^{-Ct} \leq J \leq e^{Ct} \tag{1.1.7}$$

of the Lebesgue measure by the flow.

The above discussion raises the following question. Of course, $\operatorname{div} \mathbf{b}$ is bounded *in particular* when $\mathbf{b} \in W^{1,\infty}$, but is it possible to uniquely define a solution (both to the dynamical system and the linear transport equation) *only* assuming $\operatorname{div} \mathbf{b}$ is bounded and not necessarily $\mathbf{b} \in W^{1,\infty}$? In that case, solving (1.1.1) using the solution to (1.1.2) is unclear, since for $\mathbf{b}$ not Lipschitz continuous, solving the latter equation is itself unclear. The difficulty culminates in the observation that, say for $\mathbf{b}$ not continuous, $\mathbf{b}$ is only almost everywhere defined and thus $\mathbf{b}(X(t,x))$ does not necessarily make sense in (1.1.2) if X happens to accidentally weight sets of zero Lebesgue measure.

In answer to the latter question, the correspondence between the transport equation and the differential equation for regular fields $\mathbf{b}$ was extended by R. DiPerna and the second author in [41] in order to define a notion of generalized flow of solutions for ordinary differential equations with Sobolev coefficients. To cut a long story short, it was proved in [41] (and it will be briefly summarized below) that for $\mathbf{b} \in W^{1,1}$, $\operatorname{div} \mathbf{b} \in L^\infty$, one may define a solution flow for (1.1.2) (which, as above, gives solutions continuous in time, valued in L^p), precisely because, under the same assumptions, one may solve (1.1.1). This amounts to solving (1.1.2) for *almost all* initial conditions x, and not *all* initial conditions. This is the price to pay for considering less regular fields $\mathbf{b}$:

pathological behaviors cannot be excluded, but they concern a set of initial data of zero Lebesgue measure, and remain of zero Lebesgue measure as time evolves (this is the essence of the condition on div **b**). The proof falls in essentially two steps: *formal a priori estimates* (which would suffice to show the well-posedness if regularity were present) and *regularization* procedure (using a convolution with a regularizing kernel) in order to make the estimates rigorous. When the transport equation is solved on the whole space $\mathbb{R}^d$ instead of a bounded domain with periodic boundary conditions, the assumptions above need to be complemented with the assumption $\frac{\mathbf{b}}{1+|x|} \in L^1 + L^\infty$ in order for the proof to apply.

Before we outline the elements of proof for the above claim, a few remarks, on various extensions, are in order.

The contribution [41] already contains several remarks about the “optimality” of the couple of conditions ($\mathbf{b} \in W^{1,1}$, $\operatorname{div} \mathbf{b}$ bounded). In particular, counterexamples to uniqueness were exhibited for $\mathbf{b} \in W^{1,1}$ with unbounded divergences, or for divergence-free $\mathbf{b} \in W^{s,1}$ with arbitrary $s < 1$. On the other hand, the result can be extended in the following directions:

(i) One-sided control of div **b**. A glimpse at (1.1.6) (and this will be confirmed in the outline of the proof below) shows that for solving the linear transport equation for *increasing* times t, controlling the negative part $(\operatorname{div} \mathbf{b})_-$ of the divergence is sufficient: $(\operatorname{div} \mathbf{b})_- \in L^\infty$ allows to conclude. On (1.1.6), we indeed observe that, for increasing times, the positive part of div **b** contributes, to the *expansion* of the trajectories and the negative part to their *contraction*, respectively. The latter phenomenon could lead trajectories to eventually weight sets of zero Lebesgue measure, which could possibly create problems for fields **b** only almost everywhere defined. We take this opportunity to mention that the **conservative form** of (1.1.1), namely $\partial_t u - \partial_i (u\, b_i) = 0$ may be treated by arguments similar to those of the sequel. Since (e.g.) $\int u \partial_i (u b_i) = - \int u (b_i \partial_i u)$, the proof proceeds likewise, changing div **b** into its opposite.

(ii) Correspondence PDE/ODE for $\mathbf{b} \in L^1$. It was proved in [71] that for all L^1 fields **b** (that is, not necessarily $W^{1,1}$ but satisfying the other condition(s) above), the existence and uniqueness of the *generalized* flow, in the sense of [41], is equivalent to the existence and uniqueness of the solution to the transport equation. Any additional property that ensures one of the two facts then implies the other. Some related issues are examined in [24].

(iii) Integrability of the symmetric part of $\nabla \mathbf{b}$. It was observed in [29], and it can be deduced from the proof outlined below, that, instead of a $W^{1,1}$ regularity of **b**, an L^1 integrability of **b** itself along with that of the *symmetric* part of $\nabla \mathbf{b}$, namely $\partial_i b_j + \partial_j b_i$, $1 \le i, j \le d$, are sufficient to perform the regularization step (using a more specific regularization kernel), and thus to obtain the result. More recently, it has been established that the result also holds when the **antisymmetric part of $\nabla \mathbf{b}$ is** L^1 (along with div **b**). This is a consequence of [25] addressing the case when the gradient of **b** is the singular integral of an L^1 function. This work is in particular motivated by the two-dimensional Euler system and the Vlasov-Poisson system [21–23].

(iv) Piecewise $W^{1,1}$ regularity of **b**. Similarly to the previous comment, it was noticed in [71] that one could accommodate vector fields **b** that are only piecewise-$W^{1,1}$, with suitable compatibility conditions along the locations of discontinuities so that div **b** is globally bounded, using a specific regularization kernel (varying differently in the directions tangent and normal to the "jumps"). Arguments along those lines allow to treat even more general vector fields **b**. This is the purpose of our next item.

The most important addition to the results of [41] is the following result:

(v) BV vector fields **b**. The extension of the well-posedness theory of [41] to BV vector fields instead of $W^{1,1}$ vector fields, was achieved in [5] (and further works of the same author with different collaborators; see e.g. [9], see also more recent references of the same authors).

Besides the work [5], many contributions by many scientists and in various research directions, have followed up on [41]. It is hopeless to try and summarize all of them. We begin with some of our own works:

(vi) Self-contained proof of the well-posedness of the dynamical system. The dynamical system (1.1.2) was studied in [41] using the results previously established for the linear transport equation (1.1.1) and the correspondence (1.1.3) between the two equations. In [54], a proof is performed in a self-contained way, only using the dynamical system itself. It is extended in [53] to BV vector fields.

(vii) Sobolev differentiability of the flow of the dynamical system with respect to the initial conditions of the dynamical system. Differentiability with respect to the initial conditions (under further regularity assumptions on the vector field **b**) for the solution to (1.1.2) is part of the Cauchy–Lipschitz theory in the regular setting. Similarly, analogous differentiability properties, in the framework of Sobolev spaces, were established in [64] using the particular form of the linear transport equation associated to the tangent flow. The extension of the arguments on the transport equation to the BV case has been considered in [66], but the consequences on the theory of stochastic differential equations were not specifically discussed therein. We will return to related questions below, see Section 4.3.

And we would also like to cite, as *examples* of the many recent contributions by other researchers: the extension to the case div **b** $\in$BMO in [79], the Hamiltonian setting considered by several authors and in particular in the recent study [32], etc. We also point out the lecture notes [7] summarizing the state of the art at the time of their publication, as well as [36] and [6] for slightly more recent accounts.

We mention at the end of this chapter some mathematical issues that, among many, remain unsolved.

Proof of well-posedness of transport equations with Sobolev regular coefficients. We outline here the main ingredients of the proof contained in [41], not only for self consistency of the present notes but also because these arguments will form the bottom line of the arguments we will perform on the case of equation (1). For our exposition, we recall we work on a bounded domain with periodic boundary conditions. We will additionally assume that the vector field $\mathbf{b}$ satisfies $\mathbf{b} \in W^{1,1}$ and $\operatorname{div}\mathbf{b}$ bounded.

First observe that it is easy to understand the time derivative in (1.1.1) in the sense of distributions but that the term $b_i\partial_i u$ is much more difficult since $\partial_i u$ has no regularity a priori. It is thus useful to write this term as $b_i\partial_i u = \partial_i(b_i u) - (\partial_i b_i)u$. Our assumption on the boundedness of $\operatorname{div}\mathbf{b}$ allows us to give the latter term a sense, but the former term is not as simple in an arbitrary large dimension, unless we have $u \in L^\infty$. To this end, we therefore assume, only for simplicity because other more general cases may indeed be addressed, that $u_0 \in L^\infty$ (and then we will rely on the maximum principle to obtain the desired bound on u itself). We however mention that, for unbounded initial conditions u_0, the technique introduced in [41] is to consider the notion of *renormalized* solution, that is, in short, say u is a solution when $\Phi(u)$ is a solution for the initial condition $\Phi(u_0)$, for all appropriate *bounded* functions Φ that are in addition continuous, C^1, or smooth. We already note that a similar technique will be employed in the present contribution for (1), see Section 2.3.

Observing the linearity of equation (1.1.1), one immediately realizes that the major issue is not existence, but uniqueness. The proof of uniqueness performed in [41] is articulated in two steps: *a priori estimates* (which, in passing, also provide existence with elementary approximation arguments) and *regularization*. The bound on $\operatorname{div}\mathbf{b}$ is used for the a priori estimates, the $W^{1,1}$ regularity is employed for the regularization step. We now briefly expand the outline of the proof, emphasizing its main steps and underlying ideas.

The first step consists in establishing *formal a priori estimates* on the tentative solution u. This is performed by multiplying the equation by some function $\beta'(u)$ (where β is some convenient *renormalization* function) and integrating by parts. Formally, this procedure yields

$$\frac{d}{dt}\int \beta(u) + \int (\operatorname{div}\mathbf{b})\beta(u) = 0. \tag{1.1.8}$$

Therefore, when $\operatorname{div}\mathbf{b}$ is L^∞ (and actually for positive times, only a control of the negative part is needed, as already pointed out above), we obtain $\int \beta(u)$ bounded for all times if it is bounded at initial time. In particular, $\beta(u) = |u|^p$ ($1 \le p < +\infty$) yields formal L^p bounds on the solution. Besides, the L^∞ bound is obtained by application of the maximum principle. The argument we have just outlined is the exact analogue to the argument originally performed by Liouville on the dynamical system itself (see (1.1.6)). Bearing in mind our specific study of parabolic-type equations, we observe, and we will return to this below, that the presence in the equation of a *constant nonnegative* second-order operator would not modify the above estimate (assuming the renormalization function β is convex).

Using this first step, *existence* of a solution is readily proved. The transport coefficient $\mathbf{b}$ is approximated by convolution $\mathbf{b}_\varepsilon = \rho_\varepsilon \star \mathbf{b}$, using some regularizing kernel $\rho_\varepsilon = \varepsilon^{-N}\rho(\varepsilon^{-1}\cdot)$, with $\rho \in \mathcal{D}(\mathbb{R}^N)$, $\rho \geq 0$, $\int \rho = 1$. The linear transport equation

$$\partial_t u_\varepsilon - (\mathbf{b}_\varepsilon)_i \partial_i u_\varepsilon = 0$$

admits a unique solution u_ε, by standard arguments. The above formal a priori estimate (1.1.8) can be rigorously established on u_ε:

$$\frac{d}{dt}\int \beta(u_\varepsilon) + \int (\operatorname{div}\mathbf{b})\beta(u_\varepsilon) = 0 \tag{1.1.9}$$

along with the L^∞ bound (maximum principle). As the equation is linear, passing to the (weak) limit provides a solution to (1.1.1) in a suitable functional space, typically $L^1 \cap L^\infty$ when the initial condition $u(t=0,\cdot)$ lies in that space. For this to hold, we only need $b_i \in L^1$ and $\operatorname{div}\mathbf{b} \in L^\infty$. The natural weak formulation of equation (1.1.1) also readily follows from the above argument.

The second major step is a *regularization procedure*. Estimate (1.1.8) has indeed been obtained *multiplying* $\partial_i u$ by $\beta'(u)$, a manipulation that is illicit with non-regular solutions. The necessary regularization is based upon the celebrated *commutation* lemma,which basically claims that

$$[\rho_\varepsilon, \mathbf{b}.\nabla](u) := \rho_\varepsilon \star (\mathbf{b}.\nabla u) - \mathbf{b}.\nabla(\rho_\varepsilon \star u) \xrightarrow{\varepsilon\to 0} 0 \tag{1.1.10}$$

in L^1 when, for instance, $\mathbf{b} \in W^{1,1}$ and $u \in L^\infty$. This lemma is a classical result of real analysis, sometimes called Friedrich's Lemma. We refer to [41, Lemma II.1], and in the present work, Lemmata 1 (page 32) and 10 (page 90) for precise statements. Detailed proofs, comments and applications may also be found, e.g., in [27, 80]. Notice also that (4.1.24)–(4.1.25) page 124 below formalizes the argument in an actually slightly more complex context. The need for some Sobolev regularity on $\mathbf{b}$ may be formally understood in the following manner: the above commutator basically involves a quantity of the form

$$\int u(y)\,(\mathbf{b}(y) - \mathbf{b}(x)).\nabla\rho_\varepsilon(x-y)\,dy, \tag{1.1.11}$$

where ρ_ε converges in distribution to the Dirac mass, thus the need for evaluating $\mathbf{b}(y) - \mathbf{b}(x)$ in terms of $y - x$ for $y - x$ small. This evaluation in turn requires a weak differentiability of $\mathbf{b}$. By duality, for $u \in L^\infty$, the derivative of $\mathbf{b}$ only needs to belong to L^1, thus the $W^{1,1}$ space is the appropriate space for $\mathbf{b}$.

We now briefly formalize the above sketch of proof of (1.1.10) for the reader's convenience. The same arguments show more generally this convergence in L^1 for $\mathbf{b} \in W^{1,p}$ and $u \in L^q$, $\frac{1}{p} + \frac{1}{q} = 1$, or the convergence in L^r, for $1 < r < +\infty$ for suitable conjugate exponents, along with various other convergences. We first write

$$\rho_\varepsilon \star (\mathbf{b}.\nabla u) - \mathbf{b}.\nabla(\rho_\varepsilon \star u) = \int u(y)((\mathbf{b}(y) - \mathbf{b}(x)).\nabla\rho_\varepsilon(x-y))\,dy - (u\operatorname{div}\mathbf{b}) \star \rho_\varepsilon,$$

where, since it is a convolution with a smooth approximation of the Dirac mass, the rightmost term converges to $u\operatorname{div}\mathbf{b}$ in L^1 as ε vanishes. On the other hand, the first term of the right-hand side converges also to $u \operatorname{div} \mathbf{b}$ if both u and $\mathbf{b}$ are smooth, simply noticing that $\nabla\rho(x-y) = -\nabla_y(\rho(x-y))$, next integrating by parts with the Green formula, and finally using again that ρ_ε converges to the Dirac mass in the sense of distributions. Proving (1.1.10) therefore amounts to arguing by density and estimating this first term in the suitable norms. For this purpose, we take $v \in L^p$ and $\mathbf{c} \in W^{1,q}$ arbitrary, $\frac{1}{p} + \frac{1}{q} = 1$, and ε sufficiently small so that

$$\begin{aligned}
&\left\| \int v(y)((\mathbf{c}(y) - \mathbf{c}(x)) \,.\, \nabla\rho_\varepsilon(x-y))\, dy \right\|_{L^1_x(B_R)} \\
&\quad \leq C \|v\|_{L^p_x(B_{R+1})} \left\| \int\limits_{|x-y|\leq C\varepsilon} \varepsilon^{-d-1} |\mathbf{c}(y) - \mathbf{c}(x)|\, dy \right\|_{L^q_x(B_{R+1})} \\
&\quad \leq C \|v\|_{L^p_x(B_{R+1})} \left(\int\limits_{B_{R+1}} dx \int\limits_{|z|\leq C} dz \left| \frac{\mathbf{c}(x+\varepsilon z) - \mathbf{c}(x)}{\varepsilon} \right|^q \right)^{\frac{1}{q}} \\
&\quad \leq C \|v\|_{L^p_x(B_{R+1})} \left(\int\limits_{B_{R+1}} dx \int\limits_{|z|\leq C} dz \left(\int_0^1 |\nabla \mathbf{c}(x+t\varepsilon z) \,.\, z|\, dt \right)^q \right)^{\frac{1}{q}} \\
&\quad \leq C \|v\|_{L^p_x(B_{R+1})} \|\nabla \mathbf{c}\|_{L^q(B_{R+1+C})},
\end{aligned} \tag{1.1.12}$$

using the fact that $\rho_\varepsilon = \varepsilon^{-d}\rho(\varepsilon^{-1}\cdot)$ is supported in a ball of radius of order ε, and denoting by C various irrelevant constants. The proof of (1.1.10) is now easy to complete. The field $\mathbf{b}$ is approximated by a sequence of smooth functions $\mathbf{b}_n$ converging to $\mathbf{b}$ in $W^{1,1}$. One cannot proceed exactly likewise for u since smooth functions are not dense in L^∞. Since $u \in L^\infty \subset L^2$, it is approximated by a sequence of smooth functions u_n so that $n\|\mathbf{b}_n\|_{W^{1,2}} \|u_n - u\|_{L^2} \leq 1$. The convergence (1.1.10) holds, for each n, for u_n and $\mathbf{b}_n$ as ε vanishes. On the other hand, (1.1.12) successively applied to ($v = u \in L^\infty$, $\mathbf{c} = \mathbf{b} - \mathbf{b}_n \in W^{1,1}$) and ($v = u - u_n \in L^2$, $\mathbf{c} = \mathbf{b}_n \in W^{1,2}$), along with a simple argument on the difference $(u \operatorname{div} \mathbf{b} - u_n \operatorname{div} \mathbf{b}_n) \star \rho_\varepsilon$, show that

$$\begin{aligned}
\|[\rho_\varepsilon, \mathbf{b} \,.\, \nabla](u) - [\rho_\varepsilon, \mathbf{b}_n \,.\, \nabla](u_n)\|_{L^1} &\leq C\|u\|_{L^\infty} \|\mathbf{b} - \mathbf{b}_n\|_{W^{1,1}} + C\|u - u_n\|_{L^2} \|\mathbf{b}_n\|_{W^{1,2}} \\
&\leq C\|u\|_{L^\infty} \|\mathbf{b} - \mathbf{b}_n\|_{W^{1,1}} + Cn^{-1}
\end{aligned}$$

for a constant C independent of $\varepsilon \leq 1$. The convergence (1.1.10) follows for u and $\mathbf{b}$.

We note in passing that the $W^{1,1}$ regularity may be relaxed into a BV regularity. This is the purpose of the work [5] already cited above. On the other hand, the question arises to know whether this BV regularity is the minimal one for the commutation property to hold; see our comments along this line in the next paragraph on page 10. *Uniqueness* readily follows from this commutation lemma by convolution: considering $f = g - h$ the difference of two solutions to (1.1.1), one takes the convolution of the transport equation (1.1.1) with ρ_ε, next obtains the same equation up to an error term, namely the right-hand side of

$$\partial_t f_\varepsilon - b_i \partial_i f_\varepsilon = [\rho_\varepsilon, \mathbf{b} \,.\, \nabla](f),$$

with additionally the fact that, now and because we have taken the difference of two solutions, the initial condition is zero. We next multiply both sides by $\beta'(f_\varepsilon)$ and integrate in the space variable. Letting next ε go to zero, using (1.1.10) and the bounds on f_ε, (1.1.8) is obtained. That estimate in turn yields uniqueness. We mention that we have left the regularizing kernel ρ_ε quite general, but making it more specific allows for the generalizations (iii), (iv), (v) previously mentioned.

As briefly mentioned in the introduction, the above proof for the transport equation has to be slightly adapted when the equation is posed on the whole space or on a bounded domain with specific Dirichlet boundary conditions (on the part of the boundary where $\mathbf{b}\,.\,\mathbf{n} \neq 0$, $\mathbf{n}$ denoting of course the normal to the boundary). For the former question, we refer the reader to [41]. For the latter question, we cite, e.g., [26] when $\mathbf{b}$ is Sobolev regular and [38] when $\mathbf{b}$ has bounded variation. For the parabolic equation on which we focus in these notes (starting with Section 1.3), we omit all such variations of the boundary conditions (which might be, in some cases, delicate to accommodate) and exclusively consider the periodic case.

Notice that, since existence is relatively easy and uniqueness is the major issue, one could think of proving uniqueness exploiting existence for the adjoint equation. We indeed recall that the classical argument (holding true for sufficiently regular data and solutions) proceeds as follows. The difference $u(t,x)$ of two solutions to (1.1.1) is again solution to (1.1.1), with zero initial condition $u(0,\cdot) = 0$. One thus introduces $v(s,x)$ solution to the adjoint equation

$$\frac{\partial}{\partial s} v + \operatorname{div}(\mathbf{b} v) = 0, \tag{1.1.13}$$

with initial condition $v(s = 0, \cdot) = u(t, \cdot)$ (if $\mathbf{b}$ depends on time then $\mathbf{b}(t - s)$ is to be considered in (1.1.13)). Then, writing

$$\begin{aligned}
\int |u(t,\cdot)|^2 &= \int u(t,\cdot)v(0,\cdot) - \int u(0,\cdot)v(t,\cdot) \\
&= -\int_0^t \int \frac{\partial}{\partial s}(u(t-s,\cdot)v(s,\cdot))\,ds \\
&= \int_0^t \int (\mathbf{b}\,.\,\nabla u(t-s,\cdot)v(s,\cdot) + u(t-s,\cdot)\operatorname{div}(\mathbf{b}\,v(s,\cdot)))\,ds,
\end{aligned} \tag{1.1.14}$$

where the last term, by the Green formula, vanishes. This shows $u(t,\cdot) = 0$ for all times t and thus uniqueness. The point is however that for functions u and v not sufficiently regular, the above adjoint calculus does not necessarily make sense. Giving a rigorous meaning to the above manipulations (multiplication of (1.1.1) by v and, likewise, of (1.1.13) by u, as well as use of the Green formula (that is, integration by parts)) actually requires the techniques – formal a priori estimates and regularization – exposed here. Therefore, using the classical adjoint viewpoint does not allow to circumvent the difficulty.

Some unsolved issues. As briefly mentioned above, many questions remain unsolved, even in the now classical setting of transport equations (that is, $\mathbf{a} \equiv 0$ in (1)). We would like to emphasize the following ones.

Questions such as the existence and uniqueness theory for an L^1 vector field $\mathbf{b}$ for which only the symmetric derivative $\partial_i b_j + \partial_j b_i$ is a bounded measure, or a divergence free vector field $\mathbf{b}$ for which only curl $\mathbf{b}$ is a bounded measure (a case that is very much practically relevant for fluid mechanics applications) are immediate possible extensions of the existing results to investigate. Results in this direction are contained in [21].

Furthermore, other related issues may be considered.

Growth conditions at infinity. One example of such an open question is related to the optimality of the condition $\mathbf{b} \in (1+|x|)(L^1(\mathbb{R}^d) + L^\infty(\mathbb{R}^d))$ classically imposed when the equation is set on the entire space $\mathbb{R}^d$. We are aware we have deliberately ruled out the question of behavior at infinity in these notes, restricting our study to a domain with periodic boundary condition. However, we would like to make this one exception, to show that the problem is not entirely closed in this respect either. Indeed, when $\mathbf{b}$ is assumed periodic (and of course as usual $W^{1,1}_{\mathrm{loc}}$, and, say, divergence-free), uniqueness of the *periodic* solution u to (1.1.1) of course holds, when the initial condition u_0 is assumed periodic. But one may prove uniqueness of the *non-necessarily periodic* solution u to (1.1.1) in $\mathbb{R}^d$, say in $L^1 \cap L^\infty$, *without* assuming $\mathbf{b} \in (1+|x|)(L^1 + L^\infty)$, for a general initial condition u_0, say in $L^1 \cap L^\infty$. The outline of the proof is as follows. We know by the usual argument that it suffices to prove that if $u \in L^1 \cap L^\infty$, $u \geq 0$ is a solution to the equation with zero initial condition, then $u = 0$. Consider the sum of translates

$$u_N = \sum_{|k|\leq N, k\in\mathbb{Z}^d} u(\cdot + k).$$

By linearity and periodicity of $\mathbf{b}$, it also solves (1.1.1). Since $u \geq 0$, we know, by monotone convergence, that $u_\infty = \lim_{N\to+\infty} u_N \in \mathbb{R}_+ \cup \{+\infty\}$ is in $L^1(Q)$ since

$$\int_{\mathbb{R}^d} u = \lim_{N\to+\infty} \int_Q u_N = \int_Q u_\infty,$$

where Q of course denotes the unit cell $[0,1[^d$. It immediately follows that u_∞ is almost everywhere finite. Note that in addition, again by monotonicity,

$$\int_Q |u_\infty - u_N| = \int_Q u_\infty - \int_Q u_N$$

and the latter quantity vanishes as $N \to +\infty$, thus we indeed have strong convergence in $L^1(Q)$. By construction and because of the above convergences, u_∞ is periodic. The point is, u_∞ is not necessarily in L^∞. So we remark that, since $u_N \in L^\infty$, u_N is also a renormalized solution to (1.1.1), so that $\frac{u_N}{1+u_N}$ solves (1.1.1). But clearly $\frac{u_N}{1+u_N}$ converges in L^1 to $\frac{u_\infty}{1+u_\infty} \in L^1 \cap L^\infty$, which is periodic and solves (1.1.1). Therefore, by the "periodic" result, it vanishes. This shows that $u_\infty = 0$, and that uniqueness holds. This observation clearly indicates that the *"best"* condition at infinity on $\mathbf{b}$ for uniqueness to hold is still unclear.

A similar statement actually applies to the diffusion coefficient $\mathbf{a}$ that we will be manipulating when studying parabolic equations such as (1). We only address the periodic case, but, should we consider the case when the equation is posed on the whole space $\mathbb{R}^d$, then we would have to put restrictions on the growth rate of $\mathbf{a}$ at infinity. The typical condition, present, e.g., in our previous work [64], is $\mathbf{a} \in (1+|x|^2)(L^1(\mathbb{R}^d) + L^\infty(\mathbb{R}^d))$ (which, when $\mathbf{a} = \sigma\sigma^t$ and in terms of the classical theory of stochastic differential equations, corresponds to the standard assumption $\sigma \in (1+|x|)L^\infty(\mathbb{R}^d)$). It turns out that, again, the best condition at infinity on $\mathbf{a}$, or σ, is unclear. For instance, in dimension $d = 1$, the equation $d\mathbf{X}_t = (1+\mathbf{X}_t^2)\,d\mathbf{W}_t$ has a unique strong solution although the coefficient $\sigma(x) = 1+|x|^2$ clearly violates the linear growth condition above. And the same claim actually holds for all equations $d\mathbf{X}_t = (1+\mathbf{X}_t^2)^{\frac{\alpha}{2}}\,d\mathbf{W}_t$, for α arbitrarily large (see [58, Section 5.5]).

BV regularity. Another interesting question, which we have just briefly mentioned above, is the question of the optimality of the BV assumption for the commutation property to hold. Put differently, assume that we have a divergence-free field $\mathbf{b}$ for which there exists a constant C such that, for all $u \in L^\infty$ and for all $\varepsilon > 0$, $\|[\rho_\varepsilon, \mathbf{b}\,.\,\nabla](u)\|_{L^1} \le C\|u\|_{L^\infty}$. The question then is: does this imply that $\mathbf{b} \in BV$? It turns out that the answer is positive, if one allows, in the construction of the regularization kernel ρ_ε, for *all* arbitrary smooth functions ρ. On the other hand, if one only considers radially symmetric smooth functions ρ, then the commutation only implies that the symmetric derivative $(D\mathbf{b})_s = (\partial_i b_j + \partial_j b_i)$ is a bounded measure.

Infinite-dimensional ambient space. A much more ambitious endeavor would be to try and adapt the theory to infinite-dimensional spaces. As we have already repeatedly emphasized the Lebesgue measure plays a leading role in the theory, and the intuitive interpretation of it. Besides, it also plays a crucial role in the technique of proof: the regularization is performed using a convolution, naturally involving the Lebesgue measure again. An adaptation to the infinite-dimensional context, where no Lebesgue measure exists, certainly requires significant developments. A natural idea along this direction is to consider Wiener spaces. Transport equations on such spaces have been considered in [8]. Some second-order operators were addressed in [18] and related works by the same authors, and also [76], etc. A variety of approaches are possible and many questions are open.

1.2 Stochastic differential equations

We now proceed to the case where $\mathbf{a} \not\equiv 0$ in (1), and specifically consider the case when the coefficient matrix $\mathbf{a}$ reads $\mathbf{a} = \frac{1}{2}\sigma\,\sigma^t$ for a possibly rectangular $d \times n$ matrix σ. The size n may possibly be $n = +\infty$. A particular case of this situation is the case when $\mathbf{a}$ is a symmetric positive definite matrix, and we may then write $\mathbf{a} = \frac{1}{2}\sigma\,\sigma^t$ with a square $d \times d$ matrix $\frac{1}{\sqrt{2}}\sigma = \sqrt{\mathbf{a}}$. In the change from $\mathbf{a}$ to σ and backward, some regularity is transmitted. When the matrix is positive definite, a and $\sqrt{a}$ share the same regularity. In the classical setting, one may refer, e.g., to [51, Lemma 1.1, Volume 1], where it is shown that $\sqrt{a}$ is C^m, resp. $C^{m,\alpha}$, if a is. When the matrix is only nonnegative, some regularity may be lost. In the classical setting, $a \in C^2$ implies only that σ is Lipschitz continuous for instance (see, e.g., [51, Theorem 1.2]) or, when $\mathbf{a} \in W^{2,\infty}$, we have $\sigma \in W^{1,\infty}$. In order to figure out why *some* regularity is lost, and in which manner it is lost, one just has to consider the case $\mathbf{a} = \frac{1}{2}|x|^2\,\mathrm{Id}$, $\sigma = |x|\,\mathrm{Id}$. Other cases of rectangular matrices σ are less clear for what regards regularity issues.

The classical Itô theory and related issues. Similarly to the correspondence between the transport equation (1.1.1) and the *ordinary* differential equation (1.1.2), equation (1) is, at least formally, related to the *stochastic* differential equation

$$\begin{cases} d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)\,dt + \sigma(\mathbf{X}_t)\,d\mathbf{W}_t, \\ \mathbf{X}_{t=0} = x. \end{cases} \tag{1.2.1}$$

Bearing in mind that stochastic differential calculus is based upon the Itô differentiation formula, the formal correspondence between the solution $\mathbf{X}_t^x$ to (1.2.1) and the solution u to (1) (with $\mathbf{a} = \frac{1}{2}\sigma\,\sigma^t$) is obtained by an argument similar to (1.1.4). Using the Itô formula, we indeed compute

$$\begin{aligned} d(u(t-s,\mathbf{X}_s)) = -\frac{\partial u}{\partial t}(t-s,\mathbf{X}_s)\,ds + \mathbf{b}(\mathbf{X}_s)\,.\,\nabla u(t-s,\mathbf{X}_s)\,ds \\ + \mathbf{a}(\mathbf{X}_s)\,.\,D^2 u(t-s,\mathbf{X}_s)\,ds + \sigma^t(\mathbf{X}_s)\nabla u(t-s,\mathbf{X}_s)\,.\,d\mathbf{W}_s. \end{aligned} \tag{1.2.2}$$

Integrating from 0 to t, taking the expectation of both sides and using the equation (1), we obtain the *Feynman–Kac formula*

$$u(t,x) = \mathbb{E}_x(u_0(\mathbf{X}_t)), \tag{1.2.3}$$

where, here and throughout these notes, the subscript x in $\mathbb{E}_x$ recalls that the expectation is taken for the fixed initial condition x of (1.2.1) over all the Brownian trajectories. It is evidently a generalization of (1.1.3) to the stochastic context.

Formally, if one knows a solution to (1.2.1), then one may solve (1) using (1.2.3). Actually, this formal argument is delicate to make rigorous when the coefficients are not smooth. We shall return to this. Conversely, if one knows the solution to (1) for all initial, sufficiently regular conditions u_0, then one knows (1.2.3), which in turn allows to uniquely characterize the law of $\mathbf{X}_t$ at all times. Because of the semi-group structure (that is, the Markov property), it also characterizes the joint laws such as that of $(\mathbf{X}_s, \mathbf{X}_t)$.

For a continuous in time Markov process (actually the Càdlàg property is sufficient), this completely characterizes the process in law. Note that it does *not* characterize the trajectories of $\mathbf{X}_t$ themselves, we will also return to this below.

There is actually an *adjoint* viewpoint to this discussion. If instead of (1) with $\mathbf{a} = \frac{1}{2}\sigma\,\sigma^t$, one considers the adjoint equation (namely the *Fokker–Planck equation* or *forward Kolmogorov equation*)

$$\partial_t p + \operatorname{div}(p\mathbf{b}) - \frac{1}{2}\partial^2_{ij}(\sigma_{ik}\sigma_{jk}p) = 0, \tag{1.2.4}$$

then, for regular coefficients, the fact that (1.2.4) admits at most one solution implies that two processes $\mathbf{X}_t$ solution to (1.2.1) sharing the same initial distribution at initial time have the same law for all times. Note that the latter property (uniqueness in law for all times) does not *per se* imply the *uniqueness-in-law* of the process $\mathbf{X}_t$ solution to (1.2.1). The ambiguity of the terminology should not be misleading. As above (and this will also be mentioned in the non-regular setting below, see our discussion page 51), it is only under additional assumptions on the process (semi-group structure and continuity in time) that we can then conclude to the uniqueness in law of the process itself.

The Itô theory for equations such as (1.2.1) plays the role of the Cauchy–Lipschitz theory for ordinary differential equations: for $\mathbf{b}$ and σ Lipschitz continuous, there exists a unique solution to (1.2.1) (if we work on an unbounded domain, which is *not* our setting here, additional appropriate growth - namely at most linear- conditions at infinity are required for σ and $\mathbf{b}$, see [58, p. 289] for a precise statement). Although we will not dwell into all the details of the probability setting (we refer to the excellent classical textbooks [57, 58, 83, 84, 87]) we wish to mention that the solution established by the Itô theory is called *strong* since it is constructed for a *given* probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, a *given* Brownian motion $\mathbf{W}_t$ (and possibly a *given* initial condition $\mathbf{X}_0$ replacing x). Its uniqueness holds *pathwise*, that is, $\mathbf{X}_t$ is unique. An alternate notion of solution to (1.2.1) is the notion of *weak* solution, for which $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, $\mathbf{W}_t$ (and possibly the actual random variable $\mathbf{X}_0$ corresponding to a given law of the initial condition) are themselves part of the solution, and not given beforehand.

Remark 1. We would like to mention that, throughout these notes, we assume that the ambient dimension d is strictly higher than one. The one-dimensional case, because there is a total order on the real line and thus the left-hand side and the right-hand side of any possible discontinuity point of the coefficients is well defined, allows for specific considerations and particular settings. For instance, the assumption of Lipschitz regularity of the coefficients for the Itô theory can be weakened (typically in a $C^{0,\frac{1}{2}}$ regularity). We will not proceed in this direction, given that the arguments cannot, in any event, be generalized to higher dimensions. The monograph [34] presents the specifics and collects results for existence and uniqueness of solutions to *one-dimensional* stochastic differential equations with non-regular coefficients.

In terms of (pathwise) uniqueness, the essential arguments of the classical Itô theory go as follows. Denoting by $\mathbf{X}_t^x$ the solution to (1.2.1), and likewise $\mathbf{Y}_t^y$ the solution for another initial condition y instead of x, and intending to prove that $\mathbf{X}_t^x = \mathbf{Y}_t^x$ almost surely, we may argue (i) directly on the stochastic differential equation (1.2.1) using Itô calculus, (ii) using the Fokker–Planck equation (1.2.8) in a space of doubled dimension, or (iii) using the backward Kolmogorov equation (1.2.11) (adjoint to the Fokker–Planck equation). Of course, all these viewpoints are equivalent.

In the first viewpoint, we consider (1.2.1) itself and, using the Itô formula, we compute

$$\begin{aligned}\frac{d}{dt}\mathbb{E}(|\mathbf{X}_t^x - \mathbf{Y}_t^y|^2) &= \mathbb{E}\Big(\frac{d}{dt}|\mathbf{X}_t^x - \mathbf{Y}_t^y|^2\Big)\\ &= \mathbb{E}\big(2(\mathbf{X}_t^x - \mathbf{Y}_t^y)\,.\,(\mathbf{b}(\mathbf{X}_t^x) - \mathbf{b}(\mathbf{Y}_t^y))\\ &\qquad + (\sigma(\mathbf{X}_t^x) - \sigma(\mathbf{Y}_t^y))\,.\,(\sigma^T(\mathbf{X}_t^x) - \sigma^T(\mathbf{Y}_t^y))\big).\end{aligned} \tag{1.2.5}$$

Assuming the setting is regular, that is, $\mathbf{b}$ and σ Lipschitz continuous, we infer

$$\frac{d}{dt}\mathbb{E}(|\mathbf{X}_t^x - \mathbf{Y}_t^y|^2) \leq C_0\,\mathbb{E}(|\mathbf{X}_t^x - \mathbf{Y}_t^y|^2), \tag{1.2.6}$$

which yields

$$\mathbb{E}(|\mathbf{X}_t^x - \mathbf{Y}_t^y|^2) \leq e^{C_0\,t}\,|x - y|^2, \tag{1.2.7}$$

and thus implies uniqueness of the solution $\mathbf{X}_t^x$. A minoration by $e^{-C_0\,t}\,|x-y|^2$ may be similarly obtained from (1.2.5), thereby showing that the two quantities $|\mathbf{X}_t^x - \mathbf{Y}_t^y|$ and $|x-y|$ are actually of comparable order.

We now show pathwise uniqueness using the viewpoint of partial differential equations, and we may proceed in two different manners. The notion of pathwise uniqueness is, like that of uniqueness-in-law, related to equations of type (1). Assume indeed uniqueness of the solution $p = p(t,x,y)$ to the following Fokker–Planck equation in dimension $2d$:

$$\begin{aligned}\frac{\partial p(t,x,y)}{\partial t} &+ \operatorname{div}_x(p(t,x,y)\mathbf{b}(x)) + \operatorname{div}_y(p(t,x,y)\mathbf{b}(y))\\ &- \frac{1}{2}\frac{\partial^2}{\partial x_i\partial x_j}(\sigma_{ik}(x)\sigma_{jk}(x)\,p) - \frac{\partial^2}{\partial x_i\partial y_j}(\sigma_{ik}(x)\sigma_{jk}(y)\,p)\\ &- \frac{1}{2}\frac{\partial^2}{\partial y_i\partial y_j}(\sigma_{ik}(y)\sigma_{jk}(y)\,p) = 0.\end{aligned} \tag{1.2.8}$$

This equation is of course associated with the stochastic differential equation

$$d\begin{pmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{pmatrix} = \mathbf{B}(\mathbf{X}_t,\mathbf{Y}_t) + \mathbf{\Sigma}(\mathbf{X}_t,\mathbf{Y}_t)d\begin{pmatrix}\mathbf{W}_t\\ \overline{\mathbf{W}}_t\end{pmatrix} \tag{1.2.9}$$

set in $\mathbb{R}^{2d}$, for the Brownian motion $(\mathbf{W}_t, \overline{\mathbf{W}}_t)$ in $\mathbb{R}^{2d}$, with the following drift and *non-symmetric degenerate* dispersion matrix:

$$\mathbf{B}(x,y) = \begin{pmatrix}\mathbf{b}(x)\\ \mathbf{b}(y)\end{pmatrix} \quad \text{and} \quad \mathbf{\Sigma}(x,y) = \begin{pmatrix}\sigma(x) & 0\\ \sigma(y) & 0\end{pmatrix}. \tag{1.2.10}$$

We remark that for $\mathbf{X}_t^x$ and $\mathbf{Y}_t^y$ two solutions to (1.2.1) respectively starting from x and y, the joint law of $(\mathbf{X}_t^x, \mathbf{Y}_t^y)$ solves (1.2.8). Put differently, (1.2.8) is the Fokker–Planck equation of (1.2.9). Likewise, the joint law of $(\mathbf{X}_t^x, \mathbf{X}_t^y)$ solves (1.2.8). By uniqueness, these two laws are therefore equal to one another. Formally taking the limit $y \to x$, we obtain $\mathbf{X}_t^x = \mathbf{Y}_t^x$ and thus pathwise uniqueness. The argument can be made rigorous for regular coefficients (and will indeed be made rigorous for more general coefficients in Section 2.2 below).

Alternately, one may again use an adjoint viewpoint. Instead of (1.2.8), one considers the *backward Kolmogorov equation*

$$\frac{\partial f}{\partial t} - b_i(x)\frac{\partial f}{\partial x_i} - b_i(y)\frac{\partial f}{\partial y_i} - \frac{1}{2}\sigma_{ik}(x)\sigma_{jk}(x)\frac{\partial^2 f}{\partial x_i \partial x_j} - \sigma_{ik}(x)\sigma_{jk}(y)\frac{\partial^2 f}{\partial x_i \partial y_j} - \frac{1}{2}\sigma_{ik}(y)\sigma_{jk}(y)\frac{\partial^2 f}{\partial y_i \partial y_j} = 0. \quad (1.2.11)$$

Given (1.2.3), $f(t,x,y) = \mathbb{E}(|\mathbf{X}_t^x - \mathbf{Y}_t^y|^2)$ solves equation (1.2.11) with initial condition $\psi_0(x,y) = |x-y|^2$, when $\mathbf{X}_t^x$ and $\mathbf{Y}_t^y$ are again two solutions to (1.2.1), starting from x and y respectively. A simple computation then shows that, for some sufficiently large constant C, $\overline{\psi}(t,x,y) = e^{Ct}|x-y|^2$ is a super-solution to (1.2.11) whence

$$\mathbb{E}(|\mathbf{X}_t^x - \mathbf{Y}_t^y|^2)) \leq e^{Ct}|x-y|^2$$

for all times. Taking $x = y$ yields pathwise uniqueness.

To conclude our comments on the regular setting, we would like to mention the following three, entangled points.

Firstly, as briefly mentioned above, using the solution to (1.2.1) for $\mathbf{b}$ and σ “only” Lipschitz continuous, which is the setting of the Itô theory, does not immediately allow to construct a solution to (1), even if this is evident *formally*. Indeed, assuming that u_0 is itself Lipschitz continuous, (1.2.3) only shows that u is Lipschitz continuous in the space variable, a property that does not allow to claim it is a classical solution to (1). However, arguing on the partial differential equation itself, and assuming that $\mathbf{b}$ and σ are more regular, that is, typically $W^{2,\infty}$, we do obtain directly from (1.2.3) a solution to (1), in $W_{x,t}^{2,1,\infty}$ when $u_0 \in W^{2,\infty}$. The space $W_{x,t}^{k,p,\infty}$ evidently denotes the space of functions that are $W^{k,\infty}$ in the space variable x and $W^{p,\infty}$ in the time variable t. Uniqueness of this solution, on the other hand, is an easy matter given the linearity of the equation, the regularity of the coefficients and that of the solution manipulated. For an only *continuous* (not necessarily Lipschitz continuous) initial condition u_0, and $\mathbf{b}$ and σ Lipschitz continuous as in the Itô theory, it can also be proven that a solution to (1) uniquely exists. The result is originally due to Oleinik. We note that, in that setting, the adequate notion of solution is that of *viscosity solutions*, and the solution is continuous for all times. We will briefly return to all this in Section 1.3.1. Even for such Lipschitz continuous regular coefficients $\mathbf{b}$ and σ, it is not part of the classical theory to address less regular (meaning, less than continuous) initial conditions such as $u_0 \in L^1$ for instance. One readily understands the difficulty upon considering the key formula (1.2.3) which makes the connection between the various viewpoints: the meaning of $u_0(\mathbf{X}_t)$ for an initial condition u_0 that is *not* continuous is unclear (in the absence of any specific property of $\mathbf{X}_t$).

Secondly, and in line with the previous comment, we wish to emphasize that in the case of the parabolic-type equation (1), assuming Lipschitz continuous coefficients, along with the regularity (say, continuity) of the initial condition u_0, allows to prove existence and uniqueness of the solution, *but* does not allow to control the flow, and thus to prove *continuity* with respect to the initial condition, that is, an estimate of the form $\|u(t,\cdot)\| \le C\|u_0\|$ in the suitable functional space L^p, $1 \le p < +\infty$. An example of such an estimate is (2.1.3) below. This is in sharp contrast with the case of a linear transport equation where the Lipschitz regularity allows to also control the flow, as recalled in (1.1.7) and our discussion on page 2 above. The reason for this striking difference is of course the presence of the diffusion (or, noise) term, which affects the flow. For transport equations, the purpose of considering Sobolev regular coefficients is therefore to investigate whether that control of the flow, equivalent to the continuity of the map $u_0 \to u$, can be established under weaker assumptions. For parabolic-type equations, *even* for Lipschitz continuous coefficients, we do not have that property. The study we have performed in our previous works and we continue here definitely follows a different perspective from the classical perspective. For instance, in our previous work [65], we have established some results of existence, uniqueness, and continuity with respect to the initial condition, for (1) when $\mathbf{b}$ and σ are Lipschitz continuous, for an initial condition u_0 that can be Lipschitz continuous, or H^1, provided some additional assumptions are satisfied. Some partial results were also obtained therein for $u_0 \in L^2$. These comments also relate to our next item.

A third fact worth emphasizing is that, in echo to our comment (vii) above on ordinary differential equations, the question of uniqueness naturally brings up the question of *linear tangent flows*. Taking $y = x - h\xi$, with $|\xi| = 1$ and h a small parameter in (1.2.7), and next letting h go to zero, formally yields

$$\mathbb{E}\left(\left|\frac{\partial \mathbf{X}_t^x}{\partial x}\right|^2\right) \le e^{C_0\,t}. \tag{1.2.12}$$

This gives an evaluation of the tangent flow, in the direction ξ. This corresponds to the equation

$$\begin{cases} d\mathbf{Y}_t = \mathbf{b}'(\mathbf{X}_t)\mathbf{Y}_t\,dt + \sigma'(\mathbf{X}_t)\,\mathbf{Y}_t\,d\mathbf{W}_t, \\ \mathbf{Y}_{t=0} = \xi, \end{cases} \tag{1.2.13}$$

where we have formally denoted by $\mathbf{b}'$ and σ' the derivatives of $\mathbf{b}$ and σ. The latter equation does not make rigorous sense, unless $\mathbf{b}$ and σ are sufficiently regular, say at least C^1. In the absence of regularity, the meaning of a term such as $\sigma'(\mathbf{X}_t)$ is indeed unclear, since $\mathbf{X}_t$ may *a priori* weight sets of zero Lebesgue measure. It will be a corollary of the results on the well-posedness of the parabolic equation (1) to also give sense to this linear tangent flow, as we did for the case of ordinary differential equation. We notice that works of H. Kunita have addressed tangent flows for stochastic differential equations in the classical Itô theory, see [61, 62].

The martingale formulation. Going beyond the previously mentioned *regular* setting requires to understand the notion of solution to (1.2.1) in a broader sense. This is the purpose of the *martingale problem*, introduced by Stroock and Varadhan. Recall that, in the regular setting, when $\mathbf{X}_t$ is a solution to (1.2.1) and φ is a smooth function, the Itô differentiation formula shows that

$$\varphi(t, \mathbf{X}_t) - \varphi(0, x) - \int_0^t (\partial_s \varphi + \mathbf{b} \,.\, \nabla\varphi + \frac{1}{2}\sigma\sigma^t D^2 \varphi)(s, \mathbf{X}_s)\, ds \tag{1.2.14}$$

is a stochastic integral, thus a martingale. The idea is then to choose the latter property as a characterization of a new class of solutions to equation (1.2.1). The focus is not any longer on the particular trajectories of the possible solution $\mathbf{X}_t$ but on its law. In short, one chooses as the probability space Ω the space of continuous trajectories and the problem of finding a solution to the stochastic differential equation (1.2.1) reduces to finding a convenient probability $\mathbb{P}$ on this space Ω, so that (1.2.14) is a martingale, for all (say, C^2) test-function φ. Intuitively, searching for the law of a process solution to a stochastic differential equation is much closer to searching for the solution to a partial differential equation than making a connection between the two equations using a representation formula of the type (1.2.3) (or alternately, (1.1.3)). This is exemplified by our discussion above on the connection between (1.2.1) and (1) via formula (1.2.3).

Solving the martingale problem may be shown to be equivalent to solving (1.2.1) in the weak sense. This is not elementary. Existence of a solution to the martingale problem is equivalent to the existence of a weak solution to the stochastic differential equation (1.2.1), while uniqueness of that solution is equivalent to uniqueness in law for (1.2.1). The prototypical result for the martingale problem (see, e.g., [84, p. 170, Volume 2] for a precise statement) states that (again on a bounded domain) for $\mathbf{b}$ and σ Lipschitz continuous, *or* for $\mathbf{b}$ bounded and $\frac{1}{2}\sigma(x)\sigma^t(x)$ continuous, positive definite for all x, there exists a unique solution to the martingale problem associated to equation (1.2.1).

An important remark, particularly in the context of these lecture notes, is that, actually, the uniqueness of a solution to the martingale problem (which is the key point in the well-posedness of that problem) is in fact obtained using an argument based on the existence of a solution with appropriate regularity to the parabolic partial differential equation. The well-posedness for the PDE is the celebrated Agmon–Douglis–Nirenberg theory for parabolic equations [2, 3]. We will return to this connection in Section 2.2 (page 56).

1.3 Parabolic equations

We also would like to recall here various results on the classical, and less classical theory of linear parabolic equations. We begin with the regular setting, briefly discuss the viewpoint of fundamental solutions and then recall our previous works on the irregular setting.

1.3.1 Parabolic equations with regular data, and related issues

The systematic study of (1) in the case of regular coefficients **a** and **b**, without necessarily assuming that **a** is positive definite, dates back to the works by O. Oleinik.

For coefficients $\mathbf{a} \in W^{2,\infty}$ (not necessarily positive definite), $\mathbf{b} \in W^{2,\infty}$, and initial conditions $u_0 \in W^{2,\infty}$, *a priori* estimates may be established and consequently existence and uniqueness of the solution u to (1) holds in the class $W^{2,1,\infty}_{x,t}$ (where the first exponent refers to the differentiability in space, and the second to the differentiability in time). This allows to define all the terms of (1) in the classical sense, which was of course the purpose of considering that particular functional setting. A variant of the result holds for all degrees of differentiability $k \geq 2$ instead of 2. Also, Oleinik established Lipschitz-type estimates for the solution when $\mathbf{a} = \sigma\sigma^t$, σ and **b** are Lipschitz continuous. This case corresponds to the classical probability setting.

In the case when **a** is positive definite, the main result, which we briefly mentioned in the previous section, is the Agmon–Douglis–Nirenberg result. The setting can be either that of Schauder spaces, or that of Sobolev spaces, and we only mention here the latter setting, given our focus in these lecture notes. Under the additional condition of positive definiteness, the Agmon–Douglis–Nirenberg result generalizes the results by Oleinik to L^p integrable functions instead of L^∞ in the following sense. If **a** is continuous and **b** is bounded, if the initial condition u_0 is sufficiently regular, if there is a right-hand side of (1) in L^p, for some $1 \leq p < +\infty$, then we have existence and uniqueness of the solution to (1) in $W^{2,1,p}_{x,t}$. This result, as announced above, is the key result to solve the martingale problem.

Weakening the above assumptions on **a**, **b** and u_0 immediately creates substantial difficulties. In particular, even in the case of a definitive positive second-order term, no systematic study explores the cases when the initial condition u_0 lacks regularity. One specific setting we have briefly mentioned is the case when the initial condition is considered only continuous. The theory of viscosity solutions, [35], is then the appropriate tool. Either in the case when **a** and **b** are Lipschitz continuous, or in the case when **a** and **b** are only continuous but **a** is positive definite, there is existence and uniqueness of a continuous solution to (1).

Further reducing the regularity of the initial condition essentially consists in considering $u_0 \in L^1$. This exactly brings up the question of whether the viewpoint of *fundamental solutions* is, or not, the suitable viewpoint. In the case when **a**, **b**, u_0 are regular and the second-order term is uniformly elliptic, it is indeed standard in the vein of distribution theory, to write the solution to (1) as

$$u(t,x) = \int p(t,x,y)u_0(y)\,dy, \tag{1.3.1}$$

where we write somewhat vaguely the integral and where p of course denotes the fundamental solution to (1). For non-regular u_0, even though the integral above can be given a sense (see [10] and other works by D. G. Aronson), it is unclear in which sense (1.3.1) solves (1) and especially how (1.3.1) agrees with u_0 at initial time. In any event, there is a good intuitive reason why pursuing our analysis within the context of fundamental solutions is likely to not be an appropriate strategy. When both **a** and **b** vanish – a regular setting indeed! –, in which case (1) reduces to $\partial_t u = 0$, the fundamental solution writes $p(t, x, y) = \delta_0(x - y)$. The latter function is evidently singular and therefore obtaining estimates on the fundamental solution is hopeless. One alternate option would be to consider estimates on the fundamental solution *seen as a kernel*, which amounts to actually solving the equation. Yet another option would be to manipulate a *parametrix*, that is, an approximation of the fundamental solution. It is actually unclear what can be expected from these perspectives. We will not proceed in those directions.

1.3.2 Some previous works on parabolic equations with irregular data and related issues

We now recall some essential results of our previous works regarding the solution to equations of the type (1) with non-regular (that is, typically Sobolev regular) coefficients and initial conditions. Since the settings we considered will be generalized in the present contribution, and the arguments we used are simple version of those we will use, our discussion is kept brief.

Parabolic-type equations. Of course, it is immediate to observe that *any constant nonnegative* second-order operator does not modify the standard proof performed in the case of linear transport, which we have recalled above. All the results of the linear transport equation therefore hold for (1) in that case. We will improve those results later in the text. If we additionally assume some ellipticity of this operator, a more general field **b** can be accommodated. Along these lines, the first extension of the results of [41] on transport equations to the setting of parabolic-type equations is, to our knowledge, our previous work [64]. We have studied there the parabolic equation

$$\partial_t u - b_i \partial_i u - \frac{1}{2}\Delta u = 0. \tag{1.3.2}$$

The presence of the regularizing second-order operator $-\Delta$ of course yields a better regularity on the solution u than in the pure transport case. From the formal a priori estimate

$$\frac{d}{dt}\int \frac{u^2}{2} + \int (\operatorname{div}\mathbf{b})\frac{u^2}{2} + \frac{1}{2}\int |\nabla u|^2 = 0, \tag{1.3.3}$$

similar to (1.1.8) for the transport equation, it is expected that $u \in L^2([0, T], H^1)$. This expected additional H^1 regularity is useful for the regularization step. Instead of expressing the commutator $[\rho_\varepsilon, \mathbf{b}\,.\,\nabla](u)$ using (1.1.11), we directly express it as

$$\int (\mathbf{b}(y) - \mathbf{b}(x))\,.\,\nabla u(y)\rho_\varepsilon(x - y)\,dy. \tag{1.3.4}$$

We see that an L^2 regularity on **b** is now sufficient to proceed in this regularization step and show that (1.3.4) vanishes in L^1_x as $\varepsilon \to 0$, without the need for the weak differentiability of **b**. Once the regularization is performed, the a priori estimate (1.3.3) makes rigorous sense, up to small error terms vanishing with the regularization. Controlling div **b** then typically allows to conclude that uniqueness holds, as for the transport equation. This discussion explains why and how, mimicking the arguments of [41], we have obtained in [64, Section 5] existence and uniqueness of the solution to (1.3.2) in $L^\infty([0,T], L^1 \cap L^\infty) \cap L^2([0,T], H^1)$ when $\mathbf{b} \in W^{1,1} + L^2$ and $\operatorname{div} \mathbf{b} \in L^\infty$.

A specific case of a varying matrix **a** in (1) was first examined in [65]. The outline of the proof in that case is essentially similar to that for the constant case above, but of course technicalities arise. A prototypical result obtained in [65] (and recalled in Chapter 3 below) is the existence and uniqueness for the solution to

$$\begin{cases} \partial_t u + \partial_i(u b_i) - \dfrac{1}{2}\,\partial_i(\sigma_{ik}\sigma_{jk}\partial_j\, u) = 0, \\ u|_{t=0} = u_0, \end{cases} \tag{1.3.5}$$

in the functional space

$$\{u \in L^\infty([0,T], L^\infty) : \sigma^t \nabla u \in L^2([0,T], L^2)\}$$

under the assumptions $\mathbf{b} \in W^{1,1}$, $\operatorname{div} \mathbf{b} \in L^\infty$, $\sigma \in H^1$, $u_0 \in L^\infty$. In the particular case, when the matrix $\sigma\sigma^t$ is positive definite, existence and uniqueness can be established in the space

$$\{u \in L^\infty([0,T], L^\infty) : u \in L^2([0,T], H^1)\}$$

under the assumptions $\mathbf{b} \in L^2$, $\operatorname{div} \mathbf{b} \in L^\infty$, $\sigma \in L^\infty$, $u_0 \in L^\infty$.

Both in [64] and [65], various extensions of the above prototypical result were obtained. These extensions essentially concern, on the one hand, extensions to equations of the type (1) that are not in a *divergence form* like (1.3.5) and, on the other hand, the consequences of the results on parabolic equations on the theory of stochastic differential equations. More extensions will be examined in the present contribution.

The former extensions essentially consist in accounting for the transformation of **b** into $\mathbf{b}^\sigma$ defined by $b_i^\sigma = b_i - \frac{1}{2}\partial_j(\sigma_{ik}\sigma_{jk})$ for all $1 \le i \le N$, which allows to accommodate part of the non-divergence form into the transport term. We will reiterate similar arguments (and significantly extend them) in Section 4.1. Another useful observation is that the transformation of **b** into $\mathbf{b} + \sigma\theta$ for some $\theta \in L^\infty$ does not affect most of our arguments (a fact that is indeed connected to the *Girsanov transform* in stochastic analysis, and to which we will return below in the comments after Proposition 2 in Chapter 3).

The point to stress is however that in our previous works [64, 65], we only very briefly consider an unbounded initial condition u_0, or a lower regularity of the coefficients than that mentioned above, or both settings together. In a small portion of our work [65], we have studied the case of Lipschitz continuous coefficients σ and **b** for (1) with $\mathbf{a} = \frac{1}{2}\sigma\sigma^t$, with an initial condition that is H^1, or Lipschitz continuous, or L^2. In particular, only partial results were obtained in the latter setting, and we refer the reader to [65] for more details. Actually, works are rare, in the literature, that consider such irregular settings. The only contributions known to us are the following ones. In [70, Appendix E], renormalized solutions for a specific parabolic equation with an unbounded initial condition has been considered, in particular using notions and techniques introduced in the context of Fokker–Planck–Boltzmann equations in [40]. In line with the questions we examined in [65, Section 8], existence and uniqueness of solutions for equations such as (1) were considered by A. Figalli in [46] for two main settings: (a) uniformly positive definite, Lipschitz continuous in time, matrices $\sigma\sigma^t$ and L^∞ vector fields, or (b) space-independent matrices $\sigma\sigma^t$ and BV vector fields. Besides these contributions, two groups of authors have also considered similar issues. As we will point out below, some of the techniques we develop here for renormalized solutions are of course connected to some arguments by D. Blanchard, F. Murat and their collaborators in [15, 16] and other works. The series of works [17, 19, 20] by V. I Bogachev and coworkers present a different, interesting perspective.

To conclude this paragraph, we wish to cite some works related to [65], which followed up on it, or which present some alternate approaches. An immediate extension of our results on the parabolic equation to the case of a vector field $\mathbf{b}$ with bounded variation is performed in [77], combining the techniques of [65] and [5] and following our suggestion [65, Remark 12, p. 1299]. Of interest are also the works by F. Flandoli and his collaborators (see [11, 12, 45, 48] and the nice survey [47]) that study the regularizing effect due to an extra noise term in the transport equation, a phenomenon somewhat related to the parabolic regularization mechanism we examine in the present contribution.

The second category of extensions we wish to mention concerns the theory of stochastic differential equations.

Consequences on the theory of stochastic differential equations. In [64], we considered (1.2.1) for $\sigma \equiv \mathrm{Id}$, $\mathbf{b} \in W^{1,1}$, $\operatorname{div}\mathbf{b} \in L^\infty$ (and, when the domain is unbounded, the growth condition $\frac{\mathbf{b}}{1+|x|} \in L^1 + L^\infty$). In that particular setting, we were able to define a (unique) generalized *flow* of solutions to (1.2.1), strong in the probability sense, which amounts to solving the SDE for almost all initial conditions. We cannot emphasize enough that our approach does not give any information on the existence or the uniqueness of a solution to (1.2.1) for a specific initial condition. It only holds for the *flow* or, equivalently, when the equation is considered for almost all initial conditions. Our proof was performed specifically using that σ is a constant matrix, and connecting (1.2.1) with the ordinary differential equations

$$d(\mathbf{X}_t - \sigma\mathbf{W}_t) = \mathbf{b}((\mathbf{X}_t - \sigma\mathbf{W}_t) + \sigma\mathbf{W}_t)\, dt$$

parameterized by the Brownian trajectories (we refer the reader to [64] for a more detailed discussion). Then, extending to this Sobolev regular setting the classical connection between uniqueness in law for (1.2.1) and uniqueness for (1.2.4), we were able to use the unique solvability of the Fokker–Planck equation (1.2.4) we had previously established to show that the strong solution flows to (1.2.1) all share the same law, which is the unique solution to (1.2.4).

In addition, we have proved (as recalled above in our outline of the proof) that we may also uniquely solve the Fokker–Planck equation (1.2.4) for $\mathbf{b} \in L^2 + W^{1,1}$ instead of $\mathbf{b} \in W^{1,1}$. However, it is not evident to translate this property in terms of the theory of stochastic differential equations, precisely because extending our notion of generalized flow to the case $\mathbf{b} \in L^2$ is unclear.

Intuitively, the fact that the case of a constant σ in (1) only brings a small additional difficulty with respect to the linear transport equation can be understood as follows. As said above, existence is obtained using the underlying *ordinary* differential equation. On the other hand, when it comes to uniqueness issues, the scheme of the proof is to formally subtract $d\mathbf{Y}_t = \mathbf{b}(\mathbf{Y}_t)\,dt + \sigma\,d\mathbf{W}_t$ to $d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)\,dt + \sigma\,d\mathbf{W}_t$, which again leaves us with a differential equation $d(\mathbf{X}_t - \mathbf{Y}_t) = (\mathbf{b}(\mathbf{X}_t) - \mathbf{b}(\mathbf{Y}_t))\,dt$ close to the deterministic setting since the Brownian motion term has cancelled out.

All previous arguments unfortunately collapse when σ varies. The dispersion matrix $\sigma(\mathbf{X}_t)$ then intimately modifies the trajectories of the associated deterministic dynamics and the questions of the deformation of the Lebesgue measure by the flow has to be entirely revisited. To our knowledge, this question of the *deformation of the Lebesgue measure under a stochastic flow* has never been investigated.

As mentioned in [65, Section 3], the following questions remained unsolved in [64]:
- in the case of a constant dispersion matrix σ:
  (i) the pathwise uniqueness of (generalized) strong solutions to (1.2.1) when $\mathbf{b}$ has Sobolev regularity,
  (ii) the existence and uniqueness-in-law of (generalized) weak solutions to (1.2.1) when $\mathbf{b}$ is, say, L^2,
- in the case of a varying dispersion matrix σ:
  (iii) all questions regarding, in the various senses, existence and uniqueness of the solution to (1.2.1).

The present contribution intends to clarify such outstanding questions, in particular in Sections 2.2 and 4.2.

We would like to finally emphasize that, because we have (for simplicity of exposition) restricted our study to parabolic equations posed on a bounded domain with periodic boundary conditions, one may transpose the results to the probability theoretic setting in two slightly different ways. Either, one may apply our arguments and results to stochastic differential equations posed on random processes valued in the torus $[0,1]^d$. In that case, some simplifications happen, but, on the other hand, technicalities arise because of periodic boundary conditions, the norm on the torus being slightly different from the Euclidean norm on the space $\mathbb{R}^d$. Or, one may
- firstly extend our results of the partial differential equation setting to unbounded domains; this might be performed simply combining the results of the present contribution and the techniques used in [65] to estimate tails of all integrals on $\mathbb{R}^d$ (assuming for this purpose that $\frac{\mathbf{b}}{1+|x|} \in L^1 + L^\infty(\mathbb{R}^d)$ and $\frac{\sigma}{1+|x|} \in L^2 + L^\infty(\mathbb{R}^d)$), and
- secondly transpose these "generalized" results to stochastic differential equations posed on real valued random processes.

Either way, the heart of the arguments for the connection rests on the same crucial ingredients, which we will focus upon in our Sections 2.2 and 4.2 describing the application of the results to the probability theoretic setting.

To conclude this paragraph mentioning a different, although related, perspective on the topic, we would like to cite the recent work [33]. There, the results on the stochastic differential equation are examined *on their own*, in a somewhat decoupled manner from the results on the underlying parabolic equation. The spirit of the approach is similar to what has been completed for proving existence and uniqueness of the solution to ordinary differential equations in the works of G. Crippa and C. De Lellis, directly estimating distances between flows for the differential equation (see [37]).

Relation to elliptic problems. All our arguments in these notes are focused on the parabolic equation (1) and its variants. It is worth emphasizing that these arguments apply mutatis mutandis to the static variant of this equation, namely the so-called *advection-diffusion equation*

$$-b_i \partial_i u - a_{ij} \partial^2_{ij} u = f, \tag{1.3.6}$$

for some given right-hand side f. We now briefly explain why. This will also explain why in our discussions below we will sometime consider the static variant (1.3.6) to illustrate the optimality, the sharpness, or, from another perspective, the limitations of our arguments. To start with, let us first recall, for convenience of the reader, the theory of existence and uniqueness for the equation (1.3.6) in the classical setting, that is, for regular coefficients **a**, **b** and data f. For this purpose, we will argue, without so much loss of generality on the equation

$$-\mathbf{b}\,.\,\nabla u - \Delta u = f, \tag{1.3.7}$$

which, again for simplicity, we consider on a bounded, regular domain *with, specifically here, homogeneous Dirichlet conditions*. The simplest setting is that when existence and uniqueness are simultaneously obtained using the coerciveness of the differential operator $Lu := -\mathbf{b}\,.\,\nabla u - \Delta\, u$. We notice that

$$(Lu, u) = \int |\nabla u|^2 - \int \mathbf{b}\,.\,\nabla u u,$$

or alternately that

$$(Lu, u) = \int |\nabla u|^2 + \frac{1}{2} \int (\operatorname{div} \mathbf{b}) u^2.$$

Using the Hölder inequality and the Sobolev inequality $\|u\|_{L^{2d/(d-2)}} \le C \|\nabla u\|_{L^2}$ (see the next section of these notes for the details on these inequalities and the related estimations), this respectively shows that the operator L is coercive on H^1_0 of the domain in either of the following two cases: (i) when **b** itself is sufficiently small in L^d norm, or (ii) when $\operatorname{div} \mathbf{b}$ is sufficiently small in $L^{\frac{d}{2}}$. In any event, coerciveness is obtained. Existence and uniqueness for (1.3.7) then follow from a direct application of the Lax–Milgram Lemma. We have implicitly proceeded in an ambient dimension $d > 2$, but the cases of dimensions 1 and 2 can be addressed likewise. When using periodic, instead of Dirichlet homogeneous, boundary conditions, the argument needs to be slightly adapted since now the Sobolev inequality reads

$$\|u\|_{L^{2d/(d-2)}} \leq C(\|\nabla u\|_{L^2} + \|u\|_{L^2}).$$

Then the addition of a zeroth-order term αu to the left-hand side of equation (1.3.7), with a sufficiently large constant α, is necessary to obtain coerciveness. Also, more elaborate arguments can be used (many of them will be seen later in these notes) in order to extend the assumptions needed on $\mathbf{b}$ in (1.3.7); see for instance our assumptions (2.1.2) in the statement of Theorem 1 below. Equation (1.3.6) is addressed similarly, with more difficulties because of the possible degeneracy of the second-order term.

Again in the classical setting, the more general case when L is not coercive is usually addressed using the Fredholm alternative (see, e.g., [52, Chapter 8]). One writes (1.3.7) in the form

$$L_\alpha u := \alpha u - \mathbf{b}.\nabla u - \Delta u = \alpha u + f,$$

for some constant parameter α presumably sufficiently large. That coefficient is then adjusted so that L_α is coercive (for similar reasons as L in the simple case above). Then L_α^{-1} exists and, because of the ellipticity and the Rellich Theorem, is a compact operator. The equation writes $(\mathrm{Id} - \alpha L_\alpha^{-1})u = L_\alpha^{-1} f$ and is solved using the Fredholm alternative. The fact that $\mathrm{Ker}(\mathrm{Id} - \alpha L_\alpha^{-1})$ is trivial is typically obtained using the maximum principle on the operator L. Interestingly, the existence part does not actually require compactness: a modified argument, based on monotonicity (see the details, and related considerations, in [67–69]), allows to extend the theory to unbounded domains.

Assume now we consider coefficients $\mathbf{a}$ and $\mathbf{b}$ that are not sufficiently regular for the above classical arguments to make sense. Then the techniques we develop throughout these notes on the parabolic equation (1) allow to reinstate coerciveness on the operator, possibly adding a zeroth-order term $\alpha\, u$, for a sufficiently large constant α. Put differently, depending on the context, we are able to prove that the operator L, or L_α, is coercive in the suitable topology. All this depends on the properties of the second-order operator. It is simple when that operator is elliptic, less simple, or even unclear, otherwise – we will study hypo-elliptic or sub-elliptic situations –, and will be evident later in these notes. In sharp contrast, when (1.3.6) is considered without any additional term $\alpha\, u$ and without the freedom of taking $\mathbf{b}$ small in a certain sense, the situation is less clear. In particular, uniqueness may then rely on an application of the maximum principle, which in turn requires some minimal regularity of the coefficients which we do not necessarily wish to assume. Such a situation is not addressed at all in these notes.

Additionally, the above remarks show that we will have the opportunity, in order to better understand the role played by the various assumptions we put on the coefficients, to use equation (1.3.6) as a test-bed. This will actually be how we will mainly use that equation throughout these notes.

To conclude this section, we mention that a setting very close to the elliptic variant (1.3.6) is the case where the *time-dependent* equation (1) is considered with a *sufficiently regular initial condition* u_0. Then, formally denoting (1) by

$$\partial_t u + Lu = 0,$$

and assuming that the coefficients **a** and **b** in L are *time-independent*, we observe, differentiating this equation with respect to time, that the function $\partial_t u$ formally solves the same equation. Therefore, $\partial_t u$, thus Lu, belongs to the same functional spaces as the solution u itself. Put differently, $Lu = f$, where f belongs to all the appropriate functional spaces we will derive in the sequel. Formally, at each time t, the problem is thus amenable to the techniques of advection-diffusion equations, and our techniques and results will again carry over to that case. We will return to this below, in Section 4.3, with a general approach (valid also for time-dependent coefficients).

1.3.3 Brief outline of our techniques and results

Having overviewed the existing works on related issues, we are now in a position to briefly outline our techniques and results in a slightly more detailed manner than what we did in the introduction. To this end, we note that, as repeatedly mentioned above, our proofs of existence and uniqueness proceed through two main steps: *regularization* and *a priori estimates*. These two steps can be understood by considering the following two quantities:

$$\int \mathbf{b} \,.\, \nabla u \, u \tag{1.3.8}$$

and

$$\int \left(\int \rho_\varepsilon(x-y)(\mathbf{b}(y)-\mathbf{b}(x)) \,.\, \nabla u(y)\, dy \right) u(x)\, dx. \tag{1.3.9}$$

The former quantity appears in the formal L^2 a priori estimate established upon multiplying the equation considered by u and integrating over the domain, while (1.3.9) is the typical remainder obtained when making this formal estimate rigorous upon using the commutator $[\rho_\varepsilon, \mathbf{b} \,.\, \nabla](u)$. In the now classical study of the linear transport equation, as initiated in [41], the two terms are estimated using integration by parts

$$\int \mathbf{b} \,.\, \nabla u\, u = -\int \underbrace{\operatorname{div} \mathbf{b}}_{\in L^\infty} \,.\, \underbrace{\frac{u^2}{2}}_{\in L^1} \tag{1.3.10}$$

and

$$\begin{aligned} &\int \left(\int \rho_\varepsilon(x-y)(\mathbf{b}(y)-\mathbf{b}(x)) \,.\, \nabla u(y)\, dy \right) u(x)\, dx \\ &\quad = \cdots - \int\int \underbrace{\nabla \rho_\varepsilon(x-y) \,.\, (\mathbf{b}(y)-\mathbf{b}(x))}_{\text{use } \nabla \mathbf{b} \in L^1} \underbrace{u(y)}_{\in L^\infty}\, dy\, \underbrace{u(x)}_{\in L^\infty}\, dx \end{aligned} \tag{1.3.11}$$

when, for simplicity of the exposition, the initial condition is assumed bounded. The above two-fold estimation allows to intuitively understand the classical couple of assumptions ($\operatorname{div} \mathbf{b} \in L^\infty$, $\mathbf{b} \in W^{1,1}$). We have already mentioned this.

We now move on and consider the case when a positive definite second-order operator is added to the transport, namely $-\Delta u$ for simplicity. The brief outline of the results established in our work [64] is that we may now leave (1.3.9) as is, write that

$$\int\int \underbrace{\rho_\varepsilon(x-y)(\mathbf{b}(y)-\mathbf{b}(x))}_{\text{use } \mathbf{b}\in L^2} \cdot \underbrace{\nabla u(y)}_{\in L^2}\, dy\, \underbrace{u(x)}_{\in L^\infty}\, dx, \tag{1.3.12}$$

and also conclude that existence and uniqueness hold *without* assuming $\mathbf{b} \in W^{1,1}$. We now observe (and this observation was not included in [64]) that, also using the better (namely H^1) regularity of the solution u expected from the presence of $-\Delta u$ (or a similar term), we can equally well improve our treatment of the term (1.3.8) and write, again without resorting to an integration by parts,

$$\int \underbrace{\mathbf{b}}_{\in L^d} \cdot \underbrace{\nabla u}_{\in L^2}\ \underbrace{u}_{\in L^{2d/(d-2)}}, \tag{1.3.13}$$

using the classical Hölder estimate. Notice that for simplicity we assume that the ambient dimension d is strictly larger than 2; the two-dimensional case requires specific adaptations. Estimate (1.3.13) holds in the absence of any control of $\operatorname{div} \mathbf{b}$. The consideration of (1.3.12) and (1.3.13) together shows that, in such a setting, no assumption on the weak derivatives of $\mathbf{b}$ seems necessary to conclude. We will show in the present contribution that this is indeed the case. We will also see variants and extensions of estimate (1.3.13), using Lorentz interpolation spaces. Notice that if we are ready to control $\operatorname{div} \mathbf{b}$, then the parabolic situation is also advantageous, since we may then improve (1.3.10) and the corresponding assumption $\operatorname{div} \mathbf{b} \in L^\infty$ into

$$\int \mathbf{b}.\nabla u\, u = -\int \underbrace{\operatorname{div} \mathbf{b}}_{\in L^{d/2}}\ \underbrace{\frac{u^2}{2}}_{\in L^{d/(d-2)}} \tag{1.3.14}$$

and some similar variants and extensions.

It is now obvious from (1.3.12)–(1.3.13) that our main two ingredients for the parabolic-like case are
(i) the boundedness of u and
(ii) the extra regularity provided by the second-order operator.
Our next step is to investigate how we may further generalize the above setting in both regards.

The first direction to generalize the results concerns the initial condition. Our discussion above has been conducted under the simplifying assumption $u_0 \in L^\infty$. But of course we would like to address more general initial conditions, say $u_0 \in L^p$. For such an initial condition, we need to adapt the above arguments, and we will do so using a renormalization technique as initially introduced in [41]. Intuitively, we essentially manipulate a truncation $\Phi(u)$ instead of u itself. The exponents involved in the various estimates such as (1.3.12) or (1.3.13) are affected by this change, but the bottom line for the proof remains.

Our second direction follows the observation that the *extra* regularization provided by the second-order term and which allows to address more general cases of **b** is not restricted to the case of $-\Delta u$. The Sobolev inequality

$$\|u\|_{L^{2d/(d-2)}} \leq C\|\nabla u\|_{L^2} \tag{1.3.15}$$

is a key ingredient to establish (1.3.13) or (1.3.14), but extensions of this inequality hold in a variety of settings. The immediate extension is the case when a *positive definite* second-order operator in divergence form $-\mathrm{div}(\sigma\sigma^t\nabla u)$ is added to the transport equation. This allows to readily extend the result on the Laplacian to the case of operators with varying coefficients. This was already the purpose of [65], but the generality considered therein was weaker than in the present contribution. Likewise, if the second-order operator is definite in *some* directions only, then we may obtain mixed results, using *"transport-like"* assumptions on **b** along the directions not regularized, and more general assumptions along the other directions. Note that, as soon as a second-order term with varying coefficients is considered, a second regularization step like in (1.3.9) but this time specifically on this second-order term, must be performed. And this again requires various, sometimes elaborate, assumptions, which we omit in this brief outline.

Much more generally and in the same vein, we also notice that the gain in (1.3.13) is related to the fact that the exponent $\frac{2d}{d-2}$ is strictly larger than 2, thus allowing for considering unbounded fields **b**. This suggests exploring which type of second-order operator, of the form $-\mathrm{div}(\sigma\sigma^t\nabla u)$ and not necessarily positive definite, gives rise to estimates of the type

$$\|u\|_{L^p} \leq C\|\sigma^t\nabla u\|_{L^2} \tag{1.3.16}$$

for some $p > 2$, and how we may use this to consider more general fields **b**. Obviously, (1.3.16) is, in a sense, an extension of (1.3.15). Not unexpectedly, the subject is closely related to considerations around the *hypoellipticity* of operators. We will explore these issues in Section 3.5.

To conclude this chapter, we mention that not only numerous variants of the above results along different directions will be investigated but we will also systematically consider how the results on the parabolic equation affect the theory of stochastic differential equations with irregular coefficients.

2 Operators with constant second-order term

We now get to the heart of the matter, and study (1). We devote this chapter to the case when the (we recall, symmetric) second-order operator in (1) has constant coefficients a_{ij}. For operators that have no sign, nothing general can be expected. The most general setting we may therefore consider is that of constant operators that are *nonnegative*, in the sense of symmetric operators. We have already briefly mentioned that the presence of such operators does not modify the proofs performed on the linear transport equation. We thus have *at least* the results of [41] and their extensions, already mentioned above, holding true for equation (1) in those conditions. We also have recalled in the previous section the results obtained in [64] when the operator is elliptic, that is essentially, a Laplacian. We now wish to establish in a sense the *most general* result for that setting. We note, slightly anticipating on the contents of the next chapters, that when its coefficients are constant, the second-order term is automatically *in divergence form*, a distinction that will only become important in the case of varying coefficients studied later on in these notes.

To our end, we first remark that, as in the case of linear transport, the consideration of initial conditions that are unbounded already creates a difficulty for the *definition* of the terms of the equation. Indeed, the term

$$b_i \partial_i u = \partial_i (u b_i) - (\partial_i b_i) u$$

does not necessarily make sense for an unbounded function u. The difficulty has been circumvented for transport equations using a renormalization of the solution function u. In the presence of the second-order differential operator, renormalization is, at first sight, unclear. The question will be examined later, in Section 2.3, but, to start with, we will assume, in Section 2.1 that $u_0 \in L^\infty$.

In echo to our Remark 1, we also wish to immediately make clear that we will *not* consider the one-dimensional setting for (1). That particular setting allows for very specific arguments and results (one may typically consider fields $\mathbf{b}$ that are only L^1) which may not be extended to strictly higher dimensions. Also, because the two-dimensional setting is well known to be particular, we will begin with considering three or higher-dimensional settings. We devote a paragraph of Section 2.1 to the adaptation of the results to the two-dimensional setting.

Likewise, the case of an L^1 initial condition is somewhat peculiar and will be commented upon.

Miscellaneous remarks and comments are collected in Section 2.4.

Some useful elements of functional analysis. Before we proceed, we need to now recall, for consistency, some well known facts on the Lorentz spaces $L^{p,q}$, $1 \le p < +\infty$, $1 \le q \le +\infty$, and their particular cases, the Marcinkiewicz, or *weak-L^p*, spaces $L^{p,\infty} = M^p$, $1 \le p < +\infty$. For more details, we refer the reader to the classical textbooks [14, Chapter 1], [91, Chapter 1, Section 8], and also to [90]. For any measurable

https://doi.org/10.1515/9783110635508-002

function f on $\mathbb{R}^d$, we may define the function

$$\mu_f(s) = \mathrm{meas}\{x : |f(x)| > s\} \tag{2.0.1}$$

called the *distribution function* of f, and next the *non-increasing rearrangement* of f on $(0, \infty)$ by

$$f^*(t) = \inf\{s > 0 : \mu_f(s) \leq t\}. \tag{2.0.2}$$

It is immediate to see that f and its rearrangement f^* share the same distribution function. From f^* one may construct the *non-increasing spherically symmetric rearrangement* of f, also called the *Schwarz symmetrization* of f,

$$f^\#(x) = f^*(c_d|x|^d) \tag{2.0.3}$$

for $x \in \mathbb{R}^d$, where c_d is the $(d-1)$-dimensional Lebesgue measure of the unit sphere of $\mathbb{R}^d$. For the time being, we use these rearrangements to define some functional spaces. We introduce, when $1 \leq p < \infty$, $1 \leq q < \infty$, the *Lorentz* space $L^{p,q}$ of all measurable functions f such that

$$|f|_{L^{p,q}} = \left[\int_0^\infty \frac{1}{t}\,[t^{\frac{1}{p}} f^*]^q\,dt\right]^{\frac{1}{q}}, \tag{2.0.4}$$

or, equivalently,

$$|f|_{L^{p,q}} = \left[\int_0^\infty y^{q-1}\mu_f(y)^{\frac{q}{p}}\,dy\right]^{\frac{1}{q}} \tag{2.0.5}$$

is finite. We have $L^{p,p} = L^p$, as is easily seen on (2.0.4), since the non-increasing rearrangement f^* has the same L^p norms as f. The fact that (2.0.5) agrees with (2.0.4) can be intuitively explained as follows. Assume that $f = f^*$ in one dimension, and, in addition, that it is strictly decreasing, thus invertible. In that case, by strict monotonicity, $\mu_f(s) = f^{-1}(s)$, thus (2.0.4) also writes

$$|f|_{L^{p,q}} = \left[\int_0^\infty t^{\frac{q}{p}-1}[\mu_f^{-1}(t)]^q\,dt\right]^{\frac{1}{q}} \propto \left[\int_0^\infty y^q\,d(\mu_f(y)^{\frac{q}{p}})\right]^{\frac{1}{q}} \propto \left[\int_0^\infty y^{q-1}\mu_f(y)^{\frac{q}{p}}\,dy\right]^{\frac{1}{q}},$$

which is (2.0.5), successively using a change of variable $y = \mu_f^{-1}(t)$ and an integration by parts.

In the case $q = \infty$, the space $L^{p,\infty}$ is defined as the space of measurable functions f such that

$$\sup_{t>0} t^{\frac{1}{p}} f^*(t) < +\infty, \tag{2.0.6}$$

or that

$$\sup_{s>0} s\mu_f(s)^{\frac{1}{p}} = \sup_{s>0} s\,\mathrm{meas}\{x : |f(x)| > s\}^{\frac{1}{p}} < +\infty. \tag{2.0.7}$$

Quantities (2.0.4) and (2.0.7) are however *not* norms on the corresponding spaces. They are only quasi-norms, that is, they do not satisfy the triangle inequality. One thus defines

$$f^{**} = \frac{1}{t}\int_0^t f^*(s)\,ds.$$

When $1 < p$, $1 \le q < \infty$, f (or f^*) may be replaced by f^{**} in (2.0.4) and then defines a norm on the space $L^{p,q}$. Likewise, the space $L^{p,\infty}$, $1 < p < +\infty$, can be equipped with the norm

$$\|f\|_{L^{p,\infty}} = \sup_{t>0} t^{\frac{1}{p}} f^{**}(t). \tag{2.0.8}$$

The norms thereby defined are of course equivalent to the quasi-norms defined above; see [14, Section 1.6] or [81]. In the particular case ($p = 1$, $q = \infty$), the definition (2.0.8) yields $L^{1,\infty} = L^1$ while (2.0.7) defines a non-normable set. The spaces $L^{p,\infty}$, $1 < p$, are called the *Marcinkiewicz*, or *weak-L^p*, spaces. On a bounded domain, we have the sequence of inclusions

$$L^{p,1} \subset \cdots \subset L^{p,r} \subset \cdots \subset L^{p,p} = L^p \subset \cdots \subset L^{p,s} \subset \cdots \subset L^{p,\infty} \tag{2.0.9}$$

for all $1 < p < +\infty$, $1 \le r \le p \le s \le +\infty$. We also have, on any bounded set,

$$L^{p,q} \subset L^{r,s}, \tag{2.0.10}$$

whenever $1 \le r < p \le +\infty$, $1 \le q \le +\infty$, $1 \le s \le +\infty$. The typical example that shows how these spaces are different, and in particular that $L^{p,\infty}$ generically differs from all the spaces L^q is the function $\frac{1}{|x|}$, which will indeed be very useful below. In dimension $d \ge 2$, it belongs to $L^{d,\infty}$, is "almost" in $L^d(\mathbb{R}^d)$ but belongs to none of the spaces $L^q(\mathbb{R}^d)$, $1 \le q \le +\infty$.

We will repeatedly use in the sequel the extension of the classical Hölder and Young inequalities, respectively, to functions in the Lorentz spaces. This extension is quite straightforward. We recall it here for further reference. We have

$$\|fg\|_{L^{p,q}} \le C\|f\|_{L^{p_1,q_1}}\|g\|_{L^{p_2,q_2}} \tag{2.0.11}$$

for $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$, $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2}$, $1 \le p, p_1, p_2, q, q_1, q_2 \le +\infty$ (except the cases where one of the couples is $(1,\infty)$). Likewise,

$$\|f \star g\|_{L^{p,q}} \le C\|f\|_{L^{p_1,q_1}}\|g\|_{L^{p_2,q_2}} \tag{2.0.12}$$

for $1 \le q < +\infty$, $1 + \frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$, $\frac{1}{q} \le \frac{1}{q_1} + \frac{1}{q_2}$, $1 \le p, p_1, p_2, q_1, q_2 \le +\infty$ (except again the cases where one of the couples is $(1,\infty)$).

The above Young inequality (2.0.12) allows to prove the following important property: there exists a constant c (which may be made explicit; see below) such that

$$\|u\|_{L^{2d/(d-2),2}} \le c\|\nabla u\|_{L^2} \quad \text{in dimension } d > 2 \tag{2.0.13}$$

for all smooth, compactly supported functions. This inequality improves the classical Sobolev inequality

$$\|u\|_{L^{2d/(d-2)}} \le c\|\nabla u\|_{L^2}. \tag{2.0.14}$$

Using a density argument, inequality (2.0.13) is easily extended to all functions of $\mathbb{R}^d$ such that both sides are finite. Using a suitable cut-off function and again a density argument, it is also readily extended (and this is the specific form that we will use in the sequel of this contribution, given our periodic setting) in the form

$$\|u\|_{L^{2d/(d-2),2}} \le c\big(\|\nabla u\|_{L^2} + \|u\|_{L^2}\big) \tag{2.0.15}$$

to all periodic functions. There are many possible proofs for (2.0.13). The original proof rests on the convolution $u = -\Delta u \star G = -\partial_i u \star \partial_i G$, where G is the Green function of the (periodic) Laplacian operator in dimension $d > 2$. We notice that $\partial_i u \in L^{2,2} = L^2$, $\partial_i G \in L^{\frac{d}{d-1},\infty}$ since G has a singularity in $\frac{1}{|x|^{d-2}}$. We may thus apply (2.0.12) with $p = \frac{2d}{d-2}$, $q = 2$, $p_1 = 2$, $q_1 = 2$, $p_2 = \frac{d}{d-1}$, $q_2 = \infty$, which formally yields (2.0.13). The above proof can also be formalized using the Fourier transform, writing

$$\hat{u} = ((1 + |\xi|)\hat{u})(1 + |\xi|)^{-1}$$

and adjusting the exponents in an Hölder estimate using the Lorentz spaces the function $(1 + |\xi|)^{-1}$ belong to. We would like to mention a strategy of proof that does not rest upon the use of the generalized Young inequality (2.0.12) but only on the classical Sobolev inequality (2.0.14). That strategy of proof can be adapted to more general settings; see our remarks after Lemma 11 regarding the hypoellipticity setting. A manner to actually prove (2.0.13) relies upon the idea of (a) decomposing ∇u into a sum $\nabla u = \sum \nabla \varphi_n(u)$, where the $\nabla \varphi_n(u)$ are (almost) orthogonal in L^2, and (b) use the classical Sobolev inequality for each $\varphi_n(u)$. This technique is quite general: it only requires that the classical estimate involve a first-order derivative, and automatically allows, then, to recover a second exponent $q = 2$, in the notation of the Lorentz spaces. The outline of that proof goes as follows. We introduce, for $n \in \mathbb{Z}$,

$$\varphi_n(t) = \begin{cases} 0 & \text{when } |t| \le 2^{n-1}, \\ t - 2^{n-1} & \text{when } 2^{n-1} \le |t| \le 2^{n+1}, \\ 2^{n+1} - 2^{n-1} & \text{when } 2^{n+1} \le |t|. \end{cases} \tag{2.0.16}$$

We notice that

$$\nabla \varphi_n(u) = \mathbb{1}_{2^{n-1} \le |u| \le 2^{n+1}} \nabla u, \tag{2.0.17}$$

$$\varphi_n(u) \ge 2^{n-2} \mathbb{1}_{|u| \ge 2^n - 2^{n-2}}, \tag{2.0.18}$$

two facts which will be useful below. For all n, we use the classical Sobolev inequality (2.0.14) for the function $\varphi_n(u)$:

$$\|\nabla \varphi_n(u)\|_{L^2}^2 \ge \mu \|\varphi_n(u)\|_{L^p}^2 \ge \nu 2^{2n} \big(\text{meas}\{|u| \ge 2^n - 2^{n-2}\}\big)^{\frac{2}{p}},$$

where $p = \frac{2d}{d-2}$, and μ and ν denote irrelevant positive constants independent on n and u, and where we have used (2.0.18). Adding these inequalities for all $n \ge -k$ for $k \in \mathbb{N}$, we obtain

$$\sum_{n \ge -k} \|\nabla \varphi_n(u)\|_{L^2}^2 \ge \eta \sum_{n \ge -k} 2^{2n} \big(\text{meas}\{|u| \ge 2^n\}\big)^{\frac{2}{p}} \tag{2.0.19}$$

for another positive constant η. We next note that the $L^{p,2}$ quasinorm writes

$$|u|^2_{L^{p,2}} = \int_0^\infty y\mu_u(y)^{\frac{2}{p}}\,dy.$$

By the monotonicity of μ_u, the series on the right-hand side of (2.0.19) and the above integral can be compared to one another, and we infer from (2.0.19) that

$$\sum_{n\geq -k} \|\nabla\varphi_n(u)\|^2_{L^2} \geq \int_{2^{-k}}^\infty y\mu_u(y)^{\frac{2}{p}}\,dy, \tag{2.0.20}$$

again up to an irrelevant multiplicative factor. On the other hand, we can remark that because of the overlap between the regions $2^{n-1} \leq |t| \leq 2^{n+1}$ and $2^{n-2} \leq |t| \leq 2^n$, for all n, we have

$$2\int |\nabla u|^2 \geq \sum_{n\in\mathbb{Z}} \int |\nabla\varphi_n(u)|^2. \tag{2.0.21}$$

Inserting (2.0.21) into (2.0.20), and next letting k go to infinity on the right-hand side, we obtain (2.0.13).

We conclude this paragraph recalling the following standard notation. We will write, e.g., $f \in \varepsilon L^{p_1,q_1} + L^{p_2,q_2}$ to mention that $f = f_1 + f_2$, where $f_1 \in L^{p_1,q_1}$, $f_2 \in L^{p_2,q_2}$ and $\|f_1\|_{L^{p_1,q_1}} \leq \varepsilon$. In particular, $f \in \varepsilon L^{p_1,q_1} + L^{p_2,q_2}$ for all $\varepsilon > 0$ means that, for any arbitrary small $\varepsilon > 0$, we have such a decomposition with $\|f_1\|_{L^{p_1,q_1}} \leq \varepsilon$.

2.1 Bounded initial condition

We now study the case of a constant second-order term **a** and a bounded initial condition u_0. Our main result, Theorem 1, assumes that **a** is a positive definite matrix and that the ambient dimension is $d \geq 3$. We provide many comments on our assumptions and results in Section 2.1.2 and the subsequent subsections. In particular, that section contains the adaptation of our technique of proof to the two-dimensional setting, while the case where **a** is only nonnegative is addressed in Corollary 1.

2.1.1 The main result and its proof

Theorem 1 ($d \geq 3$, **a** constant positive definite, u_0 bounded). *Assume that the ambient dimension d is larger than or equal to* 3, *and that the matrix coefficient* **a** *in* (1) *is a* constant, positive definite *symmetric matrix. Supply equation* (1) *with a bounded initial condition $u_0 \in L^\infty$. Assume also that the transport field* **b** *satisfies*

$$\mathbf{b} = \beta_1 + \beta_2, \quad \beta_1 \in L^2, \qquad \beta_2 \in W^{1,1}, \tag{2.1.1}$$

$$\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2, \quad \mathbf{b}_1 \in \varepsilon_1 L^{d,\infty} + L^d, \quad [\operatorname{div}\mathbf{b}_2]_- \in \varepsilon_2 L^{\frac{d}{2},\infty} + L^{\frac{d}{2}} \tag{2.1.2}$$

for some positive constants $\varepsilon_1, \varepsilon_2$ *which will be made explicit in the proof (see* (2.1.15)*) in terms of* $\mathbf{a}$ *and* $\mathbf{b}$*. Then there exists a unique solution* u *to* (1) *in the functional space* $L^2([0,T], H^1) \cap L^\infty([0,T], L^\infty)$*. In addition, for all* $1 < p < +\infty$*, that solution is in* $C([0,T], L^p)$*, and there exists some constant* C_0 *such that the solution satisfies*

$$\|u(t,\cdot)\|_{L^p} \le e^{C_0 t} \|u_0\|_{L^p} \tag{2.1.3}$$

for all $t \in [0,T]$*. In the case* $p = +\infty$*, estimate* (2.1.3) *holds with* $C_0 = 0$ *and we have the maximum principle*

$$\inf u_0 \le u(t,\cdot) \le \sup u_0. \tag{2.1.4}$$

An estimate in the spirit of (2.1.3) *will be established for* $p = 1$ *in Lemma 2 below, see* (2.1.24).

Some simple remarks are in order:
- The case $d = 2$ will be addressed below.
- We recall that to keep the notation simple, we denote $\mathbf{b} \in L^2$ even though $\mathbf{b}$ is valued in $\mathbb{R}^d$ and the adequate notation would be $\mathbf{b} \in (L^2)^d$.
- As will be evident from the proof, assumption (2.1.1) is useful for the regularization step of the proof, while (2.1.2) is employed for the uniqueness.
- Estimate (2.1.3) also holds true for the positive $[u(t,\cdot)]_+$ and negative $[u(t,\cdot)]_-$ parts of $u(t,\cdot)$
- Upon assuming that the initial condition u_0 is more regular, we may obtain a better regularity of the solution. A strategy for addressing this question will be briefly outlined in Section 4.3.

Besides the immediate remarks above, our major two comments on the statement of Theorem 1 are respectively related to our assumptions (2.1.2) on $\mathbf{b}$ and a specific estimation of the form (2.1.3) for the L^1 norm. We will comment later in this chapter upon the above issues and some extensions. The remaining part of this section is devoted to the proof of Theorem 1.

As mentioned above, assumptions (2.1.1) and (2.1.2) are useful for the regularization step and the uniqueness step, respectively. These two steps have been outlined in Section 1.1 for the case of transport equations. We will therefore only concentrate ourselves on the modifications necessary to adapt the scheme of proof to our present context.

Regularization. Assuming (2.1.1) allows to proceed with the regularization step as we did in our earlier work [64]. More precisely, we have:

Lemma 1 (reproduced from [64, Lemma 5.1], see also Lemma 10). *Let* $b \in L^2 + W^{1,1}$, $u \in L^\infty \cap H^1$*. Let* $\rho_\varepsilon = \varepsilon^{-d}\rho(\varepsilon^{-1}\cdot)$*, where* ρ *is a fixed, nonnegative, compactly supported, smooth function. Then*

$$[b\,.\nabla, \rho_\varepsilon](u) \xrightarrow{\varepsilon\to 0} 0 \quad \text{in } L^1. \tag{2.1.5}$$

Intuitively, Lemma 1 holds because, by linearity, the commutator in (2.1.5) is decomposed into a commutator with a $W^{1,1}$ transport field treated as in (1.1.10) for the transport equation since in particular $u \in L^\infty$, and a commutator with an L^2 transport field where we explicitly use, then, that $u \in H^1$ as we briefly outlined in (1.3.4) of Section 1.1. The detailed proof can be read in [64].

The first novelty in the proof of Theorem 1 is therefore the formal proof of uniqueness, which will be made rigorous by the above regularization step. We now give the details of that part of the proof.

Uniqueness. Assume we have a solution u to (1). We proceed formally. We multiply (1) by u and integrate over the periodic cell to obtain

$$\frac{1}{2}\frac{d}{dt}\int |u|^2 + \int a_{ij}\partial_i u \partial_j u = \int \mathbf{b}_1 \,.\, \nabla u\, u - \frac{1}{2}\int |u|^2 \operatorname{div} \mathbf{b}_2 \tag{2.1.6}$$

after some evident integration by parts. We note that the second term of the left-hand side may be bounded from below

$$\int a_{ij}\partial_i u \partial_j u \geq \underline{a}\int |\nabla u|^2, \tag{2.1.7}$$

where $\underline{a}$ is the coerciveness constant of the positive definite matrix $\mathbf{a}$. The two terms of the right-hand side are respectively bounded from above as follows. We first decompose $\mathbf{b}_1$ as $\mathbf{b}_1 = \mathbf{b}_{11} + \mathbf{b}_{12}$ with $\mathbf{b}_{11} \in \varepsilon_1 L^{d,\infty}$ (for some ε_1 which will be made precise below) and $\mathbf{b}_{12} \in L^d$. For all $\delta > 0$ presumably small, we in turn decompose $\mathbf{b}_{12}$ into

$$\mathbf{b}_{12} = \mathbf{b}_{12,\delta} + (\mathbf{b}_{12} - \mathbf{b}_{12,\delta}) \tag{2.1.8}$$

with $\|\mathbf{b}_{12,\delta}\|_{L^d} \leq \delta$ and $\mathbf{b}_{12} - \mathbf{b}_{12,\delta} \in L^\infty$. For this purpose, we simply define

$$\mathbf{b}_{12,\delta} = \mathbf{b}_{12}\,\mathbb{1}_{\{x;|\mathbf{b}_{12}(x)|\geq R\}}$$

and choose the cut-off parameter R sufficiently large so that $\|\mathbf{b}_{12,\delta}\|_{L^d} \leq \delta$. We then write

$$\begin{aligned}
\left|\int \mathbf{b}_1 \,.\, \nabla u\, u\right| &\leq \left|\int \mathbf{b}_{11} \,.\, \nabla u\, u\right| + \left|\int \mathbf{b}_{12,\delta} \,.\, \nabla u\, u\right| + \left|\int (\mathbf{b}_{12} - \mathbf{b}_{12,\delta}) \,.\, \nabla u\, u\right| \\
&\leq \|\mathbf{b}_{11}\|_{L^{d,\infty}}\|\nabla u\|_{L^2}\|u\|_{L^{2d/(d-2),2}} + \|\mathbf{b}_{12,\delta}\|_{L^d}\|\nabla u\|_{L^2}\|u\|_{L^{2d/(d-2)}} \\
&\qquad + \|\mathbf{b}_{12} - \mathbf{b}_{12,\delta}\|_{L^\infty}\|\nabla u\|_{L^2}\|u\|_{L^2} \\
&\leq c(\|\mathbf{b}_{11}\|_{L^{d,\infty}} + \|\mathbf{b}_{12,\delta}\|_{L^d})\|\nabla u\|_{L^2}^2 \\
&\qquad + (\|\mathbf{b}_{12} - \mathbf{b}_{12,\delta}\|_{L^\infty} + 1)\|\nabla u\|_{L^2}\|u\|_{L^2}
\end{aligned} \tag{2.1.9}$$

successively using the Hölder inequalities (2.0.11), inequality (2.0.15), the fact that we may always assume, say, $c(\varepsilon_1 + \delta) \leq 1$ and the fact that $L^{\frac{2d}{d-2},2} \subset L^{\frac{2d}{d-2}}$ (since $\frac{2d}{d-2} \geq 2$). The rightmost term of (2.1.9) is now estimated using the classical discrete Young inequality

$$\begin{aligned}
(\|\mathbf{b}_{12} - \mathbf{b}_{12,\delta}\|_{L^\infty} + 1)\|\nabla u\|_{L^2}\|u\|_{L^2} &\leq \varepsilon_3\big(\|\mathbf{b}_{12} - \mathbf{b}_{12,\delta}\|_{L^\infty}^2 + 1\big)\|\nabla u\|_{L^2}^2 \\
&\qquad + \frac{1}{4\varepsilon_3}\|u\|_{L^2}^2,
\end{aligned} \tag{2.1.10}$$

where the small constant ε_3 will be adjusted below. Inserting (2.1.10) into (2.1.9), we obtain

$$\left|\int \mathbf{b}_1 . \nabla u\, u\right| \le c\big(\|\mathbf{b}_{11}\|_{L^{d,\infty}} + \|\mathbf{b}_{12,\delta}\|_{L^d} + \varepsilon_3 \|\mathbf{b}_{12} - \mathbf{b}_{12,\delta}\|^2_{L^\infty} + \varepsilon_3\big)\|\nabla u\|^2_{L^2} + \frac{1}{4\varepsilon_3}\|u\|^2_{L^2}, \tag{2.1.11}$$

where we recall that $\|\mathbf{b}_{11}\|_{L^{d,\infty}} \le \varepsilon_1$ by assumption, and $\|\mathbf{b}_{12,\delta}\|_{L^d} \le \delta$ by construction.

For the rightmost term of (2.1.6), we proceed somewhat similarly, manipulating $\operatorname{div} \mathbf{b}_2$ instead of $\mathbf{b}_1$. We first split $[\operatorname{div} \mathbf{b}_2]_-$ as $[\operatorname{div} \mathbf{b}_2]_- = f_1 + f_2$ with $f_1 \in \varepsilon_2 L^{\frac{d}{2},\infty}$ and $f_2 \in L^{\frac{d}{2}}$ for a small positive parameter ε_2 yet to be determined. Next, we write

$$f_2 = f_{2,\delta} + (f_2 - f_{2,\delta})$$

with $f_{2,\delta} = f_2\, \mathbb{1}_{\{x; |f_2(x)| \ge R\}}$, a cut-off parameter R such that $\|f_{2,\delta}\|_{L^{d/2}} \le \delta$ and a remainder $f_2 - f_{2,\delta} \in L^\infty$ by construction. We now write

$$\int [\operatorname{div} \mathbf{b}_2]_- |u|^2 \le \|f_1\|_{L^{d/2,\infty}} \|u\|^2_{L^{2d/(d-2),2}} + \|f_{2,\delta}\|_{L^{d/2}} \|u\|^2_{L^{2d/(d-2)}} + \|f_2 - f_{2,\delta}\|_{L^\infty} \|u\|^2_{L^2} \tag{2.1.12}$$

using both the classical Hölder inequality and the Hölder inequality for Marcinkiewicz spaces recalled in (2.0.11). The first two terms are next treated using inequality (2.0.15). We obtain

$$\int [\operatorname{div} \mathbf{b}_2]_-\, |u|^2 \le c\big(\|f_1\|_{L^{d/2,\infty}} + \|f_{2,\delta}\|_{L^{d/2}}\big)\|\nabla u\|^2_{L^2} + \big(\|f_2 - f_{2,\delta}\|_{L^\infty} + c\big)\|u\|^2_{L^2}, \tag{2.1.13}$$

with, we recall, $\|f_1\|_{L^{d/2,\infty}} \le \varepsilon_2$, $\|f_{2,\delta}\|_{L^{d/2}} \le \delta$. We now collect (2.1.7), (2.1.11), (2.1.13), and insert them into (2.1.6) to obtain

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|u\|^2_{L^2} + \underline{a}\|\nabla u\|^2_{L^2} \le c\Big(&\|\mathbf{b}_{11}\|_{L^{d,\infty}} + \|\mathbf{b}_{12,\delta}\|_{L^d} + \varepsilon_3 \|\mathbf{b}_{12} - \mathbf{b}_{12,\delta}\|^2_{L^\infty} \\ &+ \varepsilon_3 + \|f_1\|_{L^{d/2,\infty}} + \|f_{2,\delta}\|_{L^{d/2}}\Big)\|\nabla u\|^2_{L^2} \\ &+ \Big(\frac{1}{4\varepsilon_3} + \|f_2 - f_{2,\delta}\|_{L^\infty} + c\Big)\|u\|^2_{L^2}. \end{aligned} \tag{2.1.14}$$

We choose ε_1, ε_2, δ and finally ε_3 all sufficiently small so that

$$c\,(\varepsilon_1 + \delta + \varepsilon_3 \|\mathbf{b}_{12} - \mathbf{b}_{12,\delta}\|^2_{L^\infty} + \varepsilon_3 + \varepsilon_2 + \delta) \le \frac{1}{2}\underline{a}, \tag{2.1.15}$$

thereby obtaining

$$\frac{1}{2}\frac{d}{dt}\|u\|^2_{L^2} + \frac{1}{2}\underline{a}\|\nabla u\|^2_{L^2} \le C\|u\|^2_{L^2}. \tag{2.1.16}$$

This shows both an a priori estimate on the solution u in $L^\infty([0,T], L^2) \cap L^2([0,T], H^1)$ and also, when starting from a zero initial condition, uniqueness. Both arguments can readily be made rigorous using the regularization procedure.

Existence. As already mentioned in the introduction of these notes, we can easily establish, using the above formal a priori estimates, the existence of a solution to (1). This is classical. We indeed proceed by approximation. The field **b** is approximated by a sequence $\mathbf{b}_\varepsilon$ of smooth functions, and for each ε the existence (and uniqueness) of a solution u_ε to (1) with $\mathbf{b}_\varepsilon$ instead of **b** is established by classical tools.

Given that the equation is parabolic regular and the initial condition is bounded, we first note, by a simple application of the maximum principle, that the solution is bounded. It more precisely satisfies (2.1.4). We mention in passing that this estimate is intuitively simpler when we think of the Feynman–Kac expression of the solution $\mathbb{E}_x(u_0(\mathbf{X}_t))$ obtained with the probabilistic interpretation of equation (1) (see (1.2.3) above, and more details along those lines in Section 2.2 below).

Likewise, the formal a priori estimate (2.1.16) established above has a rigorous sense for the solution u_ε to equation (1) with $\mathbf{b}_\varepsilon$. Therefore u_ε belongs to the functional space $L^2([0,T],H^1)\cap L^\infty([0,T],L^2)$.

Also, formally multiplying (1) by $pu|u|^{p-2}$ for $p>1$, we obtain

$$\frac{d}{dt}\int |u|^p + p(p-1)\int |u|^{p-2}a_{ij}\partial_i u\partial_j u = -\int |u|^p \operatorname{div}\mathbf{b}. \tag{2.1.17}$$

All the formal manipulations we have performed above following (2.1.6) can again be performed using $v=|u|^{\frac{p}{2}}$ instead of u. They lead to an estimate analogous to (2.1.16), namely

$$\frac{d}{dt}\int |u|^p + \underline{a}\int |u|^{p-2}|\nabla u|^2 \le C\int |u|^p \tag{2.1.18}$$

for some suitable constant C. This formally gives (2.1.3).

Similarly to (2.1.16), estimate (2.1.18) is rigorous for u_ε because of the regularity of the approximate transport term $\mathbf{b}_\varepsilon$. The continuity in time of u_ε with values in L^p, $1\le p<+\infty$, is then obtained classically using estimate (2.1.17) (made rigorous for u_ε) and an embedding argument.

Passing to the limit $\varepsilon\to 0$ in all the above estimates (once integrated in time) is performed as in the case of the transport equation (see the details in [41]) and provides existence of a solution u in the appropriate functional space

$$L^2([0,T],H^1)\cap C([0,T],L^p)\cap L^\infty([0,T],L^\infty),$$

for all $1\le p<+\infty$, satisfying in addition properties (2.1.3) and (2.1.4) stated in Theorem 1. This concludes the proof of Theorem 1. □

2.1.2 On our assumptions on b and their "optimality".

Recall that a direct application of the classical result of [41] in the specific context addressed here would simply allow to establish existence and uniqueness of the solution in $C([0,T],L^p)\cap L^\infty([0,T],L^\infty)$, for all $1\le p<+\infty$, under the assumption $\mathbf{b}\in W^{1,1}$ and $[\operatorname{div}\mathbf{b}]_-\in L^\infty$. Next, we need to emphasize that in [64, Proposition 5.3] we had already remarked that the presence of a positive definite second-order term with constant coefficients (that is, up to an irrelevant change of coordinates, a Laplacian) allows to extend the traditional $W^{1,1}$ assumption on the transport field $\mathbf{b}$ and to precisely assume that $\mathbf{b} \in L^2 + W^{1,1}$. But this was performed in the case of a divergence-free field $\mathbf{b}$, or essentially equivalently, of a *bounded* divergence. Theorem 1 generalizes the result to the case when the divergence (or more precisely, as always if one is only interested in positive times, its positive part) belongs to $\varepsilon_1 L^{\frac{d}{2},\infty} + L^{\frac{d}{2}}$. Note the extension to $L^{\frac{d}{2}}$ is immediate from the proof of the results contained in [64], the difficult part being the $L^{\frac{d}{2},\infty}$ part. Much more importantly, Theorem 1 proves uniqueness in a case when the divergence is *not* controlled: one should indeed notice that there is no assumption on $\operatorname{div} \mathbf{b}_1$ in (2.1.2). Now, the part $\mathbf{b}_1$ of $\mathbf{b}$ that has a divergence we do not control is intuitively the difficult part. This can be understood in particular thinking in terms of fluid mechanics. The divergence-free part (or more generally the part $\mathbf{b}_2$ of the transport field that has a controlled divergence) induces rotation of the fluid particles, like a *curl* does. It does not concentrate the flow anywhere. Mathematically, no singularity in the transport of the Lebesgue measure, owing to the action of such a transport term, is thus to be expected. This is reflected in the absence of assumption put on $\mathbf{b}_2$ itself in (2.1.2). On the other hand, the part $\mathbf{b}_1$ may concentrate, even critically, the trajectories of fluid particles, thereby creating a singularity in the solution u. Thus the need for stringent assumptions on $\mathbf{b}_1$. For all these reasons, we now concentrate our discussion on $\mathbf{b}_1$.

The natural question is whether our assumption $\mathbf{b}_1 \in \varepsilon_1 L^{d,\infty} + L^d$ in (2.1.2) is sharp. As can easily be understood considering the bilinear form $\int \mathbf{b}_1 \,.\, \nabla u\, v$ for u and v in the natural energy space for (1), that is, H^1, the space L^d plays a natural role for the continuity of this form. Recall indeed that, using the Hölder inequality,

$$\left| \int \mathbf{b}_1 \,.\, \nabla u\, v \right| \le \|\mathbf{b}\|_{L^d} \|\nabla u\|_{L^2} \|v\|_{L^{2d/(d-2)}}.$$

The Marcinkiewicz space $L^{d,\infty}$ being an adaptation of the space L^d (intuitively containing the functions that are “almost” L^d) and allowing for the extension (2.0.11) of the classical Hölder inequalities, we also understand it plays a natural role in our theory. A more subtle question is, why the size of $\mathbf{b}_1$ in the Marcinkiewicz space $L^{d,\infty}$ matters.

It is unclear whether the size of $\mathbf{b}$ can be arbitrary in $L^{d,\infty}$. But some intuitive facts, which we now collect, seem to indicate our smallness assumption is necessary. To intuitively illustrate that the transport field $\mathbf{b}_1$ cannot be arbitrarily large in $L^{d,\infty}$, we temporarily consider the stochastic differential equation associated to (1) in the case when the second-order operator is $a_{ij} = \frac{1}{2}\delta_{ij}$ and the transport field is

$$\mathbf{b} = \mathbf{b}_1 = -\lambda \frac{x}{\|x\|^2}, \tag{2.1.19}$$

where $\|x\|$ denotes the Euclidean norm in $\mathbb{R}^d$ and λ is a positive constant, the size of which will play a critical role in the sequel. Of course, $\mathbf{b}_1$ is chosen so that it belongs to $L^{d,\infty}$, and not L^d. Notice that in our argument, and this will be true for all the counterexamples we shall consider throughout these notes, we focus on the behavior at the origin: by (2.1.19), we actually mean that we consider any field $\mathbf{b}$ (compatible with all the other necessary conditions, such as, e.g., boundary conditions) that, locally, at the vicinity of the origin, behaves like $-\lambda \frac{x}{\|x\|^2}$. The stochastic differential equation associated to the parabolic equation

$$\partial_t u + \lambda \frac{x}{\|x\|^2} . \nabla u - \frac{1}{2} \Delta u = 0 \tag{2.1.20}$$

reads

$$d\mathbf{X}_t = -\lambda \frac{\mathbf{X}_t}{\|\mathbf{X}_t\|^2}\, dt + d\mathbf{W}_t. \tag{2.1.21}$$

The reader familiar with that notion immediately realizes, either seeing the infinitesimal generator in (2.1.20) or on (2.1.21), that, for λ sufficiently large, $\|\mathbf{X}_t\|$ is a Bessel process of negative dimension. Difficulties as far as existence and uniqueness of the process are thus expected when the origin is reached. Using the chain rule of stochastic calculus, we have

$$d(\ln \|\mathbf{X}_t\|)_t = \left(\frac{d-2}{2} - \lambda\right) \frac{1}{\|\mathbf{X}_t\|^2}\, dt + \frac{\mathbf{X}_t}{\|\mathbf{X}_t\|^2} . d\mathbf{W}_t. \tag{2.1.22}$$

We observe on (2.1.22) that when $\lambda > \frac{d-2}{2}$, the drift term blows up to $-\infty$ when $\mathbf{X}_t \to 0$: the trajectories get concentrated toward the origin. It is then clear that for such values of λ, mathematical pathologies will happen. They are indeed well documented in the literature. These pathologies result in the possible loss of uniqueness of the process.

Another intuitive argument that shows that an arbitrarily large $\mathbf{b}_1$ in $L^{d,\infty}$ may be incompatible with uniqueness consists, following up on our comments of Section 1.3.2, page 23, in considering the static version of (2.1.20), namely

$$\lambda \frac{x}{\|x\|^2} . \nabla u - \frac{1}{2} \Delta u = f. \tag{2.1.23}$$

To simplify matters, we consider this equation on the ball $B(0, R)$ of radius R centered at the origin, with homogeneous Dirichlet boundary conditions on the sphere. Subtracting two solutions u_1 and u_2 of (2.1.23), we obtain a solution $u = u_1 - u_2$ to (2.1.23) with zero right-hand side and homogeneous Dirichlet boundary conditions on the sphere. We now show that there are infinitely many radially symmetric such solutions, which implies that uniqueness cannot hold for (2.1.23). Indeed, in radial coordinates in $\mathbb{R}^d$, we have $-\frac{1}{2}u'' + (\lambda - \frac{d}{2} + \frac{1}{2})\frac{u'}{r} = 0$. We integrate this into $u' = cr^{2\lambda-d+1}$ for some constant c. It immediately follows from this observation that, as soon as $2\lambda - d + 1 > -1$, that is, $\lambda > \frac{d-2}{2}$, the function u' is integrable, so that one may arbitrarily prescribe the value of u at the origin. We thus have the set of solutions $u = u(0) + c\,(2\lambda - d + 2)^{-1}\, r^{2\lambda-d+2}$ with c such that u satisfies the boundary condition $u(R) = 0$. This contradicts uniqueness for equation (2.1.23) (in echo to our remarks of Section 1.3.2, page 23, note however that no zeroth-order term of the type $\alpha\, u$, for α large, is present in (2.1.23)). In some formal sense, this *suggests* (and certainly intuitively confirms) that the smallness of $\mathbf{b}$ in $L^{d,\infty}$ is required for the parabolic equation to have a unique solution.

2.1.3 The L^1 estimate

We now establish an estimate in the spirit of (2.1.3) for $p = 1$. For such an estimate, the case $p = 1$ requires stronger assumptions on the transport field $\mathbf{b}$ than the case $p > 1$. We would like to now briefly explain why. The difficulty in Theorem 1, and the added value with respect to a direct, naive adaptation of the results established in [41] for the pure transport equation, is the ability to address transport fields $\mathbf{b}$ the divergence of which is not necessarily controlled. Indeed, the usual assumptions on $\mathbf{b}$ for the transport equation recalled in Section 1.1, namely ($\mathbf{b} \in W^{1,1}$, $[\operatorname{div}\mathbf{b}]_- \in L^\infty$) are usually necessary for the regularization of the equation and the a priori estimates. This holds for the transport equation in the absence of any further regularity of the solution u. In the presence of the regularizing second-order term, however, less regular fields $\mathbf{b}$ can be addressed. In particular, for the a priori estimates, one may consider, then, a transport field the divergence of which is not completely controlled. This is the essence of our assumption (2.1.2) in the statement of Theorem 1. In the L^p setting, for $p > 1$, this intrinsically stems from the strong convexity of the function $t \to |t|^p$: regularity and integrability are gained from that strong convexity. In contrast, when $p = 1$, the function $t \to |t|$ is *not* strongly convex and we gain nothing with arguments based on direct a priori estimates. Some difficulties arise and we have to impose slightly more demanding assumptions on $\mathbf{b}$ than those of Theorem 1. Lemma 2 below considers an adequate set of such assumptions. We note in passing, and this will be the purpose of the first of our many remarks below, after the statement of Lemma 2, that strong convexity is reinstated if we consider $|t|\ln|t|$ instead of $|t|$ and work in the functional space $L\ \ln L$ instead of L^1.

Remark 2. Given the lack of strict convexity of the L^1 norm, one could also consider trying to "gain in convexity" using the classical De la Vallée Poussin Lemma. For each $u \in L^1$, we indeed know there exists a convex, non-decreasing function g, satisfying $g(0) = 0$ and $\lim_{t\to+\infty} \frac{g(t)}{t} = +\infty$, such that $g(u) \in L^1$. Therefore, multiplying the equation

$$\partial_t u - \mathbf{b}.\nabla u - \Delta u = 0$$

by $g'(u)$, one formally obtains

$$\frac{d}{dt}\int g(u) + \int (\operatorname{div}\mathbf{b})g(u) + \int g''(u)|\nabla u|^2 = 0.$$

Unfortunately, we do not know how to proceed from there, unless we *assume* additional properties on the function g. We therefore do not pursue this direction.

We also outright note that, of course, in Lemma 2, we may add to the transport field $\mathbf{b}$ any other transport field $\tilde{\mathbf{b}}$ that satisfies the stronger assumptions ($\tilde{\mathbf{b}} \in W^{1,1}$, $[\operatorname{div}\tilde{\mathbf{b}}]_- \in L^\infty$) since for that latter part, the L^1 estimate evidently holds since it already holds for the pure transport equation and the laplacian is a nonnegative operator.

We will return to the specific question of the L^1 estimate repeatedly throughout these notes, and in particular in Section 3.4 and Section 4.1.4 below.

Lemma 2 (L^1 estimate). *Consider the solution u of which the existence and uniqueness has been established in Theorem 1. Assume, in addition to the assumptions of this theorem, that*

- $\mathbf{b} \in L^{d,1}$,
- $\mathbf{b}$ *is independent of time.*

Then there exist some constants C_0 and C_1 (the latter not necessarily equal to one) such that

$$\|u(t,\cdot)\|_{L^1} \leq C_1 e^{C_0 t} \|u_0\|_{L^1} \tag{2.1.24}$$

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

Remark 3. As for (2.1.3), the L^1 estimate (2.1.24) also holds true for the positive and negative parts of u.

Before we give the proof of Lemma 2, several comments are in order.

Firstly, we comment upon a methodological point. Our L^p estimates, for $p > 1$, are obtained successively multiplying (1) by a suitable function (typically u^{p-1}), integrating over the domain, and integrating by parts. The general procedure has been outlined in formulae (2.1.17)–(2.1.18). The natural extension of such arguments for $p > 1$ to the case $p = 1$ yields estimates in the space $L \ln L$ and not L^1. Intuitively, we have, for $x > 0$ fixed, $\lim_{p\to 1} x\frac{x^{p-1}-1}{p-1} = x \ln x$ so that $x \ln x$ formally takes the role of x^p. More precisely, since

$$p\left[\frac{u^{p-1}-1}{p-1}\right]\frac{\partial u}{\partial t} = \frac{\partial}{\partial t}\left[u\frac{u^{p-1}-1}{p-1} - u\right],$$

the same formal observation tells us that the limit of our arguments – when $p \to 1$ – involves the function $u \ln u - u$. Therefore, it is easy, adapting our arguments, to obtain an extension of estimation (2.1.3) in $L \ln L$ instead of L^p. An interesting application of that type of reasoning to our particular context is made slightly more precise in Remark 36 below. Notice that this observation is quite general, methodologically. Another classical illustration of such a phenomenon is given by the properties (see [91, p. 84]) of the (Hardy–Littlewood) *maximal function*

$$Mf(x) = \sup_{B \ni x} \fint_B |f(y)|\, dy$$

for which the classical inequalities

$$\|Mf\|_{L^p} \leq C\|f\|_{L^p}$$

(on bounded domains) for $1 < p < +\infty$ are extended, for $p = 1$, to a “weak” inequality

$$\|Mf\|_{L^{1,\infty}} \leq C\|f\|_{L^1},$$

or alternately to the inequality

$$\|Mf\|_{L^1} \leq C + C\||f| \ln(2 + |f|)\|_{L^1}$$

indeed involving the space $L \ln L$, since $f \in L^1$ alone does not imply that $Mf \in L^1_{\mathrm{loc}}$ (see more details on all these issues, e.g., in [39]).

Secondly, we emphasize, and this will be clear in the proof, that, for now (see Section 3.4 and Section 4.1.4), we are only able to establish the result when **b** is independent of time (which, in the context of this manuscript, means that, in contrast to all the other results presented herein, the "usual" adaptation of our proof to time-dependent transport fields is *not yet* possible).

Thirdly, we have emphasized in the statement of Lemma 2 that the constant C_1 in the L^1 estimate (2.1.24) is not necessarily one. Only under a more stringent assumption on **b**, namely $[\operatorname{div}\mathbf{b}]_- \in L^\infty$, is (2.1.24) true with $C_1 = 1$. To start with, we observe that, if $[\operatorname{div}\mathbf{b}]_- \in L^\infty$, then the formal estimate (2.1.17) shows (2.1.24) with $C_1 = 1$. That fact is already true (see (1.1.8)) for the transport equation (1.1.1). In both cases, the estimate may be made rigorous using suitable arguments from [41]. As it turns out, the condition $[\operatorname{div}\mathbf{b}]_- \in L^\infty$ is actually *necessary* for (2.1.24) to hold with $C_1 = 1$. Indeed, let us assume $C_1 = 1$ and prove that, then, we necessarily have a better bound on $\operatorname{div}\mathbf{b}$ than that we have assumed. Take the initial condition $u_0 \geq 0$, so that by the maximum principle $u \geq 0$ for all times, and thus the L^1 estimate (2.1.24) reads

$$\int u(t,\cdot) \leq e^{C_0\,t} \int u_0,$$

that is,

$$\int \frac{u(t,\cdot) - u_0}{t} \leq \frac{e^{C_0\,t} - 1}{t} \int u_0.$$

Formally taking the limit when t goes to zero (and this can be justified rigorously), we obtain

$$\int \frac{\partial u(t,\cdot)}{\partial t}\bigg|_{t=0} \leq C_0 \int u_0.$$

But since

$$\int \frac{\partial u}{\partial t} = \int \mathbf{b}.\nabla u + \int \Delta u = -\int u \operatorname{div}\mathbf{b} \;+\; 0,$$

this shows $\int (-\operatorname{div}\mathbf{b})u_0 \leq C_0 \int u_0$, which holds for all initial conditions $u_0 \geq 0$, first in L^∞, the setting we have worked in so far, and next by density in L^1. This implies $(-\operatorname{div}\mathbf{b}) \leq C_0$, or equivalently $[\operatorname{div}\mathbf{b}]_- \leq C_0$, a condition which was *not* assumed beforehand. Our point in Lemma 2 is, given the (non-degenerate) parabolic setting, to obtain an L^1 estimate with a weaker assumption. This comes at the price of having a constant C_1 possibly larger than unity.

Fourthly, we wish to further illustrate the fact that an estimate in L^1 norm such as (2.1.24) is sensitive to many details of the problem under consideration. To this end, we now change the perspective and take the probability theoretic viewpoint. Using the representation formula (1.2.3), an L^1 estimate such as (2.1.24) writes

$$\int \mathbb{E}(u_0(\mathbf{X}_t^x))\,dx \leq C \int u_0(x)\,dx$$

for some time-dependent constant C, that is,

$$\mathbb{E}\Big(\int u_0(\mathbf{X}_t^x)\,dx\Big) \leq C \int u_0(x)\,dx, \tag{2.1.25}$$

which exactly means, on average in probability, having a control on the deformation of the Lebesgue measure, a concept the importance of which we have repeatedly mentioned. Equation (2.1.25) may be expressed in terms of the fundamental solution p to the parabolic equation as

$$\iint p(t,x,y)\mathbb{1}_A(y)\,dy\,dx \le C|A|$$

for all measurable sets A, that is,

$$\int p(t,x,y)\,dx \le C. \tag{2.1.26}$$

We notice that this estimate does not yield any information on the regularity of the fundamental solution. It is a relatively weak information, much weaker indeed than the information classically obtained in the probability setting by *fixing* x, and studying $\mathbb{E}_x(u_0(\mathbf{X}_t^x))$ for $u_0 \in L^1$. For a nondegenerate stochastic differential equation with regular coefficients, an estimate on p, *for x fixed*, is then obtained. Typically, it reads $\mathbb{E}_x(u_0(\mathbf{X}_t^x))) \le C\int u_0$, that is, $\int p(t,x,y)\mathbb{1}_A(y)\,dy \le C|A|$, thus, concentrating the set A to a point,

$$p(t,x,y) \le C, \tag{2.1.27}$$

almost everywhere in x and y. Property (2.1.27) is a very strong property as compared to (2.1.26). As we have already mentioned, the consideration of the very particular case ($\mathbf{a} = 0$, $\mathbf{b} = 0$), for which $p(t,x,y) = \delta(x-y)$, shows that, in full generality, only the weaker of these two properties can be expected: (2.1.26) holds since $\int \delta(x-y)\,dx = 1$, but clearly (2.1.27) does not. Put differently, our viewpoint consists in studying the property of transport of the Lebesgue measure, in sharp contrast to the probability viewpoint which estimates the property of diffusion of the process.

Fifthly, we would like to stress the "almost optimality" of the assumption $\mathbf{b} \in L^{d,1}$ in the following sense. We first show that the assumption $\mathbf{b} \in L^{d,\infty}$ is not sufficient for the L^1 estimate (2.1.24) to hold and next we indicate why $L^{d,p}$, for $1 < p < +\infty$ is not expected to be sufficient either. To show the former fact, we again consider equation (2.1.20):

$$\partial_t u + \lambda\frac{x}{\|x\|^2}\,.\nabla u - \frac{1}{2}\Delta u = 0,$$

where we note that, whatever λ, $\mathbf{b} = -\lambda\frac{x}{\|x\|^2}$ belongs to $L^{d,\infty}$. We consider this equation in the whole space and supply it with the initial condition

$$u(t=0,\cdot) = (2\pi\delta)^{-\frac{d}{2}}e^{-\frac{|x|^2}{2\delta}} \tag{2.1.28}$$

for some presumably small parameter $\delta > 0$. Notice that, when δ vanishes, the initial condition (2.1.28) converges to the Dirac mass at zero. With this particular initial condition, we may solve (2.1.20) explicitly. To this end, we simply write (2.1.20) in the radially symmetric form

$$\partial_t u - \frac{1}{2}u'' + \left(\lambda - \frac{d}{2} + \frac{1}{2}\right)\frac{u'}{r} = 0$$

(which is actually a radially symmetric heat equation in "dimension" $d-2\lambda$), seek its solution under the form $\beta(t)e^{-\frac{|x|^2}{2\alpha(t)}}$, identify α, β and eventually obtain

$$u(t,x) = (2\pi)^{-\frac{d}{2}}\delta^{-\lambda}(t+\delta)^{-\frac{d}{2}+\lambda}e^{-\frac{|x|^2}{2(t+\delta)}}.$$

We then have $\int u(t,\cdot) \propto \delta^{-\lambda}(t+\delta)^{\lambda}$. We now choose $\lambda > 0$, say $\lambda = 1$, so that, when $\delta \to 0$ and t is small (say $t = \sqrt{\delta}$), $\int u(t,\cdot)$ takes arbitrarily large values. This contradicts the L^1 estimate (2.1.24).

To further understand the optimality, we also indicate the following argument, which addresses a case when $\mathbf{b} \notin L^{d,1}$ but $\mathbf{b}$ belongs to spaces that are included in $L^{d,\infty}$. As already done above, we look at the elliptic variant of our problem. The proof of Lemma 2 below proceeds upon considering the adjoint equation (see (2.1.30) below) and proving its solution satisfies an L^∞ estimate. In the case $\mathbf{b} \in L^{d,p}$, for all $1 < p < +\infty$, we now show that the static version of this adjoint equation (2.1.30), namely

$$-\operatorname{div}(v\,\mathbf{b}) - \frac{1}{2}\Delta v = 0, \tag{2.1.29}$$

does not have a bounded solution. This is an indication toward a contradiction to (2.1.24) (although we acknowledge that, *strictly speaking*, this does not *show* (2.1.24) does not hold) for $\mathbf{b} \in L^{d,p}$, $1 < p < +\infty$. Indeed, consider $v = (-\ln|x|)^{\gamma}$, for $0 < \gamma < 1$, and

$$\mathbf{b} = -\frac{1}{2}\nabla \ln v = \frac{1}{2}\gamma\frac{1}{|x|(-\ln|x|)}\frac{x}{|x|}.$$

We observe that

$$\int_0^\infty \frac{1}{t}\left[t^{\frac{1}{d}}\frac{1}{t^{\frac{1}{d}}(-\ln t)}\right]^p dt = \int_0^\infty \frac{1}{t(-\ln t)^p}\,dt < +\infty$$

for all $1 < p < +\infty$. Given the definition (2.0.4) of the norm in the Lorentz spaces and the fact that $\mathbf{b}$ is radially symmetric and non-increasing, this shows that $\mathbf{b} \in L^{d,p}$ for all $1 < p < \infty$ at the vicinity of the origin. On the other hand, v solves (2.1.29) and is not bounded.

We now turn to:

Proof of Lemma 2. We do not know a direct proof of (2.1.24). We therefore proceed by duality. Assuming for simplicity that $a_{ij} = \frac{1}{2}\delta_{ij}$, we introduce, for t a fixed time, the adjoint equation to (1)

$$\partial_s v - \operatorname{div}(v\mathbf{b}(t-s)) - \frac{1}{2}\Delta v = 0, \tag{2.1.30}$$

posed for $s \geq 0$ and which we supply with an arbitrary initial condition $v(s=0,\cdot) = w$ in L^∞. Formally at least (we have already used that argument in (1.1.14), and it may be made rigorous using regularization), we have

$$\int u(t,\cdot)w - \int u_0 v(t,\cdot) = \int_0^t \frac{\partial}{\partial s}\int u(s,\cdot)v(t-s,\cdot)\,ds = 0.$$

It follows that

$$\begin{aligned}\|u(t,\cdot)\|_{L^1} &= \sup_{w\in L^\infty}\frac{\left|\int u(t,\cdot)w\right|}{\|w\|_{L^\infty}}\\ &= \sup_{w\in L^\infty}\frac{\left|\int u_0 v(t,\cdot)\right|}{\|w\|_{L^\infty}}\\ &\le \|u_0\|_{L^1}\sup_{w\in L^\infty}\frac{\|v(t,\cdot)\|_{L^\infty}}{\|w\|_{L^\infty}}\end{aligned}$$

and therefore estimate (2.1.24) on $\|u(t,\cdot)\|_{L^1}$ may be established proving the dual estimate

$$\|v(t,\cdot)\|_{L^\infty} \le C_1 e^{C_0\,t}\|v(t=0,\cdot)\|_{L^\infty} \tag{2.1.31}$$

for the solution v to (2.1.30). The remainder of the proof aims at proving (2.1.31). We write (2.1.30) under the form

$$\partial_s v - \frac{1}{2}\Delta v = \operatorname{div}(v\,\mathbf{b}(t-s)) \tag{2.1.32}$$

and first argue in the case when (2.1.30) is posed on the entire space with zero boundary condition at infinity. In a final step, we will adapt our argument to the case of a bounded domain with periodic boundary conditions. We therefore consider the solution v to (2.1.30) that writes, using the Duhamel formula and a simple integration by part,

$$v(t,x) = w \star p(t,x) - \int_0^t\int \nabla p(t-s,x-y)v(s,y)\mathbf{b}(t-s,y)\,dy\,ds, \tag{2.1.33}$$

where

$$p(t,x) = \frac{1}{(2\pi t)^{\frac{d}{2}}}e^{-\frac{|x|^2}{2t}} \tag{2.1.34}$$

denotes the heat kernel associated to the left-hand side of (2.1.32). Our aim is to prove that, for t sufficiently small, (2.1.33) implies (2.1.31). Actually, we will only be able to prove this fact under the additional assumption that $\mathbf{b}$ is independent of time, as was announced in the statement of Lemma 2. Evidently, the first term of (2.1.33) is estimated using

$$\|w\star p(t,x)\|_{L^\infty([0,t],L^\infty_x)} \le \|w\|_{L^\infty_x}. \tag{2.1.35}$$

The rightmost term of (2.1.33) is more difficult to address. We write

$$\begin{aligned}&\left|\int_0^t\int \nabla p(t-s,x-y)v(s,y)\mathbf{b}(t-s,y)\,dy\,ds\right|\\ &\quad\le \|v\|_{L^\infty([0,t],L^\infty_x)}\int_0^t\int |\nabla p(t-s,x-y)||\mathbf{b}(t-s,y)|\,dy\,ds\\ &\quad\le \|v\|_{L^\infty([0,t],L^\infty_x)}\left\|\left(\int_0^t|\nabla p(t-s,\cdot)|\,ds\right)\star|\mathbf{b}|\right\|_{L^\infty},\end{aligned} \tag{2.1.36}$$

where, in the latter estimate, we have explicitly used that $\mathbf{b}$ does not depend on time. We now observe that, because of (2.1.34),

$$\nabla p(s,x) \quad \text{behaves as} \quad \frac{|x|e^{-\frac{|x|^2}{2s}}}{s^{\frac{d}{2}+1}}, \tag{2.1.37}$$

thus, up to irrelevant multiplicative constants and after an easy change of variable,

$$\int_0^t |\nabla p(s,x)|\, ds \propto \frac{1}{|x|^{d-1}} \int_{|x|^2/2t}^{+\infty} e^{-\sigma}\sigma^{\frac{d}{2}-1}\, d\sigma. \tag{2.1.38}$$

It follows that $\int_0^t |\nabla p(s,\cdot)|\, ds \in L^{\frac{d}{d-1},\infty}$ with a norm that is bounded from above uniformly in t. However, that norm is not necessarily strictly smaller than 1, so noticing

$$\left\| \left(\int_0^t |\nabla p(t-s,\cdot)|\, ds \right) \star |\mathbf{b}| \right\|_{L^\infty} \leq \left\| \int_0^t |\nabla p(s,\cdot)|\, ds \right\|_{L^{\frac{d}{d-1},\infty}} \|\mathbf{b}\|_{L^{d,1}}$$

and inserting this into (2.1.36) is a priori not enough to infer an L^∞ bound from (2.1.33). We therefore proceed as follows. We note that, in addition to belonging to $L^{\frac{d}{d-1},\infty}$, the right-hand side of (2.1.38) vanishes, for all $x \neq 0$ fixed, as $t \to 0$. It follows that the function $\int_0^t |\nabla p(s,\cdot)|\, ds$ goes to zero, as $t \to 0$, in all the spaces $L^{\alpha,\infty}$ for $1 < \alpha < \frac{d}{d-1}$. We next approximate $\mathbf{b} \in L^{d,1}$ by some $b_\delta \in L^{\beta,1}$ so that $\beta > d$, $\alpha^{-1} + \beta^{-1} = 1$, $b - b_\delta$ is smaller in $L^{d,1}$ than the arbitrarily small parameter δ and we write

$$\begin{aligned} \left\| \left(\int_0^t |\nabla p(t-s,\cdot)|\, ds \right) \star |\mathbf{b}| \right\|_{L^\infty} &\leq \left\| \int_0^t |\nabla p(s,\cdot)|\, ds \right\|_{L^{d/d-1,\infty}} \|\mathbf{b} - \mathbf{b}_\delta\|_{L^{d,1}} \\ &\quad + \left\| \int_0^t |\nabla p(s,\cdot)|\, ds \right\|_{L^{\alpha,\infty}} \|\mathbf{b}_\delta\|_{L^{\beta,1}}. \end{aligned} \tag{2.1.39}$$

Upon choosing δ sufficiently small, this L^∞ norm can therefore be taken arbitrarily small (and thus smaller than one) when $t \to 0$. We next insert (2.1.35), (2.1.36), (2.1.39) into (2.1.33), and obtain

$$\|v\|_{L^\infty([0,t],L^\infty_x)} \leq \|w\|_{L^\infty_x} + c_t \|v\|_{L^\infty([0,t],L^\infty_x)}$$

for a constant c_t that for t sufficiently small may be taken smaller than one. This shows estimate (2.1.31) for small times, and, consequently, for all times.

Concluding the proof now amounts to adapting the above argument to the case of a bounded domain with periodic boundary conditions. For this purpose, we simply remark that the heat kernel for that domain, say without loss of generality a cubic unit cell $[0,1[^d$, reads

$$p_{\text{per}} = \sum_{k\in\mathbb{Z}^d} p(\cdot + k),$$

where p is the heat kernel on the whole space given by (2.1.34). The behavior (2.1.38), which is the crucial ingredient of the above proof, actually implies

$$\int_0^t |\nabla p_{\mathrm{per}}(s,\cdot)|\, ds \propto \sum_{k\in\mathbb{Z}^d} \frac{1}{|x+k|^{d-1}} \int_{\frac{|x+k|^2}{2t}}^{+\infty} e^{-\sigma}\sigma^{\frac{d}{2}-1}\, d\sigma, \tag{2.1.40}$$

where the terms of the series are exponentially decaying in $|k|$, and where we note that only the term $k = 0$ is singular in the domain $[0,1[^d$. That term is amenable to the techniques we have just developed above for the case of the problem set on the entire space. The proof can be straightforwardly adapted from then on. □

Remark 4. It is immediate to realize, on (2.1.36) and the subsequent estimations, that the proof readily extends to a *time-dependent* $\mathbf{b}$ in the space $L_x^{d,1}(L_t^\infty)$. Notice that, unfortunately, it does not extend to the space $L_t^\infty(L_x^{d,1})$, so that a genuine, new argument must be found to cover the "general" time-dependent setting.

2.1.4 The two-dimensional setting

Our Theorem 1 is restricted to dimensions higher than or equal to 3. This owes to the fact we have repeatedly employed (2.0.13), and actually also the slightly weaker, classical inequality $\|u\|_{L^{2d/(d-2)}} \le c\|\nabla u\|_{L^2}$. It is well known that in dimension 2, we do not have the embedding $H^1 \subset L^\infty$, but that we only "almost" have it. More precisely, using Trudinger's Theorem [1, Theorem 8.27, p. 277], we know that

$$\|u^2\|_{L_{\exp}} \le C\|u\|_{H^1}^2, \tag{2.1.41}$$

where $L_{\exp}$ denotes the Orlicz space of exponentially integrable functions (see, e.g., [1, Chapter 8] for the definitions of Orlicz spaces and the properties of those spaces which we will use below; see also [86]). It follows, using the generalized Hölder inequality for Orlicz spaces, that

$$\int |f\,.\,u^2| \le C\|f\|_{L\,\mathrm{Log}\,L}\|u^2\|_{L_{\exp}} \le C\|f\|_{L\,\mathrm{Log}\,L}\|u\|_{H^1}^2, \tag{2.1.42}$$

where $L\,\mathrm{Log}\,L$ denotes the space of functions f such that $|f|\,\mathrm{Log}_+|f| \in L^1$. We will now use these properties to adapt our result, Theorem 1, to the two-dimensional setting.

For the two-dimensional version of the result, all the arguments that do not resort to Sobolev embeddings are not affected. This is in particular the case of the regularization step. Our assumption (2.1.1), namely

$$\mathbf{b} = \beta_1 + \beta_2, \quad \beta_1 \in L^2,\ \beta_2 \in W^{1,1},$$

is still the appropriate one, and the proof we have performed above carries over to the two-dimensional setting. On the other hand, the step where a priori estimates are established, for which Sobolev embeddings are key ingredients, need to be adapted.

The crucial estimate is that of the term $\int \mathbf{b} . \nabla u . u$, which, mimicking our arguments following (2.1.6)–(2.1.7) of the proof of Theorem 1, we split as follows:

$$\begin{aligned}\left| \int \mathbf{b} . \nabla u . u \right| &\leq \left| \int \mathbf{b}_1 . \nabla u . u \right| + \left| \frac{1}{2} \int (-\operatorname{div} \mathbf{b}_2) u^2 \right| \\ &\leq \frac{1}{8\underline{a}} \int |\nabla u|^2 + 32\underline{a} \int |\mathbf{b}_1|^2 |u|^2 + |\frac{1}{2} \int (-\operatorname{div} \mathbf{b}_2) u^2|. \end{aligned} \tag{2.1.43}$$

We now notice that, when $\mathbf{b}_1 = \mathbf{b}_{11} + \mathbf{b}_{1\infty}$, with $\mathbf{b}_{11} \in L^2 \operatorname{Log} L$ and $\mathbf{b}_{1\infty} \in L^\infty$, we may, as in (2.1.8), choose $\delta > 0$ sufficiently small so that

$$\mathbf{b}_{11} - \mathbf{b}_{11,\delta} \in L^\infty \quad \text{and} \quad \|\mathbf{b}_{11,\delta}\|_{L^2 \operatorname{Log} L} \leq \delta, \tag{2.1.44}$$

where, of course, $L^2 \operatorname{Log} L$ is the space of functions f such that $|f|^2 \operatorname{Log}_+ |f| \in L^1$. This allows to estimate

$$\begin{aligned}\int |\mathbf{b}_1|^2 |u|^2 &\leq 4\big(\|\mathbf{b}_{11} - \mathbf{b}_{11,\delta}\|^2_{L^\infty} + \|\mathbf{b}_{1\infty}\|^2_{L^\infty}\big) \int |u|^2 + C \||\mathbf{b}_{11,\delta}|^2\|_{L \operatorname{Log} L} \|u\|^2_{H^1} \\ &\leq 4\big(\|\mathbf{b}_{11} - \mathbf{b}_{11,\delta}\|^2_{L^\infty} + \|\mathbf{b}_{1\infty}\|^2_{L^\infty}\big) \int |u|^2 + C\delta^2 \|u\|^2_{H^1} \\ &\leq 4\big(\|\mathbf{b}_{11} - \mathbf{b}_{11,\delta}\|^2_{L^\infty} + \|\mathbf{b}_{1\infty}\|^2_{L^\infty} + 1\big) \int |u|^2 + C\delta^2 \|\nabla u\|^2_{L^2}, \end{aligned}$$

where we have used (2.1.42), (2.1.44) and that δ is small. Choosing δ even smaller, we may thus have

$$32\underline{a} \int |\mathbf{b}_1|^2 |u|^2 \leq \frac{1}{8\underline{a}} \int |\nabla u|^2 + C \int |u|^2. \tag{2.1.45}$$

Similarly, assuming that $[\operatorname{div} \mathbf{b}_2]_- = f_{21} + f_{2\infty}$, with $f_{21} \in L \operatorname{Log} L$ and $f_{2\infty} \in L^\infty$ and writing $f_{11} = (f_{11} - f_{11,\delta}) + f_{11,\delta}$ with, on the one hand, $f_{11} - f_{11,\delta} \in L^\infty$ and, on the other hand, $\|f_{11,\delta}\|_{L \operatorname{Log} L} \leq \delta$, we obtain

$$\left| \frac{1}{2} \int (-\operatorname{div} \mathbf{b}_2) u^2 \right| \leq \frac{1}{8\underline{a}} \int |\nabla u|^2 + C \int |u|^2. \tag{2.1.46}$$

Inserting (2.1.45) and (2.1.46) into (2.1.43) yields

$$\left| \int \mathbf{b} . \nabla u . u \right| \leq \frac{3}{8\underline{a}} \int |\nabla u|^2 + C \int |u|^2.$$

It is then straightforward to obtain the L^2 estimate (2.1.16). The proof then easily proceeds.

The above outlined argument clearly shows that a simple adaptation of the proof of Theorem 1 yields the following.

Theorem 2. *Assume that the ambient dimension is $d = 2$. Then, under the assumptions of Theorem 1* except *that* (2.1.2) *is replaced by*

$$\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2, \quad \mathbf{b}_1 \in L^2 \operatorname{Log} L + L^\infty, \; [\operatorname{div} \mathbf{b}_2]_- \in L \operatorname{Log} L + L^\infty, \tag{2.1.47}$$

the conclusions of Theorem 1 hold true.

Following the statement of Theorem 1, we have made some remarks regarding the optimality of our assumptions (2.1.2). In particular, we have considered in (2.1.19) some $\mathbf{b} = -\lambda \frac{x}{\|x\|^2}$ that belongs to $L^{d,\infty}$. It turns out that the calculations we have performed then are still valid in dimension $d = 2$. When $d = 2$, these calculations show that, whatever the value of λ, uniqueness cannot hold on the stochastic differential equation (2.1.21) and on the elliptic variant (2.1.23) of our parabolic equation. Intuitively, Theorem 2 is thus unlikely to hold true when $\mathbf{b} = \mathbf{b}_1 \in L^{2,\infty}$.

Actually, considering $\mathbf{b} = \lambda \frac{x}{\|x\|^2} \ln |x|$ (for $\lambda > 0$), suggests that Theorem 2 cannot hold true for $\mathbf{b} = \mathbf{b}_1 \in L^{2,p}$ for any $p > 1$. Indeed, a similar argument to that we performed for (2.1.20) shows that $\mathbf{X}_t$ solution to

$$d\mathbf{X}_t = \lambda \frac{\mathbf{X}_t}{\|\mathbf{X}_t\|^2} \ln \|\mathbf{X}_t\| \, dt + d\mathbf{W}_t$$

satisfies

$$d(\ln \|\mathbf{X}_t\|) = \lambda \frac{1}{\|\mathbf{X}_t\|^2} \ln \|\mathbf{X}_t\| \, dt + \frac{\mathbf{X}_t}{\|\mathbf{X}_t\|^2} \, d\mathbf{W}_t.$$

We again observe a concentration of the trajectories around the origin, thus a possible lack of uniqueness.

Similarly, we have the following lemma, analogous to Lemma 2 above:

Lemma 3. *In the two-dimensional setting, under the assumptions of Theorem 2 and the additional assumptions* $\mathbf{b} \in L^{2,1}$ *and* $\mathbf{b}$ *is independent of time, the* L^1 *estimate* (2.1.24) *of Lemma 2 holds true.*

The proof of Lemma 3 directly derives from that of Lemma 2, performed now under the assumptions of Theorem 2. Indeed, our arguments in the proof of Lemma 2 are valid in all dimensions $d \geq 2$.

Similarly to the remarks we made after the statement of Lemma 2, we notice that $L^{2,1}$ seems critical for Lemma 3 to hold. Indeed, our counterexample

$$\mathbf{b} = -\lambda \frac{x}{\|x\|^2} \in L^{2,\infty}$$

shows on the parabolic equation (2.1.20) that Lemma 3 does not hold in that case. Likewise,

$$\mathbf{b} = -\frac{\theta}{\|x\| |\operatorname{Log} \|x\||} \frac{x}{\|x\|} \in L^{2,p}$$

(for all $p > 1$) which we have considered there, shows that the solution to the elliptic variant (2.1.29) is then unbounded, a strong indication toward the optimality of the space $L^{2,1}$.

Interestingly, the proof of Lemma 3 actually provides another uniqueness result similar to Theorem 2, this time when $\mathbf{b}_1 \in L^{2,1}$. That particular observation is irrelevant in dimensions $d \geq 3$ since $\mathbf{b}_1 \in L^{d,1}$ (assumption of Lemma 2) is then contained in our assumption $\mathbf{b}_1 \in \varepsilon_1 L^{d,\infty} + L^d$ of (2.1.2) in Theorem 1, so the uniqueness is already established then. In dimension 2, $\mathbf{b} \in L^{2,1}$ does allow for the regularization procedure, but it does not allow for our a priori estimate to hold true, because the term $\int \mathbf{b} \, . \, \nabla u \, u$

cannot be controlled using only $\|\nabla u\|_{L^2}$ and $\|u\|_{L^2}$. So we have to proceed otherwise, actually using the proof of Lemma 3. We nevertheless have the following.

Lemma 4. *In the two-dimensional setting, under the assumptions on* $\mathbf{a}$ *and* u_0 *of Theorem 2, and the assumptions* $\mathbf{b} \in L^{2,1}$ *and* $\mathbf{b}$ *is independent of time, we have existence and uniqueness of the solution to* (1) *in* $L^2([0,T],H^1) \cap C([0,T],L^p) \cap L^\infty([0,T],L^\infty)$ *for all* $1 \le p < +\infty$. *The solution satisfies in addition the* L^p *estimate* (2.1.3) *for all* $1 < p < +\infty$, *the* L^∞ *estimate* (2.1.4) *and the* L^1 *estimate* (2.1.24).

Proof of Lemma 4 (outline). We approximate $\mathbf{b} \in L^{2,1}$ by some smooth $\mathbf{b}_\varepsilon$. For the latter transport field, we have a regular solution u_ε. For that regular solution, we have the maximum principle (2.1.4) by standard considerations and the L^1 estimate (2.1.24) with constants (C_0, C_1) uniform in ε, using the proof of Lemma 3. By the Hölder inequality, this shows an estimate in L^p analogous to (2.1.24), again uniformly in ε. Passing to the limit, this proves existence of a solution u satisfying (2.1.4), the L^1 estimate (2.1.24), and the corresponding L^p estimates for all p. In addition, since for all ε, we have

$$\int |u_\varepsilon|^2 + \frac{1}{2}\underline{a}\int_0^T\int |\nabla u_\varepsilon|^2 \le \frac{8}{\underline{a}}T\|\mathbf{b}_\varepsilon\|_{L^2}^2\|u(t=0,\cdot)\|_{L^\infty}^2 + \int |u(t=0,\cdot)|^2,$$

we obtain a similar estimate on the limit u. This shows the solution u constructed belongs to $L^2([0,T],H^1) \cap C([0,T],L^p) \cap L^\infty([0,T],L^\infty)$. As for uniqueness, we subtract the equation satisfied by the two tentative solutions u_1 and u_2, obtain an equation on $u = u_1 - u_2$ with zero initial condition. Because $\mathbf{b} \in L^{2,1}$, we may apply *the proof of* Lemma 3, which yields the L^1 estimate and thus, since $u_0 = 0$, $u = 0$. Uniqueness follows and this concludes the proof. □

When a is only nonnegative. We briefly mention here a straightforward adaptation of Theorem 1 in the case when the constant symmetric matrix $\mathbf{a}$ is not definite positive but only *nonnegative*:

Corollary 1. *As in Theorem 1, we assume that the ambient dimension d is larger than or equal to* 3, *and that the matrix coefficient* $\mathbf{a}$ *in* (1) *is a* constant symmetric *matrix. We however only assume that* $\mathbf{a}$ *is* nonnegative, *and no longer necessarily definite positive. We denote by* $V = \operatorname{Im}\mathbf{a}$ *the image space of the matrix* $\mathbf{a}$, *by* P_V *the orthogonal projector onto that space, by* $V^\perp = \operatorname{Ker}\mathbf{a}$ *its kernel, and by* $r = \dim V$ *its rank. We assume* $r \ge 3$ *(see the comments underneath the statement of this corollary for the other cases). Again as in Theorem 1, we supply equation* (1) *with a bounded initial condition* $u_0 \in L^\infty$. *We assume, instead of* (2.1.1) *and* (2.1.2), *that we respectively have*

$$\mathbf{b} = \beta_1 + \beta_2, \quad \beta_1 = P_V\beta_1 \in L^2,\ \beta_2 \in W^{1,1}, \tag{2.1.48}$$

$$\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2, \quad \begin{cases} \mathbf{b}_1 = P_V\mathbf{b}_1 \in L^\infty(V^\perp, \varepsilon_1 L^{r,\infty}(V) + L^r(V)), \\ [\operatorname{div}_V P_V\mathbf{b}_2]_- \in L^\infty(V^\perp, \varepsilon_2 L^{\frac{r}{2},\infty}(V) + L^{\frac{r}{2}}(V)), \\ [\operatorname{div}_{V^\perp} P_{V^\perp}\mathbf{b}_2]_- \in L^\infty, \end{cases} \tag{2.1.49}$$

for some constants $\varepsilon_1, \varepsilon_2 > 0$. *Then there exists a unique solution u to* (1) *in the functional space* $C([0,T], L^p) \cap L^\infty([0,T], L^\infty)$, *for all* $1 \le p < +\infty$, *with* $P_V \nabla u \in L^2([0,T], L^2)$. *In addition, estimates* (2.1.3), (2.1.4) *and, under the appropriate assumptions, the* L^1 *estimate* (2.1.24) *of Lemma 2 hold.*

The proof of this corollary follows the exact same pattern as those of Theorem 1 and Lemma 2. In the subspace V where **a** acts as a positive definite matrix, we obtain integrability of $P_V \nabla u$ and use it. In the orthogonal subspace $V^\perp$, the equation agrees with a linear transport equation and is therefore treated correspondingly. For brevity, we omit the adaptation of Corollary 1 to the two-dimensional setting, an adaptation that can be performed similarly to the adaptation we performed above for Theorem 1. Likewise, in any ambient dimension d when the rank is $r = 1$ or 2, the result and its proof have to be amended, but there is no difficulty.

2.2 Probability theoretic viewpoint

We devote this section to the interpretation and the consequences of our results in the probabilistic setting, and more precisely for the theory of stochastic differential equations. Of course, the connection with such questions is not restricted to the case considered in the present section of *constant* second-order differential operators. Many actual interesting phenomena and issues in fact arise for non-constant coefficients operators, that is, the case considered for the partial differential equation from Chapter 3 on. Nevertheless, we find it useful to already explore the connections with stochastic analysis in more details. We will even allow ourselves to slightly anticipate on the results on varying coefficients, in order to lay some groundwork.

To start with, let us mention that, in essence, the content of this Section is to construct a functional setting, together with the suitable notions, and to introduce the necessary techniques, so that all the formal manipulations establishing the correspondence between the partial differential equations viewpoint and the stochastic differential viewpoint and briefly recalled in Section 1.2 make sense beyond the classical regular setting.

We also recall our discussion from page 21 regarding the adaptations to the context of stochastic processes of the *periodic boundary conditions* we adopted throughout the sections addressing the partial differential equations. For simplicity, we will work in this section (and likewise in Section 4.2) *as if* all the necessary conditions (on the drift **b**, the diffusion matrix $\mathbf{a} = \sigma\sigma^T$, the law u_0 of the initial condition) were fulfilled such that our results on the parabolic equations associated to the stochastic differential equations considered hold on the whole space $\mathbb{R}^d$. An alternate option is to work with processes valued in the torus. Then, there are some technicalities arising in our arguments, due to the specific geometry of the torus, which we then omit for brevity.

2.2.1 Generalities and notion of solution

Generalities. As was briefly mentioned in Section 1.2, equation (1), say in the simple form (1.3.2)

$$\partial_t u - \mathbf{b}\,.\,\nabla u - \frac{1}{2}\Delta u = 0,$$

with initial condition u_0, is formally connected to the stochastic differential equation

$$d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)\,dt + d\mathbf{W}_t. \tag{2.2.1}$$

In this case of a constant second-order operator, (2.2.1) is essentially an *ordinary* differential equation parameterized by the Brownian motion, see below and Section 1.3.2, page 20. The connection is performed by the Feynman–Kac formula (1.2.3):

$$u(t,x) = \mathbb{E}_x(u_0(\mathbf{X}_t)).$$

An analogous link exists with the equation adjoint to (1.3.2), namely (1.2.4):

$$\partial_t u + \operatorname{div}(\mathbf{b}\,u) - \frac{1}{2}\Delta u = 0,$$

the solution of which is the law of the process $\mathbf{X}_t$. The consideration of these partial differential equations allows to establish results (existence, uniqueness, etc.) on the stochastic differential equation that hold on average, that is, in law. One may also consider, in addition, these two parabolic equations in a space of augmented (actually doubled) dimension. This gives rise to equations (1.2.8) and (1.2.11), respectively, and provides a useful technical tool to prove results that hold pathwise on the stochastic differential equation. All the formal connections we have just mentioned hold in a rigorous sense when all these equations have regular data. As already mentioned, one of the purposes of the present contribution is to explore these connections in a less regular setting.

We have already partially completed the program for the specific case of equation (2.2.1) in [64]. Our results have been outlined in the last paragraph of Section 1.3.2 above. The questions left open in [64] have also been recalled there.

Our technique of proof in [64] was bound to the fact that σ is constant in (2.2.1). We have indeed recalled in Section 1.3.2 that we simply write (2.2.1) in the form $d(\mathbf{X}_t - \mathbf{W}_t) = \mathbf{b}((\mathbf{X}_t - \mathbf{W}_t) + \mathbf{W}_t)\,dt$ and, for each Brownian trajectory, use the deterministic theory of [41] for the latter equation: we used the definition, and next the construction of the flow of an *ordinary* differential equation, performed in [41], to obtain that of equation (2.2.1). We are now going to develop some new elements of a more general theory. For simplicity, we will again expose them in the case of a constant σ (which is the setting of a second-order operator with constant coefficients considered throughout the present Chapter 2). In sharp contrast with our arguments of [64] though, we construct a notion of solution and flow *directly* on equation (2.2.1). Our techniques and results of this section readily carry over to the equation

$$d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)\,dt + \sigma(\mathbf{X}_t)\,d\mathbf{W}_t, \tag{2.2.2}$$

with a *varying* σ provided it is *regular*. Depending on whether or not the second-order operator is positive definite, the assumptions on $\mathbf{b}$ might slightly vary, but in essence they are similar to the case treated here. On the other hand, in order for our arguments and results to be transferable to equation (2.2.2) for a Sobolev regular coefficient σ, adaptations are necessary. Those will be the matter of Section 4.2. As usual, explicit dependency upon time of $\mathbf{b}$ and σ can also be accommodated.

Definition of a notion of solution. To start with, we retain from the approach of [41] that the construction of the flow for the differential equation derives from that of the solution to the underlying partial differential equation. We thus begin with considering the expression $u(t,x) = \mathbb{E}_x(u_0(\mathbf{X}_t))$. We wish to understand what information is, and what information is not, provided by the solution to the parabolic equation (1.3.2).

We note that the law of the process $\mathbf{X}_t$ is not entirely determined by the parabolic equation. We observe that
(i) the partial differential equation determines the law of $\mathbf{X}_t$ at all times t,
(ii) if we impose a semi-group (i.e. Markov) property on the flow, as is performed in [41] for the deterministic flow, then, for any N-tuple of times $t_1 \leq t_2 \cdots \leq t_N$, the law of $(\mathbf{X}_{t_1}, \mathbf{X}_{t_2}, \ldots, \mathbf{X}_{t_N})$ is then also determined,
(iii) if we additionally impose that the trajectories of the process $\mathbf{X}_t$ are *continuous*, then (ii) in turn determines the law of the process $\mathbf{X}_t$.

It follows from this series of observations that, besides the evident semi-group property to impose on the flow (it is indeed a *necessary* property if we want existence and uniqueness of that flow), we have to impose conditions (on $\mathbf{b}$, and, possibly, σ) so that we expect the trajectories of $\mathbf{X}_t$ to be continuous in time. This continuity is not immediate since, for instance, $\int_0^t \mathbf{b}(\mathbf{X}_s)\,ds$ is a Lipschitz continuous, thus continuous function of time t when $\mathbf{b}$ is bounded, but the functions $\mathbf{b}$ we manipulate are not necessarily bounded. In addition, and as already mentioned earlier, even the meaning of $\mathbf{b}(\mathbf{X}_s)$ is unclear, for a non-regular $\mathbf{b}$.

We claim that the following condition ensures that our solution will have, for almost all initial conditions x, continuous trajectories: there exists some constant c such that

$$\int \mathbb{E}_x(\mathbb{1}_A(\mathbf{X}_t))\,dx \leq c|A| \tag{2.2.3}$$

for all times $t \geq 0$ and all Borel sets A, or, equivalently,

$$\int \mathbb{E}_x(|f(\mathbf{X}_t)|)\,dx \leq c\int |f| \tag{2.2.4}$$

for all times $t \geq 0$ and for all $f \in L^1$. In other words, the process $\mathbf{X}_t$, seen as a probability measure, will only weight continuous trajectories. We now expose why the claim holds true. Our arguments below are formal and will be made rigorous for regularized solutions in the proof of existence of a solution which we will perform later. Notice also that (2.2.3) is the natural extension to the stochastic setting of the conservation (1.1.5) (for $p = 1$) owing to the control of the inverse Jacobian for ordinary differential equations. Also, it is typically true in the classical, regular setting because, using the Feynman–Kac formula (1.2.3), $\mathbb{E}_x(|f(\mathbf{X}_t)|)$ is actually the solution u to (1.3.2), that is, $\partial_t u - \mathbf{b}\,.\nabla u - \frac{1}{2}\Delta u = 0$ for the initial condition $u_0 = |f|$. Put differently, (2.2.3) controls the evolution of the L^1 norm. Notice that its variant, discussed in Remark 8 below, similarly encodes the controlled evolution of the L^p norm. We may also remark that, formally at least and rigorously in the regular setting again, (2.2.3) reads as an L^∞ bound on the density $p(t,\cdot)$ of $\mathbf{X}_t$ when the Lebesgue measure is taken as initial law. Indeed,

$$\int \mathbb{E}_x(|f(\mathbf{X}_t)|)\,dx = \int p(t,y)|f(y)|\,dy.$$

Before we start, we therefore bear in mind the following. Estimate (2.2.3) is formally equivalent to an L^1 estimate for the solution to the associated partial differential equation, in the spirit of the L^1 estimate (2.1.24): $\mathbb{E}_x(\mathbb{1}_A(\mathbf{X}_t))$ is the solution at time t and point x for the initial condition $\mathbb{1}_A$. This equivalence is true in the classical setting. It will also hold for the regularized solutions we construct, as mentioned above, and therefore, passing to the limit, for the solution we construct.

Intuitively, property (2.2.3) expresses that the possible modifications of the almost everywhere defined field $\mathbf{b}$ on sets of zero Lebesgue measure do no affect the solution *on average*. Put differently, the Lebesgue measure is not significantly perturbed by the stochastic flow we construct, a property the importance of which we have already repeatedly emphasized in the deterministic setting. In some vague sense, (2.2.3) is a version of (1.1.5)–(1.1.7) *averaged* over the Brownian trajectories. We now formalize this. Our solution $\mathbf{X}_t$ can be decomposed as

$$\mathbf{X}_t = \mathbf{M}_t + \mathbf{A}_t \quad \text{with } \mathbf{M}_t = x + \mathbf{W}_t,\ \mathbf{A}_t = \int_0^t \mathbf{b}(\mathbf{X}_s)\,ds,$$

where $\mathbf{M}_t$ is a martingale, obviously continuous in time given the properties of the Brownian motion $\mathbf{W}_t$. Proving the continuity of $\mathbf{X}_t$ therefore amounts to proving that of $\mathbf{A}_t$. Consider $\mathbf{b} \in L^p$ for some $1 \le p < +\infty$. We observe that

$$\int \mathbb{E}_x\Big(\Big|\frac{d}{dt}\mathbf{A}_t\Big|^p\Big)\,dx = \int \mathbb{E}_x(|\mathbf{b}(\mathbf{X}_t)|^p)\,dx \le C\int |\mathbf{b}|^p\,dx < +\infty, \tag{2.2.5}$$

successively using the definition of $\mathbf{A}_t$ and property (2.2.3). It is immediate to deduce from (2.2.5) and a similar estimate on A_t itself that

$$\int \mathbb{E}_x(\|\mathbf{A}_t\|^p_{W^{1,p}})\,dx \le C\int |\mathbf{b}|^p. \tag{2.2.6}$$

This shows that $\mathbf{A}_t$ is continuous, almost everywhere in x and almost surely.

Slightly anticipating on the case (2.2.2) of a varying matrix σ, and in order to demonstrate that, as announced above, our arguments carry over to that case, we briefly mention how to complement the above argument to address the continuity of

$$\mathbf{X}_t = \mathbf{M}_t + \mathbf{A}_t \quad \text{with } \mathbf{M}_t = x + \int_0^t \sigma(\mathbf{X}_s)\,d\mathbf{W}_s,\ \mathbf{A}_t = \int_0^t \mathbf{b}(\mathbf{X}_s)\,ds. \tag{2.2.7}$$

Of course, $\mathbf{A}_t$ is left unchanged thus its continuity is proved as above. On the other hand, the martingale $\mathbf{M}_t$ now needs a specific argument. Assume, e.g., that $\sigma \in L^p$ for some $p > 2$. We remark that

$$\int \mathbb{E}_x([\mathbf{M}_t]^p_{W^{s,p}})\,dx = \int \mathbb{E}_x \int_0^1\int_0^1 \frac{|\mathbf{M}_{t'} - \mathbf{M}_t|^p}{|t'-t|^{1+sp}}\,dt\,dt'\,dx, \tag{2.2.8}$$

by definition of the (Gagliardo) semi-norm $[\,\cdot\,]$ in $W^{s,p}$ (so that $\|f\|^p_{W^{s,p}} = \|f\|^p_{L^p} + [f]^p_{W^{s,p}}$). We will assume henceforth that $s < \frac{1}{2}$ because we cannot expect a better regularity of $\mathbf{M}_t$, given that of the Brownian motion, and we will also assume $\frac{1}{p} < s$ for a reason that will be clear below. We next observe that

$$\mathbb{E}_x(|\mathbf{M}_{t'} - \mathbf{M}_t|^p) = \mathbb{E}_x\left(\left|\int_t^{t'} \sigma(\mathbf{X}_\tau)\,d\mathbf{W}_\tau\right|^p\right) \le C_p \mathbb{E}_x\left(\left(\int_t^{t'} |\sigma(\mathbf{X}_\tau)|^2\,d\tau\right)^{\frac{p}{2}}\right), \tag{2.2.9}$$

using a martingale moment inequality, cf. [58, p. 163], for a universal constant C_p depending only on the exponent p. Next,

$$\mathbb{E}_x\left(\left(\int_t^{t'} |\sigma(\mathbf{X}_\tau)|^2\,d\tau\right)^{\frac{p}{2}}\right) \le |t'-t|^{\frac{p}{2}-1}\mathbb{E}_x\left(\int_t^{t'} |\sigma(\mathbf{X}_\tau)|^p\,d\tau\right),$$

using the Hölder inequality. We now integrate in x to obtain

$$\begin{aligned}
\int \mathbb{E}_x(|\mathbf{M}_{t'} - \mathbf{M}_t|^p)\,dx &\le C_p|t'-t|^{\frac{p}{2}-1} \int \mathbb{E}_x\left(\int_t^{t'} |\sigma(\mathbf{X}_\tau)|^p\,d\tau\right)dx \\
&= C_p|t'-t|^{\frac{p}{2}-1} \int_t^{t'}\left(\int \mathbb{E}_x(|\sigma(\mathbf{X}_\tau)|^p)\,dx\right)d\tau \\
&\le C|t'-t|^{\frac{p}{2}-1} \int_t^{t'}\left(\int |\sigma|^p\,dx\right)d\tau \\
&= C|t'-t|^{\frac{p}{2}} \int |\sigma|^p,
\end{aligned} \tag{2.2.10}$$

using property (2.2.3). We next insert this information into (2.2.8) to obtain

$$\int \mathbb{E}_x([\mathbf{M}_t]^p_{W^{s,p}})\,dx \le C \int_0^1\int_0^1 \frac{1}{|t'-t|^{1+sp-\frac{p}{2}}}\,dt\,dt' \int |\sigma|^p, \tag{2.2.11}$$

where the double integral in time on the right-hand side is finite since $\frac{1}{p} < s < \frac{1}{2}$ implies that $1 - \frac{p}{2} < 1 + sp - \frac{p}{2} < 1$. An easier argument (simply setting $t = 0$ in (2.2.10) and integrating in time) shows a similar conclusion on $\int \mathbb{E}_x(\|\mathbf{M}_t\|^p_{L^p})\,dx$ so that it is true for the whole $W^{s,p}$ norm:

$$\int \mathbb{E}_x(\|\mathbf{M}_t\|^p_{W^{s,p}})\,dx \le C \int |\sigma|^p. \tag{2.2.12}$$

Arguing as above, we deduce from (2.2.11) the continuity of $\mathbf{M}_t$, almost everywhere in x and almost surely. Let us notice that the above arguments are essentially formal, but that they make rigorous sense on regularizations of our problems, and therefore, in the limit, for the notion of solution we will define below. This concludes our discussion on assumption (2.2.3) which guarantees that the trajectories are almost surely continuous in time, for almost all initial conditions.

We are now in a position to define our notion of flow and solution for stochastic differential equations. The rationale behind our construction rests on our observations above, which intuitively show that, once integrated in x, the possible singularities of $\mathbf{b}$ and σ essentially disappear, so that, on average, the properties of the solution are those of a classical solution. Our approach to define a suitable notion of solution is therefore to consider not the solution for each initial condition separately, but to consider altogether the family of solutions for various initial conditions. This is only for this set of solutions, *globally* constructed for almost all x, that uniqueness will hold true. Likewise, when arguing in the weak sense, we do not define a process but a family $(\mathbb{P}_x)_{x\in\mathbb{R}^d}$ of laws on the Wiener space $C([0,T],\mathbb{R}^d)$, measurably depending on the parameter $x \in \mathbb{R}^d$. Note there is no reason to expect more than measurability, that is, e.g., continuity, in x.

Definition 1. We say that $\mathbf{X}_t(x)$ (abbreviated as $\mathbf{X}_t$ when there is no ambiguity) is a (family of) solution(s) to (2.2.1), parameterized by the initial condition x, or a solution flow to (2.2.1), that is,

$$d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)\,dt + \,d\mathbf{W}_t,$$

when

- $\mathbf{X}_t$ satisfies (2.2.3), that is, there exists a constant c such that

  $$\int \mathbb{E}_x(\mathbb{1}_A(\mathbf{X}_t))\,dx \le c|A|$$

  for all times $t \ge 0$ and for all Borel sets A,
- $\mathbf{X}_t$ satisfies

  $$\mathbf{X}_t = x + \mathbf{W}_t + \int_0^t \mathbf{b}(\mathbf{X}_s)\,ds \tag{2.2.13}$$

  for all times $t \ge 0$, almost surely and for almost all initial conditions x,
- $\mathbf{X}_t$ satisfies the semi-group property

  $$\mathbf{X}_{t+s} = \mathbf{X}_t(\mathbf{X}_s)$$

  for all times $s, t \ge 0$ (where our notation assumes that the Brownian motion used to construct $\mathbf{X}_t$ is the one used to construct $\mathbf{X}_{t+s}$ from time s).

Remark 5. The first two properties in Definition 1 imply that

$$\begin{aligned}&\mathbf{X}_t,\ \text{almost surely and for almost all initial conditions } x,\\ &\text{has continuous trajectories in time.}\end{aligned} \tag{2.2.14}$$

The argument has been outlined above, immediately before the statement of Definition 1.

Remark 6. The solution $\mathbf{X}_t$ defined in Definition 1 is of course adapted to the filtration generated by the initial condition x and the Brownian motion $\mathbf{W}_t$. In the probability sense, it is a *strong* solution: it is constructed from a given Brownian motion.

Remark 7. In relation with the last property of Definition 1, we remark that the solution we will construct below satisfies, almost everywhere in x,

$$\mathbb{E}_x(\varphi(\mathbf{X}_{T+s})\psi(\mathbf{X}_{t_1},\ldots,\mathbf{X}_{t_n})) = \mathbb{E}_x(\mathbb{E}_{\mathbf{X}_T}(\varphi(\mathbf{X}_s))\psi(\mathbf{X}_{t_1},\ldots,\mathbf{X}_{t_n})) \tag{2.2.15}$$

for all $n \in \mathbb{N}$, $T > 0$, $s > 0$, $t_k \in [0, T]$, $1 \le k \le n$, and all smooth functions φ and ψ. In that sense, the solution is "almost everywhere in x", a Markov process. Notice that this property is satisfied by the unique solution we will construct below, but we do not claim we may replace one of the properties of Definition 1 by this generalized Markov property and proceed likewise.

Remark 8. For any $1 \le p < +\infty$, a variant of Definition 1 holds with property (2.2.3) replaced by

$$\int \mathbb{E}_x(|f(\mathbf{X}_t)|)\,dx \le c\Big(\int |f|^p\Big)^{\frac{1}{p}} \tag{2.2.16}$$

for all $f \in L^p$. Of course, (2.2.16) for $p = 1$ agrees with (2.2.4). In that case, the estimate is the analogous estimate, in the probability theoretic language, of the L^1 estimate (2.1.24) of Lemma 2. In the case of a more general exponent p, (2.2.16) corresponds to the probability theoretic version of the L^p estimate (2.1.3) of Theorem 1. In the sequel, we proceed with the L^1 setting, which is the natural setting in probability, although this L^1 setting will be more demanding in terms of regularity assumptions on the data. All the below may be adapted to the L^p setting, with $1 < p < +\infty$, in which case more general assumptions on the data may be accommodated, but some technicalities arise in the arguments. The essential reason why the L^p case with $1 < p < +\infty$, is less demanding on the data than the L^1 case, is, we recall (see our comments at the beginning of Section 2.1.3), that regularity/integrability may then be bootstrapped on the solution of the equation using the classical a priori estimates (multiplying the equation, integrating, using the strong convexity of the L^p norm for $p > 1$).

Remark 9. An analogous definition holds for equation (2.2.2). See Section 4.2.

Remark 10. Of course, this definition agrees with the definition we have given in [64] in the special case of equation (2.2.1) using the deterministic solution flow of the associated ordinary differential equation.

Remark 11. Because of condition (2.2.3), Definition 1 allows to make sense of the stochastic equation even if the drift $\mathbf{b}$ has only some local integrability property and therefore, in that sense, formalizes the extension mentioned in our earlier work [64, Section 5.2, p. 128].

We begin with proving uniqueness in law, next turn to the existence of the weak solution. We will then address pathwise uniqueness and existence of strong solutions.

2.2.2 Existence and uniqueness in law

Existence and uniqueness in law of the solution to the stochastic differential equation both rely upon the results we have obtained on parabolic-type equations in the ambient dimension d. In sharp contrast, strong existence and pathwise uniqueness, which will be established in Section 2.2.3, require considering a parabolic equation in an augmented space, of dimension $2d$, as already seen in Section 1.2 for the regular setting.

Uniqueness-in-law of the solution. Consider $(\mathbf{X}_t)_{0\le t\le T}$ a solution to (2.2.1) in the sense of Definition 1. We now prove that this solution is unique in law.

For this purpose, we first recall the following classical argument. If the data were regular, we could consider the solution u to the partial differential equation (1.3.2) and show, using the Itô formula (1.2.2), that the Feynman–Kac formula (1.2.3), $u(t,x) = \mathbb{E}_x(u_0(\mathbf{X}_t))$ holds true, for all initial conditions u_0, thereby showing the uniqueness of the law of $\mathbf{X}_t$ at all times t. Using the semi-group property and the continuity of paths, this would in turn imply uniqueness in law of the process $(\mathbf{X}_{0\le t\le T})$. In the classical, regular setting, the uniqueness in probability is therefore based either on the unique solvability of the Fokker–Planck equation (1.2.4)

$$\partial_t p + \operatorname{div}(p\mathbf{b}) - \frac{1}{2}\partial^2_{ij}(\sigma_{ik}\sigma_{jk}p) = 0,$$

or on the *existence* of a sufficiently regular solution to

$$\partial_t u - \mathbf{b}.\nabla u - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial^2_{ij}u = 0,$$

on which Itô calculus can be performed. The second option is for instance the case for the classical uniqueness statement for the martingale problem, obtained from the solvability of the Kolmogorov equation; see, e.g., [58, Theorem 4.28]. In our non-regular setting, our argument will in some sense be reminiscent of this observation.

Consider the (unique) solution u to (1.3.2) provided by Theorem 1 (say we work in dimension 3, under the assumptions of that theorem, but of course all the other settings considered so far for partial differential equations are equally valid). In view of the above formal argument, all we need to have is a solution to (1.3.2) on which we may perform the classical calculations needed to establish the representation formula (1.2.3) (or, in other words, on which we may perform Itô calculus). This is precisely the case for our solution, since it has been obtained by regularization. Therefore we will not explicitly use, for the uniqueness part, *all* the assumptions necessary for the uniqueness of the solution of the partial differential equation. As the reader now knows it well, the proof of the latter uniqueness is based upon a regularization step and an a priori estimate. Here we will use the regularization (for our Itô formula), but the a priori estimate is already encoded in our notion of solution: it is property (2.2.3) of our process solution $\mathbf{X}_t$, which is the probabilistic formulation of that estimate. So, in a sense, we do, implicitly if not explicitly, use all aspects of our proof of uniqueness for the partial differential equation.

In our proof of uniqueness of our solution u, we have used that we may regularize equation (1.3.2) by convolution with a regularizing kernel ρ_ε. We may indeed infer from (1.3.2) that

$$\partial_t u^\varepsilon - \mathbf{b}\,.\,\nabla u^\varepsilon - \frac{1}{2}\Delta u^\varepsilon = -r_\varepsilon,$$

where we have denoted by $u^\varepsilon = \rho_\varepsilon \star u$ and $r_\varepsilon = [\mathbf{b}\,.\,\nabla, \rho_\varepsilon](u)$, a commutator the properties of which have been repeatedly seen; see for instance (2.1.5) in Lemma 1. The function u^ε is, by construction, regular. On the other hand,

$$\mathbf{X}_t = x + \mathbf{W}_t + \int_0^t \mathbf{b}(\mathbf{X}_s)\,ds,$$

because of Definition 1, has all the required properties that allow for proving (by, say, passing to the limit in a discretization in time) the Itô formula on u^ε, for ε fixed. Precisely, we consider

$$\begin{aligned}
-u^\varepsilon(t,x) + u^\varepsilon(0,\mathbf{X}_t) &= \int_0^t \frac{d}{ds}\,u^\varepsilon(t-s,\mathbf{X}_s)\,ds \\
&= -\int_0^t \frac{\partial u^\varepsilon}{\partial t}(t-s,\mathbf{X}_s)\,ds + \int_0^t \mathbf{b}(\mathbf{X}_s)\,.\,\nabla u^\varepsilon(t-s,\mathbf{X}_s)\,ds \\
&\quad + \frac{1}{2}\int_0^t \Delta u^\varepsilon(t-s,\mathbf{X}_s)\,ds + \int_0^t \nabla u^\varepsilon(t-s,\mathbf{X}_s)\,.\,d\mathbf{W}_s, \\
&= \int_0^t r_\varepsilon(t-s,\mathbf{X}_s)\,ds + \int_0^t \nabla u^\varepsilon(t-s,\mathbf{X}_s)\,.\,d\mathbf{W}_s,
\end{aligned} \tag{2.2.17}$$

where we note that, because of (2.2.3) satisfied by $\mathbf{X}_t$ and because of the L^2 integrability of ∇u^ε, the rightmost integral is a stochastic integral. We indeed have

$$\int\int_0^t \mathbb{E}_x|\nabla u^\varepsilon(t-s,\mathbf{X}_s)|^2\,ds\,dx \le C\|\nabla u^\varepsilon\|^2_{L^2([0,t],L^2)} \tag{2.2.18}$$

thus, for almost all x, $\nabla u^\varepsilon(t-s,\mathbf{X}_t) \in L^2(\Omega\times[0,t])$. It follows, taking the expectation of both sides of (2.2.17), that

$$-u^\varepsilon(t,x) + \mathbb{E}_x(u_0(\mathbf{X}_t)) = \int_0^t \mathbb{E}_x(r_\varepsilon(t-s,\mathbf{X}_s))\,ds.$$

We next remark that when we integrate over x the integrand on the right-hand side, we have, again using property (2.2.3) satisfied by definition by the solution $\mathbf{X}_t$,

$$\int \mathbb{E}_x(|r_\varepsilon|(t-s,\mathbf{X}_s))\,dx \le c\int |r_\varepsilon|(t-s,x)\,dx. \tag{2.2.19}$$

We have therefore obtained

$$\int |u^{\varepsilon}(t,x) - \mathbb{E}_x(u_0(\mathbf{X}_t))|\,dx \le c\int_0^t \int |r_\varepsilon|(t-s,x)\,dx\,ds.$$

The right-hand side vanishes with ε because, for time-dependent functions, the remainder r_ε of the commutation Lemma 1 is well known to converge to zero in $L^1_{t,x}$. In the left-hand side, u^{ε} converges to u by construction. The Feynman–Kac formula

$$u(t,x) = \mathbb{E}_x(u_0(\mathbf{X}_t))$$

follows, thereby *characterizing*, and thus in particular proving the uniqueness of, the law of $\mathbf{X}_t$ for all times t. Now, we have precisely assumed on $\mathbf{X}_t$ all the necessary properties (semi-group property and continuity of the trajectories) so that this implies the uniqueness-in-law of the process $\mathbf{X}_t$ itself.

Remark 12. Note that in fact, in this proof of uniqueness, we only make use of a form of (2.2.3) *integrated in time*.

Corollary 2. *Consider the setting of Theorem 1, and assume we have the conditions required for performing the regularization of the parabolic equation, namely* (2.1.1). *Then a solution to the stochastic equation* (2.2.1) *in the sense of Definition 1 is unique in law.*

Remark 13. In the case of the variant of Definition 1 mentioned in our Remark 8, the above argument is modified as follows. We first notice that (2.2.18) is still valid using that $|\nabla u^{\varepsilon}|^{\frac{2}{p}} \in L^p$. Next, (2.2.19) now reads

$$\int \mathbb{E}_x(|r_\varepsilon|(t-s,\mathbf{X}_s))\,dx \le c\Big(\int |r_\varepsilon|^p(t-s,x)\,dx\Big)^{\frac{1}{p}}, \tag{2.2.20}$$

and we are able to conclude whenever the remainder r_ε vanishes in L^p. The latter property requires the corresponding assumptions on the transport field $\mathbf{b}$ for the regularization step. While our commutation lemma, Lemma 1, addresses the case of a remainder vanishing in L^1, we can similarly obtain a remainder vanishing in L^p assuming that $\mathbf{b} \in L^{\frac{2p}{2-p}} + W^{1,p}$ if $p \le 2$, $\mathbf{b} \in L^\infty + W^{1,p}$ if $p = 2$, or $\mathbf{b} \in W^{1,p}$ if $p > 2$. The corresponding result may then be easily stated in a variant of Corollary 2.

Existence of a solution-in-law. In order to establish the existence of a solution in law to (2.2.1), we essentially follow the pattern of the classical proof of existence in the regular setting. We first introduce the martingale problem associated to our notion of solution, and next prove existence of a solution for this martingale problem by regularization.

First, we define the martingale problem associated to our stochastic differential equation. We equip the space of continuous functions $\mathcal{W} = \mathcal{C}([0,T],\mathcal{D})$ with the canonical filtration, which we denote by $\mathcal{B}_t$. Notice that the domain $\mathcal{D}$ may be either the whole space or the torus $\mathbb{T} = [0,1]^d_{\mathrm{per}}$ in the informal presentation we give here. As mentioned at the beginning of the section, technical arguments are necessary to address the issues at the boundary: in the classical setting $\mathbb{R}^d$, we need conditions on the behavior at infinity of $\mathbf{b}$ (and σ when later this matrix will be a varying matrix), and in the case of the torus $\mathbb{T}$, we need some technical modifications of the norm manipulated.

We say that $\mathbb{P} = (\mathbb{P}_x)_{x \in \mathcal{D}}$ on $\mathcal{W}$ is a family of probabilities solution to the *generalized martingale problem* if

(i) for almost all $x \in \mathcal{D}$ (in the sense of the Lebesgue measure), $\mathbb{P}_x(y(t = 0) = x) = 1$,

(ii) for all times t, there exists a constant C such that the marginal probability $\mathbb{P}_{x,t}$ of $\mathbb{P}_x$ at time t satisfies

$$\iint_{\mathcal{W}} |f(y(t))| \mathbb{P}_{x,t}(dy)\, dx \le C \int |f(z)|\, dz \tag{2.2.21}$$

for all L^1 function f, and

(iii) for almost all x, under the probability $\mathbb{P}_x$, the canonical process $\mathbf{X}_t : y \to y(t)$ from $\mathcal{W}$ to $\mathcal{D}$ solves the martingale formulation (1.2.14) in a weak sense, which we express upon saying that for all times $0 \le s \le t$, all regular function φ, and all bounded continuous $\mathcal{B}_s$-measurable function g,

$$\int \left[\varphi(y(t)) - \varphi(y(s)) - \int_s^t \left(\mathbf{b} \,.\, \nabla\varphi + \frac{1}{2}\Delta\varphi \right)(y(\tau))\, d\tau \right] g(y) \mathbb{P}_x(dy)\, dx = 0. \tag{2.2.22}$$

Notice that, because of (2.2.21), the term in $\mathbf{b} \,.\, \nabla\varphi$ makes sense. In a more concise and self-explanatory notation, (2.2.22) essentially expresses that

$$\varphi(\mathbf{X}_t) - \int_0^t A_\varphi(\mathbf{X}_s)\, ds$$

is a martingale, $\mathbb{P}_x$ almost everywhere.

We now note that a weak (respectively a fortiori, a strong) solution to the stochastic differential equation (2.2.1) in the sense of Definition 1 gives rise to a solution to the above martingale problem. In order to avoid confusion, denote temporarily by $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathcal{F}}_t, \widetilde{\mathbb{P}})$ the probability space and by $\widetilde{\mathbf{X}_t}$ the weak solution on that space. Next define $\mathbb{P} = \widetilde{\mathbb{P}}(\widetilde{\mathbf{X}_t}^{-1})$ the image measure on $\mathcal{W}$ of the original probability $\widetilde{\mathbb{P}}$ on $\widetilde{\Omega}$, and by $\mathbf{X}_t$ the canonical process obtained. First of all, $\int_0^t \mathbf{b} \,.\, \nabla\varphi(s, \mathbf{X}_s)\, ds$ makes sense for φ regular and for the class of (only locally integrable) non-regular fields $\mathbf{b}$ we consider, because $\widetilde{\mathbf{X}_t}$ satisfies property (2.2.3) and thus $\mathbb{P}$ satisfies (2.2.21). Next, for a fixed regular function φ, we may proceed on $\varphi(t, \mathbf{X}_t)$ with Itô differential rule as we did above for $u^\varepsilon(t, \mathbf{X}_t)$ in our proof of uniqueness in law. We end up with

$$\varphi(t, \mathbf{X}_t) - \varphi(s, \mathbf{X}_s) - \int_s^t \left(\partial_\tau\varphi + \mathbf{b} \,.\, \nabla\varphi + \frac{1}{2}\Delta\varphi \right)(\tau, \mathbf{X}_\tau)\, d\tau = \int_s^t \nabla\varphi(t - \tau, \mathbf{X}_\tau) \,.\, d\widetilde{\mathbf{W}}_\tau. \tag{2.2.23}$$

And, as above in (2.2.18), this integral is a stochastic integral, thus a martingale, and the martingale problem is thus solved.

The reciprocal property, namely that a solution to the martingale problem yields a weak solution to the stochastic differential equation (that is, a probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and a *couple* (W_t, X_t), adapted to $\mathcal{F}_t$, where W_t is a Brownian motion and X_t a solution to (2.2.1) for that particular setting) is unclear in our generalized setting. In the classical setting, the proof, given $\mathbb{P}$ a solution to the martingale problem, to establish the existence of a weak solution to the stochastic differential equation consists in taking as test function in the martingale problem first the coordinate fields $\varphi(t, x) = x_i$, $1 \leq i \leq d$, thereby reconstructing a process $\mathbf{X}_t$. Next, one uses $\varphi(t, x) = x_i x_j$, $1 \leq i, j \leq d$ to show the cross-variations of $\mathbf{X}_t$ enjoy the necessary absolute continuity in time in order to apply the Doob representation Theorem and conclude to the existence of a Brownian motion $\mathbf{W}_t$ such that $\mathbf{X}_t$ solves the stochastic differential equation.

What we prove here is the existence of a solution for the martingale problem. The proof in some sense mimics that of the regular setting (see, e.g., [58, p. 323]): prove convergence of the sequence of probabilities solutions to the regularized formulation (in the classical setting, this is completed proving tightness, below it will be "tightness once integrated in x") and next pass to the limit in the regularized formulation. We make here a detailed proof of the key ingredient of the proof: we show that the sequence of probabilities converges in an appropriate sense.

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ denote the probability space. For $\varepsilon > 0$ fixed, we may solve (2.2.1), with the initial condition x, in a classical way, for a regularized field $\mathbf{b}^\varepsilon$, converging, as ε vanishes, to $\mathbf{b}$ in a sense made precise below. Let $\mathbf{X}_t^{x,\varepsilon}$ denote the solution obtained. Defining $\mathbb{P}_x^\varepsilon = \mathbb{P}(\mathbf{X}_t^{x,\varepsilon})^{-1}$, we obtain a solution to the martingale problem. We now assume $\mathbf{b}$ satisfies the assumptions that make Lemma 2 valid, and that $\mathbf{b}^\varepsilon$ approaches $\mathbf{b}$ in the functional spaces of the statement of that Lemma, so that the L^1 estimate (2.1.24) is valid *uniformly in* ε. It follows that, again uniformly in ε, $\mathbf{X}_t^{x,\varepsilon}$ satisfies (2.2.3). Since the bounds we have obtained in (2.2.6) and (2.2.12) (in the case of a varying σ but it of course applies also to the case of a constant σ, in which case we in fact deal with the Brownian motion), depend only upon estimate (2.2.3), it also follows that, for $\frac{1}{p} < s < \frac{1}{2}$, $\int \mathbb{E}_x(\|\mathbf{X}_t^{x,\varepsilon}\|_{W^{s,p}}^p)\,dx$ is bounded uniformly in ε, which, given the definition of $\mathbb{P}_x^\varepsilon$, is exactly saying that

$$\int \mathbb{E}_x^\varepsilon(\|\mathbf{X}_t\|_{W^{s,p}}^p)\,dx \tag{2.2.24}$$

is bounded uniformly in ε, where we have of course denoted by $\mathbb{E}_x^\varepsilon$ the expectation under the probability $\mathbb{P}_x^\varepsilon$.

We now fix an arbitrary function $\varphi \in L_x^q$, for $\frac{1}{p} + \frac{1}{q} = 1$, and consider the measure on $\mathcal{W}$ defined by

$$\mu_\varphi^\varepsilon = \int \varphi(x)\mathbb{P}_x^\varepsilon\,dx.$$

It has total mass

$$\mu_\varphi^\varepsilon(\mathcal{W}) = \int \varphi(x)\,dx.$$

The bound (2.2.24) implies that, for all R,

$$\begin{aligned}\mu_\varphi^\varepsilon(\|\mathbf{X}_t\|_{W^{s,p}} \geq R) &= \int \varphi(x)\,\mathbb{P}_x^\varepsilon(\|\mathbf{X}_t\|_{W^{s,p}} \geq R) \\ &\leq \frac{1}{R}\int |\varphi|\,\mathbb{E}_x^\varepsilon(\|\mathbf{X}_t\|_{W^{s,p}}^p)^{\frac{1}{p}} \\ &\leq \frac{1}{R}C\Big(\int |\varphi|^q\Big)^{\frac{1}{q}}. \end{aligned}\tag{2.2.25}$$

We know that, in particular $\frac{1}{p} < s$, thus $W^{s,p}$ is continuously embedded in the space of continuous functions of the real line. The sequence μ_φ^ε is thus tight, and therefore (up to an extraction we omit), narrowly converges. We observe that its limit reads as the bounded measure $\int \varphi(x)\,\mu_x\,dx$ for $\mu_x \in L_x^p(\mathcal{M}_b(\mathcal{W}))$, where $\mathcal{M}_b(\mathcal{W})$ denotes the space of bounded measures on $\mathcal{W}$. We of course have $\mu_x \geq 0$. We also have $\mu_x(\mathcal{W}) \leq 1$, for almost all x. Indeed, since, for almost all x, $\int \mathbb{1}_\mathcal{W}\mathbb{P}_x^\varepsilon = 1$, we have, by integration,

$$\int \varphi(x)\,dx = \int \varphi(x)\mathbb{1}_\mathcal{W}\mathbb{P}_x^\varepsilon\,dx.$$

Using positiveness, estimating the right-hand side from below, using weak convergence, we easily get

$$\liminf \int \varphi(x)\,dx \geq \int \varphi(x)\mu_x(\mathcal{W})\,dx,$$

which shows $\mu_x(\mathcal{W}) \leq 1$. The fact that the whole mass is recovered in the limit, so that μ_x is indeed a probability for almost all x, is a consequence of the tightness. We have

$$\lim_{\varepsilon\to 0}\int \varphi(x)\mathbb{P}_x^\varepsilon\mathbb{1}_\mathcal{W}\,dx = \lim_{\varepsilon\to 0}\int \mu_\varphi^\varepsilon\mathbb{1}_\mathcal{W} = \int \varphi(x)\mu_x(\mathcal{W})\,dx,$$

and since the left-hand side is clearly $\int \varphi(x)\,dx$, this shows that $\mu_x(\mathcal{W}) = 1$ for almost all x. The limit of the sequence $\mathbb{P}_x^\varepsilon$ is therefore, for almost all x, a probability, which we may denote by $\mathbb{P}_x$. Having obtained that the limit $\mathbb{P}_x$ of $\mathbb{P}_x^\varepsilon$ is a family of probabilities, it remains to pass to the limit in all the properties of $\mathbb{P}_x^\varepsilon$, including the martingale formulation, to complete the proof of existence.

We have actually outlined the proof of the following:

Corollary 3. *Assume that* $\mathbf{b}$ *is such that the solution to the parabolic equation satisfies the* L^1 *estimate* (2.1.24) *of Lemma 2. Then there exists a solution in law to the stochastic equation* (2.2.1)*, in the sense that there exists a solution to the* generalized martingale problem *introduced above, formally equivalent to that equation in the sense of Definition 1.*

Remark 14. Notice that the argument we gave above may be seen as the classical argument for the existence of a solution to the martingale problem, which proceeds using Prokhorov's Theorem, except that, above, we need to "integrate all arguments in x", loosely speaking.

2.2.3 Pathwise uniqueness and strong existence

Pathwise uniqueness of the solution. As in the classical (regular) setting, we need to consider the parabolic equation of an augmented space of doubled dimension. In the simple case of a constant σ, (1.2.11) writes

$$\frac{\partial u}{\partial t} - \mathbf{b}(x).\nabla_x u - \mathbf{b}(y).\nabla_y u \\ - \frac{1}{2}\sigma_{ik}\sigma_{jk}\frac{\partial^2 u}{\partial x_i \partial x_j} - \sigma_{ik}\sigma_{jk}\frac{\partial^2 u}{\partial x_i \partial y_j} - \frac{1}{2}\sigma_{ik}\sigma_{jk}\frac{\partial^2 u}{\partial y_i \partial y_j} = 0. \tag{2.2.26}$$

Introducing the change of variables

$$\begin{cases} \eta = x + y, \\ \xi = x - y, \end{cases} \qquad \begin{cases} x = \frac{1}{2}(\eta + \xi), \\ y = \frac{1}{2}(\eta - \xi), \end{cases}$$

and denoting by

$$\begin{cases} \mathbf{b}^\eta = \frac{1}{2}(\mathbf{b}(x) + \mathbf{b}(y)) = \frac{1}{2}\Big(\mathbf{b}\Big(\frac{1}{2}(\eta + \xi)\Big) + \mathbf{b}\Big(\frac{1}{2}(\eta - \xi)\Big)\Big), \\ \mathbf{b}^\xi = \frac{1}{2}(\mathbf{b}(x) - \mathbf{b}(y)) = \frac{1}{2}\Big(\mathbf{b}(\frac{1}{2}(\eta + \xi)\Big) - \mathbf{b}\Big(\frac{1}{2}(\eta - \xi)\Big)\Big), \end{cases} \tag{2.2.27}$$

this equation writes

$$\frac{\partial u}{\partial t} - \mathbf{b}^\eta(\eta, \xi).\nabla_\eta u - \mathbf{b}^\xi(\eta, \xi).\nabla_\xi u - \frac{1}{2}\sigma_{ik}\sigma_{jk}D^2_{\eta_i\eta_j}u = 0. \tag{2.2.28}$$

Even in the case when $\sigma\sigma^t$ is positive definite (which we do not necessarily assume here), the second-order term in (2.2.28) is not positive definite, given the absence of ellipticity in the direction ξ at least. To proceed with a result of existence and uniqueness for the solution to (2.2.28), we therefore need to assume on $\mathbf{b}$ suitable conditions. Intuitively, to regularize a partial differential equation of the type (2.2.28), the situation is as follows.

- In any direction where there is no parabolic regularization, and this is at least the case in all directions ξ, and possibly, depending upon the directions of ellipticity of $\sigma\sigma^t$, also some directions in η, we need assumptions *as for the transport equation*. We mean, for simplicity, assumptions in the spirit of

  $$\mathbf{b} \in W^{1,1}, \quad \operatorname{div} \mathbf{b} \in L^\infty,$$

  respectively for the regularization and for the a priori estimate.
- On the other hand, in any direction where a parabolic regularization operates, we may consider more general assumptions *as for the heat equation*, namely of the type

  $$\mathbf{b} \in L^2, \quad (\mathbf{b} \in L^d \text{ or } \operatorname{div} \mathbf{b} \in L^{\frac{d}{2}}),$$

  again for the regularization and the estimate.

In either case, all these assumptions must of course additionally enjoy agreeable properties of integrability in the remaining directions.

To make all this precise, we shall only, for the sake of simplicity, consider the case where $\sigma\sigma^t$ is indeed positive definite. This allows us to work with Theorem 1 as a basis for our results. The modifications of our results when the constant matrix $\sigma\sigma^t$ is only nonnegative are not difficult, and all in the vein of how we have adapted the results of Theorem 1 to those of Corollary 1.

It is immediate to realize that the adaptation of assumptions (2.1.1) and (2.1.2) to our situation implies that we have to assume

$$\begin{cases} \mathbf{b}^\eta = \beta_1^\eta + \beta_2^\eta, \quad \beta_1^\eta \in L^2_{\eta,\xi}, \ \beta_2^\eta \in L^1_\xi(W^{1,1}_\eta), \\ \mathbf{b}^\xi \in L^1_\eta(W^{1,1}_\xi), \end{cases} \tag{2.2.29}$$

for the regularization step, while, for the a priori estimate, we need to assume

$$\begin{cases} \mathbf{b}^\eta = \mathbf{b}_1^\eta + \mathbf{b}_2^\eta, \quad \mathbf{b}_1^\eta \in L^\infty_\xi(\varepsilon_1 L^{d,\infty}_\eta + L^d_\eta), \ [\operatorname{div}_\eta \mathbf{b}_2^\eta]_- \in L^\infty_\xi(\varepsilon_2 L^{\frac{d}{2},\infty}_\eta + L^{\frac{d}{2}}_\eta), \\ [\operatorname{div}_\xi \mathbf{b}^\xi]_- \in L^\infty_{\eta,\xi}. \end{cases} \tag{2.2.30}$$

These assumptions are "sharp", in the following sense: "as sharp as our assumptions for Theorem 1". It turns out that, in the specific setting of a drift $(\mathbf{b}^\eta, \mathbf{b}^\xi)$ of the form (2.2.27) (which *is* particular because $(\mathbf{b}^\eta + \mathbf{b}^\xi)(\eta, \xi)$ is a function of only $\frac{1}{2}(\eta + \xi)$ and $(\mathbf{b}^\eta - \mathbf{b}^\xi)(\eta, \xi)$ a function of only $\frac{1}{2}(\eta - \xi)$), these assumptions reduce to more classical, and actually more restrictive, assumptions. Let us indeed express (2.2.29) and (2.2.30) in terms of the original drift term $\mathbf{b}$ as a function of the single d-dimensional variable x.

One may readily see that conditions (2.2.29), which (even if we were working on an unbounded domain) are *local* conditions since they are the conditions for the regularization step, are equivalent to $\mathbf{b} \in W^{1,1}$. Indeed, the second line of (2.2.29) in fact reads as $\mathbf{b}^\xi$ and $\nabla_\xi \mathbf{b}^\xi$ both in $L^1_{\eta,\xi}$, which, given the definition (2.2.27), amounts to saying that $\mathbf{b}(x) - \mathbf{b}(y)$ and $\nabla \mathbf{b}(x) + \nabla \mathbf{b}(y)$ are both in $L^1_{x,y}$. This is exactly $\mathbf{b} \in W^{1,1}$. One proceeds similarly for the conditions (2.2.30) and realize that they are equivalent to the strongest condition $[\operatorname{div} \mathbf{b}]_- \in L^\infty$. In the case of an ambient dimension $d \le 2$, the argument may be easily modified, and we eventually obtain the same conditions on $\mathbf{b}$.

Remark 15. The fact that, for what concerns pathwise uniqueness, we do not *gain* any regularity because of the parabolic regularization, and that we have to assume ($\mathbf{b} \in W^{1,1}$, $[\operatorname{div} \mathbf{b}]_- \in L^\infty$), is intuitively clear. Since the two solutions share the same Brownian motion, their difference indeed formally satisfies an *ordinary* differential equation, for which we need the above "usual" assumptions on $\mathbf{b}$ to perform the regularization.

We thus have:

Corollary 4. *As in Theorem 1, we assume that the matrix coefficient $\sigma\sigma^t$ is a* constant symmetric positive definite *matrix. The ambient dimension needs not be larger than or equal to* 3. *We assume* ($\mathbf{b} \in W^{1,1}$, $[\operatorname{div} \mathbf{b}]_- \in L^\infty$). *Then, for any initial condition $u_0 \in L^\infty$, there exists a unique solution u to* (2.2.28) *(or* (2.2.26)*) in the functional* *space* $C([0,T],L^p)\cap L^\infty([0,T],L^\infty)$, *for all* $1\le p<+\infty$, *with* $\nabla_\eta u\in L^2([0,T],L^2)$. *In addition, estimates* (2.1.3) *and* (2.1.4) *hold.*

Given Corollary 4, we are now in a position to establish pathwise uniqueness and, next, existence, for the (strong) solution to (2.2.1) in the sense of Definition 1, provided we assume we work in the conditions of that Corollary, of course. Assume that $\mathbf{X}_t$ and $\hat{\mathbf{X}}_t$ are two (families of) solutions, both in the sense of Definition 1, to (2.2.1), for the *same* probability space $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbb{P})$, the *same given* Brownian motion $\mathbf{W}_t$ on that probability space, and the *same* initial condition. The couple $(\mathbf{X}_t,\hat{\mathbf{X}}_t)$ is therefore a solution to the system of stochastic differential equations (1.2.9) (with (1.2.10)) which we introduced in Section 1.2. We intend to prove that these two families of solutions $\mathbf{X}_t$ and $\hat{\mathbf{X}}_t$ are identical, that is, $\mathbf{X}_t$ and $\hat{\mathbf{X}}_t$ are almost surely equal for almost all initial conditions.

We first show that the law of the couple $(\mathbf{X}_t^x,\hat{\mathbf{X}}_t^y)$ is associated to (2.2.26). Indeed, for any sufficiently regular function u_0, we may prove, arguing exactly as we did above on page 57 (see (2.2.17) and subsequent formulae), that the solution to (2.2.26) writes as

$$u(t,x,y)=\mathbb{E}(u_0(\mathbf{X}_t^x,\hat{\mathbf{X}}_t^y))$$

almost everywhere in x,y. This holds precisely because we have assumed on $\mathbf{b}$ the hypotheses of Corollary 4 which allow to construct, and characterize in law, the (unique) solution to (2.2.26), *and* because we have assumed that both $\mathbf{X}_t$ and $\hat{\mathbf{X}}_t$ satisfy (2.2.3), so that our calculations of page 57 leading to the above Feynman–Kac formula make sense.

In particular, the law of the couple $(\mathbf{X}_t^x,\hat{\mathbf{X}}_t^y)$ is entirely determined. Likewise, the same conclusions hold on the couple $(\mathbf{X}_t^x,\mathbf{X}_t^y)$, evidently. It follows, by uniqueness, that

$$\mathbb{E}(u_0(\mathbf{X}_t^x,\hat{\mathbf{X}}_t^y))=\mathbb{E}(u_0(\mathbf{X}_t^x,\mathbf{X}_t^y)).$$

Taking $u_0=|\cdot|$ where $|\cdot|$ denotes the Euclidean norm throughout the rest of this section, or, in the one-dimensional setting, the absolute value (and, if need be, approximating this function by a sequence of regular functions u_0 and passing to the limit), we obtain

$$\mathbb{E}(|\mathbf{X}_t^x-\hat{\mathbf{X}}_t^y|)=\mathbb{E}(|\mathbf{X}_t^x-\mathbf{X}_t^y|). \tag{2.2.31}$$

If the process were continuously depending upon their initial conditions, we would like to take the limit $|x-y|\to 0$ and deduce from (2.2.31) that

$$\mathbb{E}(|\mathbf{X}_t^y-\hat{\mathbf{X}}_t^y|)=\mathbb{E}(|\mathbf{X}_t^y-\mathbf{X}_t^y|)=0. \tag{2.2.32}$$

However, we do not know that this continuity holds. A classical argument to circumvent that difficulty is based upon the Lebesgue Differentiation Theorem (see, for instance, [44, pp. 43–44]). In the case at hand, it reads as follows. First, up to changing $\mathbf{X}$ into

$$\mathbf{Z}=\frac{\mathbf{X}}{\sqrt{1+|\mathbf{X}|^2}}$$

(that is, $\mathbf{X}_t^x$ into $\mathbf{Z}_t^x$, $\hat{\mathbf{X}}_t^x$ into $\hat{\mathbf{Z}}_t^x$, $\hat{\mathbf{X}}_t^y$ into $\hat{\mathbf{Z}}_t^y$, all formulae that are invertible, so proving $\mathbf{X}_t^x=\hat{\mathbf{X}}_t^x$ amounts to proving that $\mathbf{Z}_t^x=\hat{\mathbf{Z}}_t^x$), we may always assume that $\mathbf{X}$ is bounded, thus in particular in L^2. We may then consider $u_0 = |\cdot|^2$ instead of $u_0 = |\cdot|$ in formulae (2.2.31)–(2.2.32) above. Next, we integrate both sides of (2.2.31) on a ball $B(y,\delta)$, centered at y and of radius $\delta > 0$, and we obtain

$$\frac{1}{|B(y,\delta)|}\int_{B(y,\delta)} \mathbb{E}(|\mathbf{X}_t^x - \hat{\mathbf{X}}_t^y|^2)\,dx = \frac{1}{|B(y,\delta)|}\int_{B(y,\delta)} \mathbb{E}(|\mathbf{X}_t^x - \mathbf{X}_t^y|^2)\,dx. \tag{2.2.33}$$

We intend to take the limit $\delta \to 0$ and show that the right-hand side vanishes in this limit, almost everywhere in y. Successively using the Young inequality and (2.2.31), we will then have

$$\begin{aligned}\mathbb{E}(|\mathbf{X}_t^y - \hat{\mathbf{X}}_t^y|^2) &= \frac{1}{|B(y,\delta)|}\int_{B(y,\delta)} \mathbb{E}(|\mathbf{X}_t^y - \hat{\mathbf{X}}_t^y|^2)\,dx \\ &\le \frac{2}{|B(y,\delta)|}\int_{B(y,\delta)} \mathbb{E}(|\mathbf{X}_t^y - \mathbf{X}_t^x|^2)\,dx + \frac{2}{|B(y,\delta)|}\int_{B(y,\delta)} \mathbb{E}(|\mathbf{X}_t^x - \hat{\mathbf{X}}_t^y|^2)\,dx \\ &= \frac{4}{|B(y,\delta)|}\int_{B(y,\delta)} \mathbb{E}(|\mathbf{X}_t^y - \mathbf{X}_t^x|^2)\,dx \end{aligned} \tag{2.2.34}$$

and because the right-hand side of (2.2.33) vanishes, the conclusion (2.2.32) holds for almost all y, which proves uniqueness. Temporarily denoting by $f(y,\omega) = \mathbf{X}_t^y(\omega)$ (for t fixed, since time does not play any role in the argument below), the right-hand side of (2.2.33) reads as follows:

$$\begin{aligned}&\frac{1}{|B(y,\delta)|}\int_{B(y,\delta)} \mathbb{E}(|f(x,\omega) - f(y,\omega)|^2)\,dx \\ &\quad= \mathbb{E}(|f(y,\omega)|^2) + \frac{1}{|B(y,\delta)|}\int_{B(y,\delta)} \mathbb{E}(|f(x,\omega)|^2)\,dx \\ &\qquad - 2\frac{1}{|B(y,\delta)|}\int_{B(y,\delta)} \mathbb{E}(f(x,\omega)f(y,\omega))\,dx. \end{aligned} \tag{2.2.35}$$

Because of the local integrability of $\mathbb{E}(|f|^2)$ (that is, $\mathbb{E}(|\mathbf{X}_t|^2)$), and because of the Lebesgue Differentiation Theorem, the second term converges, as $\delta \to 0$, almost everywhere in y, to $\mathbb{E}(|f(y,\omega)|^2)$. The rightmost term of (2.2.35) requires a specific treatment because of the presence of the extra variable ω that lives in a not necessarily countable set (otherwise, the whole argument is much easier, see again [44, pp. 43–44]). We temporarily freeze the first factor $f(x,\omega)$ in this term and denote by

$$C_\delta(c,y) := \frac{1}{|B(y,\delta)|}\int_{B(y,\delta)} \mathbb{E}(f(x,\omega)c(y,\omega))\,dx \tag{2.2.36}$$

for a parameter function $c(y,\omega)$ in L^2. If $c(y,\omega)$ were a product function

$$c(y,\omega) = g(y)h(\omega),$$

then one would write

$$C_\delta(c,y) = \frac{1}{|B(y,\delta)|} \int\limits_{B(y,\delta)} \mathbb{E}(f(x,\omega)c(y,\omega))\,dx$$
$$= g(y)\frac{1}{|B(y,\delta)|} \int\limits_{B(y,\delta)} \mathbb{E}(f(x,\omega)h(\omega))\,dx, \tag{2.2.37}$$

and would obtain, again by the Lebesgue Differentiation Theorem this time applied to the function $\mathbb{E}(f(\cdot,\omega)h(\omega))$, that $C_\delta(c,y)$ converges, as $\delta \to 0$, almost everywhere in y, to $C_0(c,y) := g(y)\mathbb{E}(f(y,\omega)h(\omega)) = \mathbb{E}(f(y,\omega)c(y,\omega))$. The same argument would allow to conclude if $c(y,\omega)$ were a *finite sum* of such product functions:

$$c(y,\omega) = \sum_{k=1}^{K} g_k(y)h_k(\omega).$$

In the general case, we now use the density of such sums. For any $\eta > 0$, we know there exists a function c_K, sum of product functions of the variables y and ω separately, such that

$$\|C_\delta(c,y) - C_\delta(c_K,y)\|_{L^1} \leq \eta, \tag{2.2.38}$$

uniformly in δ. The density of the functions c_K holds by a direct application of the Hahn–Banach Theorem: functions of (y,ω) orthogonal to products $g(y)h(\omega)$, and therefore sums of such products, are necessarily identically zero (consider the marginal in ω, prove it vanishes using the Hahn–Banach Theorem, and then proceed similarly for the function itself). Property (2.2.38) follows. In addition, for this particular function c_K, we know that

$$C_\delta(c_K,y) \xrightarrow{\delta\to 0} C_0(c_K,y)$$

almost everywhere in y and also in L^1. Along with (2.2.38), this proves, for all $c \in L^2$, the convergence of $C_\delta(c,y)$ to $C_0(c,y)$ in L^1, as $\delta \to 0$. Thus, extracting a subsequence if necessary (and a subsequence is in any event all what we need to conclude), we obtain the limit almost everywhere in y, for a general function c. Consequently, applying this to $c \equiv f$ and collecting all the three terms of the right-hand side of (2.2.35), we obtain the convergence to zero, again almost everywhere in y, and thus our conclusion. We have thus proved the following result.

Corollary 5. *Assume the setting of Corollary 4, and likewise* ($\mathbf{b} \in W^{1,1}$, $[\operatorname{div}\mathbf{b}]_- \in L^\infty$). *Then a solution, in the sense of Definition 1, to* (2.2.1) *is pathwise unique.*

Existence of a strong solution. We prove, for a *given* probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and a *given* Brownian motion $\mathbf{W}_t$ on that probability space, the existence of a (strong) solution to (2.2.1) in the sense of Definition 1.

At this stage it is useful, for pedagogic purposes, to briefly recall the proof performed in [41] of the existence of the solution to the *ordinary* differential equation analogous to (2.2.1) for a drift $\mathbf{b}$ such that, say, ($\mathbf{b} \in W^{1,1}$, $\operatorname{div}\mathbf{b} = 0$). The proof is performed by regularization. A sequence of regularized drifts $\mathbf{b}^\varepsilon$ that approaches $\mathbf{b}$ (in the suitable functional space) is considered. For all $\varepsilon > 0$, one may solve the differential equation $\dot{\mathbf{X}}^\varepsilon(t,x) = \mathbf{b}^\varepsilon(\mathbf{X}^\varepsilon(t,x))$ in the classical sense. Bounds (independent of ε) on the sequence $\mathbf{X}^\varepsilon$ are obtained, while, for each ε, since $\mathbf{b}^\varepsilon$ is regular, $u^\varepsilon(t,x) = u_0(\mathbf{X}^\varepsilon(t,x))$ solves the corresponding transport equation (for initial condition u_0). One infers from the "stability" when $\varepsilon \to 0$ of the family of partial differential equations that $\mathbf{X}^\varepsilon(t,x)$ converges to some $\overline{\mathbf{X}}(t,x)$, which satisfies the renormalized ordinary differential equation $\partial_t\Phi(\overline{\mathbf{X}}(t,x)) = (\nabla\Phi)(\overline{\mathbf{X}}(t,x)) . \mathbf{b}(\overline{\mathbf{X}}(t,x))$ in the sense of distributions. It is next proven that $\overline{\mathbf{X}}$ satisfies the conservation of the Lebesgue measure (implied by the condition $\operatorname{div}\mathbf{b} = 0$), that it is a solution to the ordinary differential equation not only in the renormalized and distributions sense, but for almost all x and all t, while the suitable regularities on $\overline{\mathbf{X}}(t,x)$ are eventually established.

Evidently, the key step in the existence proof is the passage to the limit in the transport equation with regularized drift, which gives rise to some solution u, which is identified to be the function $u_0(\overline{\mathbf{X}})$ for some $\overline{\mathbf{X}}$, thereby *creating the limit object* $\overline{\mathbf{X}}$ from the sequence of solutions $\mathbf{X}^\varepsilon$. The fact that this limit $\overline{\mathbf{X}}$ satisfies the differential equation and all the necessary properties is then a simple consequence of the setting provided. In our stochastic setting, we are going to proceed similarly, but we have two additional difficulties

(i) after we regularize the drift and remark that $\mathbb{E}_x(u_0(\mathbf{X}_t^{x,\varepsilon}))$ solves the parabolic equation, all we can expect to obtain is an information on $\mathbf{X}_t^{x,\varepsilon}$ *in expectation*, and certainly not a pathwise information, since only *averages* (over the Brownian trajectories) of $\mathbf{X}_t^{x,\varepsilon}$ matter in the partial differential equation. So, as above for our proof of pathwise uniqueness, we will have to resort to the parabolic equation in the space of doubled dimension in order to pass to the limit *pathwise* on $\mathbf{X}_t^{x,\varepsilon}$.

(ii) in contrast to the setting of the transport equation, the preservation of the Lebesgue measure by the flow is not immediately guaranteed by the control of $\operatorname{div}\mathbf{b}$: we also have to consider specific solutions to the stochastic equation that preserve *on average* this property, so (2.2.3) will play a crucial role to prove that the limit $\overline{\mathbf{X}}_t$ of $\mathbf{X}_t^\varepsilon$ solves, in the suitable sense, the stochastic differential equation.

We now make the proof slightly more precise.

For all $\varepsilon > 0$, we may solve the stochastic differential equation (2.2.1), in a classical way, for the regularized drift $\mathbf{b}^\varepsilon$ and the fixed Brownian motion considered:

$$\mathbf{X}_t^\varepsilon = x + \mathbf{W}_t + \int_0^t \mathbf{b}^\varepsilon(\mathbf{X}_s^\varepsilon)\, ds.$$

Evidently, the solution X^ε enjoys the properties stated in Definition 1. Also evidently, $\mathbb{E}_x(u_0(\mathbf{X}_t^{x,\varepsilon}))$ solves the parabolic equation for all regular function u_0 and satisfies the associated a priori estimates we have established above. We may infer from these bounds on $\mathbf{X}_t^{x,\varepsilon}$, holding in the L^p sense in the variables (t,x,ω) and all independent of ε, that (up to an extraction, which we omit in our notation) $\mathbf{X}_t^{x,\varepsilon}$ converges to some $\overline{\mathbf{X}}_t^x$ *weakly* in the variables (t,x,ω). The point is to prove the *strong* convergence.

For this purpose, we consider, for $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$, and for a given regular initial condition u_0, the equation

$$\begin{aligned}\frac{\partial u}{\partial t} &- \mathbf{b}^{\varepsilon_1}(x).\nabla_x u - \mathbf{b}^{\varepsilon_2}(y).\nabla_y u \\ &- \frac{1}{2}\sigma_{ik}\sigma_{jk}\frac{\partial^2 u}{\partial x_i \partial x_j} - \sigma_{ik}\sigma_{jk}\frac{\partial^2 u}{\partial x_i \partial y_j} - \frac{1}{2}\sigma_{ik}\sigma_{jk}\frac{\partial^2 u}{\partial y_i \partial y_j} = 0, \end{aligned} \tag{2.2.39}$$

which we may easily, uniquely solve, given the regularity of the fields $\mathbf{b}^{\varepsilon_1}$ and $\mathbf{b}^{\varepsilon_2}$. We denote its solution by $u(t, x, y) = u^{\varepsilon_1,\varepsilon_2}(t, x, y)$. We know that this solution reads as

$$u^{\varepsilon_1,\varepsilon_2}(t, x, y) = \mathbb{E}^{x,y}(u_0(\mathbf{X}_t^{x,\varepsilon_1}, \mathbf{X}_t^{y,\varepsilon_2})).$$

One way to proceed to now prove the strong convergence of $\mathbf{X}_t^{x,\varepsilon}$ to $\overline{\mathbf{X}}_t^x$ is to choose $u_0(x, y) = |x - y|^2$ and consider, for ρ_n a sequence of approximation of the Dirac mass, the quantity

$$\begin{aligned} e(\varepsilon_1, \varepsilon_2, n) &= \mathbb{E}\iint |\mathbf{X}_t^{x,\varepsilon_1} - \mathbf{X}_t^{y,\varepsilon_2}|^2 \rho_n(x - y)\,dx\,dy \\ &= \iint u^{\varepsilon_1,\varepsilon_2}(t, x, y)\rho_n(x - y)\,dx\,dy. \end{aligned} \tag{2.2.40}$$

From our results on parabolic equations, we know that, when both ε_1 and ε_2 vanish, $u^{\varepsilon_1,\varepsilon_2}(t, x, y)$ converges to $\overline{u}(t, x, y)$ solution to the parabolic equation with drift $\mathbf{b}$ and same initial condition u_0, so that, for n fixed,

$$e(\varepsilon_1, \varepsilon_2, n) \xrightarrow{\varepsilon_1,\varepsilon_2 \to 0} \iint \overline{u}(t, x, y)\rho_n(x - y)\,dx\,dy =: e(n).$$

Next, as $n \to +\infty$, the right-hand side vanishes:

$$e(n) = \iint \overline{u}(t, x, y)\rho_n(x - y)\,dx\,dy \xrightarrow{n \to +\infty} \int \overline{u}(t, x, x)\,dx = 0. \tag{2.2.41}$$

Indeed, since $\overline{u}(t, x, y)$ is the solution for the initial condition $u_0(x, y) = |x - y|^2$ and the latter function is a super solution to the equation, we have (at least formally but this may be proven rigorously using regularization) for all times, $|\overline{u}(t, x, y)| \leq C|x - y|^2$ and thus $\overline{u}(t, x, x) = 0$. Intuitively, this, once inserted in (2.2.40), shows that the sequence $\mathbf{X}_t^{x,\varepsilon}$ is a Cauchy sequence and thus strongly converges to its limit $\overline{\mathbf{X}}_t^x$. But since we do not know $\overline{\mathbf{X}}_t^x$ continuously depends upon its initial condition x, we have to work a little more, and this is where the local integration (that is, convolution with ρ_n) is useful. This is exactly similar to what we performed above in (2.2.33). We return to (2.2.40) and use convexity to remark that

$$\limsup_{\varepsilon_1,\varepsilon_2} \mathbb{E}\int \left|\mathbf{X}_t^{x,\varepsilon_1} - \int \mathbf{X}_t^{y,\varepsilon_2}\rho_n(x - y)\,dy\right|^2 dx \leq \limsup_{\varepsilon_1,\varepsilon_2} e(\varepsilon_1, \varepsilon_2, n) = e(n).$$

Now, we observe that because of the weak convergence of $\mathbf{X}_t^{x,\varepsilon}$ to $\overline{\mathbf{X}}_t^x$, we also have the weak convergence

$$\int \mathbf{X}_t^{y,\varepsilon_2}\rho_n(x - y)\,dy \xrightarrow{\varepsilon_2 \to 0} \int \overline{\mathbf{X}}_t^y \rho_n(x - y)\,dy =: \overline{\mathbf{X}}_t^{x,n}. \tag{2.2.42}$$

Therefore

$$\limsup_{\varepsilon_1} \mathbb{E} \int \left| \mathbf{X}_t^{x,\varepsilon_1} - \overline{\mathbf{X}}_t^{x,n} \right|^2 dx \leq \limsup_{\varepsilon_1,\varepsilon_2} \mathbb{E} \int \left| \mathbf{X}_t^{x,\varepsilon_1} - \int \mathbf{X}_t^{y,\varepsilon_2} \rho_n(x-y)\, dy \right|^2 dx$$
$$\leq e(n). \tag{2.2.43}$$

It now remains to let $n \to +\infty$. We use that $\overline{\mathbf{X}}_t^{x,n}$ strongly converges to $\overline{\mathbf{X}}_t^{x}$ on the left-hand side, while the right-hand side vanishes because of (2.2.41). We obtain the pathwise convergence:

$$\lim_{\varepsilon_1} \mathbb{E} \int \left| \mathbf{X}_t^{x,\varepsilon_1} - \overline{\mathbf{X}}_t^{x} \right|^2 dx = 0. \tag{2.2.44}$$

We now establish that $\overline{\mathbf{X}}_t^{x}$ solves the stochastic differential equation in the sense of Definition 1. This is the step where property (2.2.3) proves instrumental. Because of the strong convergence $\mathbf{X}_t^{x,\varepsilon} \to \overline{\mathbf{X}}_t^{x}$ we know that the limit (in all appropriate functional spaces) of the solution $u^\varepsilon(t,x) = \mathbb{E}_x(u_0(\mathbf{X}_t^{x,\varepsilon}))$ to the parabolic equation actually reads as $\overline{u}(t,x) = \mathbb{E}_x(u_0(\overline{\mathbf{X}}_t^{x}))$. So all the estimates we have on u^ε pass to the limit into estimates for $\overline{u}$ and give information on $\overline{\mathbf{X}}_t^{x}$. In particular, we may pass to the limit in the regularized stochastic differential equation (at least, for the moment) in the sense of distributions in (t,x,ω). Likewise, since we know that, for all $f \in L^1$, we have a constant c uniform in ε such that (2.2.3) holds,

$$\int \mathbb{E}_x(|f|(\mathbf{X}_t^{x,\varepsilon}))\, dx \leq c \int |f|,$$

we know that $|f|(\mathbf{X}_t^{x,\varepsilon})$ converges in L^1 to $|f|(\overline{\mathbf{X}}_t^{x})$, thus property (2.2.3) follows for the limit process. But this property in turn implies that we obtain the stochastic differential equation not only in the sense of distributions but almost everywhere in x, and in time. It is next obtained for *all* times t because of the continuity of the trajectories itself indeed implied by the same property (2.2.3) applied to the drift $\mathbf{b}$ itself (given that the diffusion is constant and the Brownian motion has continuous trajectories). As for the semi-group property, it directly follows from uniqueness of the solution to the parabolic equation.

We have therefore established:

Corollary 6. *Assume the setting of Corollary 4, and likewise* ($\mathbf{b} \in W^{1,1}$, $[\operatorname{div} \mathbf{b}]_- \in L^\infty$) *and that the L^1 estimate* (2.1.24) *of Lemma 2 holds. Then there exist a solution, in the sense of Definition 1, and strong in the probability sense, to* (2.2.1).

2.3 Initial conditions in L^p, $p < +\infty$

We now consider *unbounded* initial conditions u_0, and, to keep things simple, assume our constant second-order operator is $-\frac{1}{2}\Delta$. To start this section, and although we already mentioned it at the beginning of Chapter 2, let us recall what the difficulty is.

When the initial condition u_0 we supply equation (1) with is *not* bounded, there is no reason to expect that the solution $u(t,\cdot)$ will be bounded. The very definition of the term $\mathbf{b}.\nabla u$ in (1) is then immediately an issue. For the regularization step (assuming that we again work in the setting of positive definite second-order operators), we know that we will have to take $\mathbf{b} \in L^2 + W^{1,1}$. For $\mathbf{b} \in L^2$, giving a sense to $\mathbf{b}.\nabla u$ requires that $\nabla u \in L^2$. However, when $u_0 \in L^p$ for some $1 \leq p < 2$, we cannot expect $\nabla u \in L^2$, since this property does not even hold for the heat equation. On the other hand, when $\mathbf{b} \in W^{1,1} \subset L^{\frac{d}{d-1}}$, expecting $\nabla u \in L^d$ is a fortiori hopeless. So we clearly have to write

$$\mathbf{b}.\nabla u = \operatorname{div}(u\mathbf{b}) - u \operatorname{div} \mathbf{b}.$$

Now, since by Sobolev embeddings, $\mathbf{b} \in L^{\frac{d}{d-1}}$, giving a sense to the product $u\mathbf{b}$ in the former term (the divergence operator would then be understood in the sense of distributions) requires that $u \in L^d$. But this property is unclear. Similarly, the second term requires some integrability of $\operatorname{div} \mathbf{b}$ better than L^1. Addressing an unbounded initial condition u_0 for (1) is thus a difficulty.

Intuitively, we *should* be able to succeed. Indeed, we know that, in the case of a bounded initial condition (and under appropriate additional assumptions), we have, for $1 < p < +\infty$, the L^p estimate (2.1.3) of Theorem 1. This shows that the map $u_0 \to u(t, \cdot)$ is a linear map, continuous from L^p to L^p, with uniform in time estimates (on finite time intervals). By density of L^∞ in L^p, it follows that there exists a unique extension for initial values u_0 only in L^p. The crucial point is, however, to prove that the object constructed is indeed a solution to (1), and that this is the only solution to that equation. Note that in the formal discussion above we leave aside the more intricate case of the space L^1, for which estimate (2.1.24) of Lemma 2 has only been proven under some restrictive assumption. Intuitively, the result should however also extend to that case by the same formal argument.

To address unbounded initial conditions for linear transport equations, the idea introduced by R. DiPerna and the second author was renormalization. We refer the reader to [41] for all the details of the theory and the technical details for transport equations. Since the equation on u does not necessarily make sense, given the difficulties outlined above to define the transport term in the equation, we will write the equation not on u, but on *a function of* u. In our parabolic setting, and in short, it will be said that u is a *renormalized* solution to (1) associated to the initial unbounded condition u_0 when $\Phi(u)$ for all appropriately chosen, smooth, bounded functions Φ is a solution to (1) for the bounded initial condition $\Phi(u_0)$, that is,

$$\partial_t \Phi(u) - \mathbf{b}.\nabla\Phi(u) - \frac{1}{2}\Delta\Phi(u) + \frac{1}{2}\Phi''(u)|\nabla u|^2 = 0. \tag{2.3.1}$$

Loosely speaking, since Φ is bounded, we expect the arguments of the previous sections to apply mutatis mutandis and we thus expect to be able to prove the existence and, foremost, uniqueness of such a renormalized solution. Formally, if we then know the unique solution for all Φ, we actually uniquely know u itself. We therefore understand that once the case of an L^∞ initial condition has been settled, the L^p case for $1 \leq p < +\infty$ may be addressed using this type of truncations. The task is however expected to be slightly more delicate than in the case of transport equations because the renormalized equation is not exactly the original equation applied to $\Phi(u)$: the second-order operator has generated the extra term $\frac{1}{2}\Phi''(u)|\nabla u|^2$. In that sense, our setting is closer to the case of the Fokker–Planck–Boltzmann-type equations considered in [40]. Note also that (2.3.1) is a *nonlinear formulation of a linear equation*.

Remark 16. The notion of renormalized solution allows to extend the notion of solution to less regular initial conditions (essentially only measurable initial conditions, see [41] for the details). It thus appears as a less demanding notion of solution. On the other hand, for a locally integrable initial condition, the notion is also, in essence, stronger than that of solution in the sense of distributions, or weak solution: under minimal assumptions on the coefficients, letting Φ approximate unity, one sees that a renormalized solution is *in particular* a solution, but a solution that enjoys additional properties. This is prototypical of a generic situation in the analysis of partial differential equations. One weakens the notion of solution, to allow for more generality. Existence of solution follows. Uniqueness, however, is then an issue, and one may have to “step back”, adding more requirements on the notion of solution so that, eventually, uniqueness holds. The notion of entropic solution for nonlinear conservation laws is one example. The notion of renormalized solution, here, is another example.

We now have to give a sense to the above formal equation (2.3.1). For this purpose we will have to put suitable conditions on the functions Φ. For our arguments of the previous sections to apply to $\Phi(u)$, we have to impose that $\Phi(u)$ is bounded and that $\nabla\Phi(u) \in L^2$. The former condition suggests that we take Φ a smooth bounded function, while the second, which also reads $\Phi'(u)\,\nabla u \in L^2$, suggests that Φ' has bounded support. We therefore typically think of functions Φ of the form of cut-off functions

$$\Phi_R(t) = \begin{cases} -R & \text{when } t < -R, \\ t & \text{when } -R + \eta \le t \le R - \eta, \\ R & \text{when } R < t, \end{cases} \tag{2.3.2}$$

for η arbitrarily small and with smooth transitions on $[-R, -R + \eta]$ and $[R - \eta, R]$.

We are now in a position to precisely define the notion of renormalized solutions we adopt.

Definition 2. Given some initial condition $u_0 \in L^p$, $1 \le p < +\infty$, we say that u is a renormalized solution to (1), that is,

$$\partial_t u - \mathbf{b}\,.\,\nabla u - \frac{1}{2}\Delta u = 0,$$

when for all smooth, real-valued, bounded functions Φ defined on $\mathbb{R}$, with compactly supported derivative Φ', the function $\Phi(u)$ solves (2.3.1), that is,

$$\partial_t \Phi(u) - \mathbf{b}\,.\,\nabla\Phi(u) - \frac{1}{2}\Delta\Phi(u) + \frac{1}{2}\Phi''(u)|\nabla u|^2 = 0$$

in the distributional sense, with initial condition $\Phi(u_0)$, together with the following conditions:

$$u \in C^0([0,T], L^p), \quad \Phi(u) \in L^2([0,T], H^1) \tag{2.3.3}$$

and

$$u^{\frac{p}{2}} \in L^2([0,T], H^1) \quad \text{for } p > 1, \tag{2.3.4}$$

$$\iint |\nabla u|^2 \mathbb{1}_{|u|<R} \le CR \quad \text{for } p \ge 1, \tag{2.3.5}$$

for all R and some constant C independent of R.

Remark 17. Of course, an analogous definition can be considered for other second-order operators than the Laplacian. In dimension 2, the assumptions to consider are those of Theorem 2. The adaptations to the settings of Corollary 1 and various other settings can also be performed. We leave these extensions to the reader.

Remark 18. In order to avoid any unnecessary technicalities in this pedagogic exposition, we will use in many instances throughout this section the notation ∇u *as usual*. However, it is important to note that this is only a *notation* for convenience, because ∇u is not necessarily locally integrable. In full generality, it only belongs to the space $L^0 = \{v : \text{meas}\{|v| > \lambda\} < +\infty \text{ for all } \lambda > 0\}$, cf. [41]. In the case $p > 1$, the situation is simpler. Then we have $u^{\frac{p}{2}} \in H^1$, and when we write $|u|^{\frac{p}{2}-1}\nabla u$, we mean $\frac{2}{p}\nabla(|u|^{\frac{p}{2}})$. Similarly, in any case, we have $\nabla\Phi(u)$ locally integrable (actually in L^2) and this is what our *notation* $\Phi'(u)\nabla u$ means. Otherwise, we have to define $T_R(u) = (u \wedge R) \vee (-R)$. We then consider $\nabla T_R(u)$, which we slightly abusively denote by $\nabla u \mathbb{1}_{|u|<R}$, as for instance in (2.3.5). On the meaning of ∇u, we refer to Lemma 7 below, the subsequent lemmata and the related comments we make there.

Before we proceed further, we would like to motivate properties (2.3.4)–(2.3.5) which we impose in our Definition 2 of a renormalized solution. When $p > 1$, we know from (2.1.18) that the solutions issued from L^p initial conditions are such that $u^{\frac{p}{2}} \in L^2([0,T], H^1)$. This is formal in general and we have shown it is rigorous when u_0 is also L^∞. This readily shows the property $\iint |\nabla u|^2 \mathbb{1}_{|u|<R} \le CR$. Indeed, if $u_0 \in L^p$, $p \ge 2$, then (say on a bounded domain) $u_0 \in L^2$ and in that case $\nabla u \in L^2$ so that this property is trivially satisfied, while, when $1 < p < 2$, we have

$$\iint |\nabla u|^2 |u|^{p-2} \ge \iint |\nabla u|^2 |u|^{p-2} \mathbb{1}_{|u|<R} \ge \frac{1}{R} \iint |\nabla u|^2 \mathbb{1}_{|u|<R}, \tag{2.3.6}$$

since $|u| < R$ and $-1 < p-2 < 0$ implies that $|u|^{p-2} > R^{p-2} > R^{-1}$. Estimate (2.3.5) also follows.

On the other hand, for $p = 1$, we cannot multiply the equation by $|u|^{p-2}u$ to obtain (2.1.18). To establish (2.3.5), we argue formally. The precise argument is in fact performed on regularized solutions; see the first step of the proof of Theorem 3 below. For each cut-off radius R, we define $T_R(u) = (u \wedge R) \vee (-R)$, which, without changing our notation, we mollify around the values R and $-R$. Multiplying (1) by $T_R(u)$ and denoting by

$$S_R(t) = \int_0^t T_R(s)\,ds,$$

a function that away from the mollifications at $\pm R$ essentially looks like

$$\begin{cases} \frac{|t|^2}{2} & \text{when } |t| \le R, \\ R|t| - \frac{R^2}{2} & \text{when } |t| \ge R, \end{cases}$$

we obtain

$$\frac{d}{dt}\int S_R(u) + \frac{1}{2}\int |\nabla T_R(u)|^2 - \int \mathbf{b}.\nabla S_R(u) = 0, \tag{2.3.7}$$

which is nothing else but a formalization of (1.3.3) for renormalized solutions. Notice that, intuitively, $S_R(u)$ and $T_R(u)$ play the roles of $\frac{u^2}{2}$ and u, respectively. To keep things simple, we assume $[\operatorname{div}\mathbf{b}]_- \in L^\infty$. The more general assumptions (2.1.2) will be discussed in Section 2.4. We therefore obtain

$$\int S_R(u) \le e^{Ct}\int S_R(u_0)$$

and

$$\iint |\nabla T_R(u)|^2 \le C_T \int S_R(u_0). \tag{2.3.8}$$

Since $S_R(u_0) \le R|u_0|$, $u_0 \in L^1$ and $\nabla T_R(u) = \nabla u \mathbb{1}_{|u|<R}$, (2.3.8) shows (2.3.5). This expresses that we do not necessarily have $\nabla u \in L^2$ but we know, in the worst case scenario, how the integral blows up.

The very important observation, which is crucial for establishing uniqueness of the renormalized solution (see Remark 21) is that, for all p and by construction, our renormalized solution satisfies a specific estimate over annular regions. Estimate (2.3.5) already gives information on the L^2 integrability of ∇u (in the case $p = 1$, the integral may diverge but, as we have just mentioned, (2.3.5) tells us how). The estimate we now prove refines this information:

Lemma 5. *For all* $1 \le p < +\infty$, *a renormalized solution of Definition 2 satisfies*

$$\lim_{R\to+\infty} \frac{1}{R}\iint |\nabla u|^2 \mathbb{1}_{\frac{R}{2}<|u|<R} = 0. \tag{2.3.9}$$

Remark 19. We notice that some conditions in the spirit of (2.3.5) or (2.3.9) were considered in [15] and related works.

Remark 20. For $\mathbf{b} = 0$ and the initial condition $u_0 = \delta_0$, the solution is of course the fundamental solution, that is, on the whole space to keep things simple, the Gaussian function (2.1.34). We then have

$$\nabla u(t,x) \propto \frac{|x|e^{-\frac{x^2}{2t}}}{t^{\frac{d}{2}+1}},$$

thus, up to constants in the calculations that are irrelevant for our purpose,

$$\begin{aligned}\int_0^T\int |\nabla u|^2 \mathbb{1}_{|u|\le R} &= \int_0^T \int_{t^{-\frac{d}{2}}e^{-\frac{|x|^2}{2t}}\le R} \frac{|x|^2 e^{-\frac{|x|^2}{t}}}{t^{d+2}}\,dx\,dt\\ &= \int |y|^2 e^{-|y|^2} \int_{t^{-\frac{d}{2}}e^{-\frac{|y|^2}{2}}\le R} \frac{1}{t^{\frac{d}{2}+1}}\,dt\,dy\\ &\propto R\int |y|^2 e^{-\frac{|y|^2}{2}}\,dy.\end{aligned}$$

It follows that (2.3.5) is true not only for $u_0 \in L^1$ but also, for instance, for δ_0. A similar calculation shows, on the other hand, that $\iint |\nabla u|^2 \mathbb{1}_{\frac{R}{2}<|u|<R} \propto R$ and thus u cannot satisfy (2.3.9). The L^1 integrability of the initial condition is therefore necessary to get (2.3.9).

Remark 21. In line with Remark 20, we would like to emphasize the essence of property (2.3.9). In order to hope for, and establish, uniqueness, we have to be able to discriminate between the solution of the equation for the initial condition $u_0 = 0$ and the solution for its "smallest perturbation" $u_0 = \delta_0$. As shown by Remark 20, condition (2.3.9) *in particular* completes this, since it discriminates between the solution for an initial condition in L^1_{loc} and the fundamental solution, obtained for δ_0.

Remark 22. An interesting unsolved question is to fully understand the case of an initial condition $u_0 \in \mathcal{M}_b$. Works by H. Osada, E. Carlen, V. Sverak, and notably works devoted to the two-dimensional Navier–Stokes equation, address some similar issues.

Proof of Lemma 5. Assume first that $p > 1$. Since by definition $u^{\frac{p}{2}} \in L^2([0,T], H^1)$, we have

$$\begin{aligned}C \ge \iint |\nabla u|^2|u|^{p-2} &\ge \iint |\nabla u|^2|u|^{p-2}\mathbb{1}_{\frac{R}{2}<|u|<R}\\ &\ge \min(1, 2^{2-p})R^{p-2}\iint |\nabla u|^2 \mathbb{1}_{\frac{R}{2}<|u|<R}\\ &\propto R^{p-1}\frac{1}{R}\iint |\nabla u|^2\mathbb{1}_{\frac{R}{2}<|u|<R}\end{aligned}$$

which, since $R^{p-1} \to +\infty$ with R, shows (2.3.9) in that case.

Consider now the slightly more technical case when $p = 1$. We define the function

$$\Phi_R(t) = \begin{cases} 0 & \text{when } t < \frac{R}{2},\\ (t-\frac{R}{2})^2 & \text{when } \frac{R}{2} \le t \le R,\\ \frac{Rt}{2} - \frac{R^2}{4} & \text{when } t > R,\end{cases}$$

suitably mollified around the values $\frac{R}{2}$ and R so that it is a smooth function which, with a slight abuse of notation, we again denote by Φ_R. We next extend the function to an even function defined on the whole line $\mathbb{R}$. Note that, essentially (and this is why we have constructed such a Φ_R), $\Phi_R'' = \mathbb{1}_{\frac{R}{2}<|t|<R}$. Up to irrelevant error terms caused by the above mollification and which can easily be eliminated after a passage to the limit, the integration of (2.3.1) yields

$$\frac{d}{dt}\int \Phi_R(u(t,\cdot)) + \frac{1}{2}\int \Phi_R''(u)|\nabla u|^2 \leq C\int \Phi_R(u(t,\cdot)), \tag{2.3.10}$$

where the constant C only depends on $\|\operatorname{div}\mathbf{b}\|_{L^\infty}$. An immediate Gronwall-type argument shows that, for all $t \in [0,T]$, $\int \Phi_R(u(t,\cdot)) \leq C_T \int \Phi_R(u_0)$ for a certain constant C_T, an information that we readily insert in (2.3.10) integrated from 0 to T to obtain

$$\int_0^T\int |\nabla u|^2 \mathbb{1}_{\frac{R}{2}<|u|<R} \leq C_T\int \Phi_R(u_0)$$

for another constant C_T. Using that $\Phi_R(|t|) \leq \frac{1}{2}R|t|\mathbb{1}_{\frac{R}{2}<|t|}$, this implies

$$\int_0^T\int |\nabla u|^2 \mathbb{1}_{\frac{R}{2}<|u|<R} \leq C_T\frac{R}{2}\int |u_0|\mathbb{1}_{\frac{R}{2}<|u_0|},$$

where we note that the rightmost integral vanishes as $R \to +\infty$ since $u_0 \in L^1$. This proves (2.3.9) in the case $p = 1$ and concludes the proof of Lemma 5. □

Theorem 3. *Under the same assumptions, in particular* (2.1.1)*, that is,* $\mathbf{b} \in L^2 + W^{1,1}$*, as those of Theorem 1* except
- *that we assume the initial condition is* L^p *for some* $1 \leq p < +\infty$*, instead of* L^∞*, and*
- *that we consider the simple assumption*

$$[\operatorname{div}\mathbf{b}]_- \in L^\infty, \tag{2.3.11}$$

instead of (2.1.2) *(see Remark 23 and Section 2.4 below),*

there exists a unique renormalized solution to (1) *in the sense of Definition 2. That solution additionally satisfies property* (2.3.9) *of Lemma 5.*

Remark 23. We have stated Theorem 3 under the simple assumption (2.3.11). The extensions of the result to assumptions of the form (2.1.2) will be examined in Section 2.4 after we complete the proof in that simple case.

Remark 24. As pointed out above, and as will be clear in the proof of Theorem 3, our proof of uniqueness of a renormalized solution makes an essential use of (2.3.9) but the proof of existence by convergence of some regularized solutions does not involve this property.

Remark 25. Definition 2 and Theorem 3 admit natural generalizations to the cases when the second-order operator is not the Laplacian, but $a_{ij}\partial_{ij}$ for a symmetric *constant* matrix $\mathbf{a}$. Depending upon whether $\mathbf{a}$ is definite positive or not, $\sqrt{\mathbf{a}}\nabla u$ or $\operatorname{Proj}_{\operatorname{Im}\mathbf{a}}\nabla u$ respectively replace ∇u in the statements and proofs. We skip these extensions here.

Proof of Theorem 3. As always in these lecture notes, we only outline the essential ingredients and steps of the proof. In this specific case, it turns out that a detailed proof has appeared in [70, Appendix E, p. 183], in the particular case when $\mathbf{b} \in L^2$ and $\operatorname{div} \mathbf{b} = 0$ (with $\mathbf{b} . \vec{n} = 0$ and Neumann boundary conditions on u). That proof is, of course, similar in its structure, its milestones and its key ingredients to the one we give here, although some details may vary. Also, in that reference, the various truncations (involving operators such as S_R and T_R defined as above on page 72) are notably made precise. For simplicity of exposition, and with a view to be as pedagogic as possible, we mainly omit these truncations here. Throughout the proof, when we write ∇u, we most of the time mean $\nabla T_R(u)$.

Consider the initial condition u_0 and a sequence u_0^ε of initial conditions in L^∞ that converges to u_0 strongly in L^p as ε vanishes. For ε fixed, we first establish some properties, all uniform in ε, of the sequence u^ε of solutions to (1) associated to those initial conditions. Our second step then consists in passing to the limit as $\varepsilon \to 0$, proving that u^ε converges and showing that the limit u of u^ε is a renormalized solution to (1) in the sense of Definition 2 for the initial condition u_0. Our final, third step establishes uniqueness of the renormalized solution.

Step 1: Existence and properties of a solution for a regularized initial condition. Since $u_0^\varepsilon \in L^\infty$, we may apply for each ε the results of the previous sections. We know there exists a unique solution to (1) with the properties stated in Theorem 1.

Case $p > 1$. We have the a priori estimate (2.1.18) on u^ε, that is,

$$\frac{d}{dt} \int |u^\varepsilon|^p + \underline{a} \int |u^\varepsilon|^{p-2} |\nabla u^\varepsilon|^2 \leq C \int |u^\varepsilon|^p, \tag{2.3.12}$$

for some constant C that is independent of ε. Integrating in time, this estimate in turn shows (2.1.3) on u^ε, that is,

$$\|u^\varepsilon(t,\cdot)\|_{L^p} \leq e^{C_0 t} \|u_0^\varepsilon\|_{L^p}. \tag{2.3.13}$$

Given that the initial condition $u_0^\varepsilon \in L^\infty$ strongly converges to u_0 in L^p, this shows that u^ε is bounded in $C([0,T], L^p)$, uniformly in ε. Actually, taking two indices ε and η, and considering by linearity the equation solved by $u^\varepsilon - u^\eta$, estimates (2.3.12) and (2.3.13), applied to the difference $u^\varepsilon - u^\eta$, respectively read as

$$\frac{d}{dt} \int |u^\varepsilon - u^\eta|^p + \underline{a} \int |u^\varepsilon - u^\eta|^{p-2} |\nabla u^\varepsilon - \nabla u^\eta|^2 \leq C \int |u^\varepsilon - u^\eta|^p \tag{2.3.14}$$

and

$$\|u^\varepsilon(t,\cdot) - u^\eta(t,\cdot)\|_{L^p} \leq e^{C_0 t} \|u_0^\varepsilon - u_0^\eta\|_{L^p}. \tag{2.3.15}$$

First, it follows from (2.3.15) that u^ε is a Cauchy sequence in $C([0,T], L^p)$ because u_0^ε converges, thus is a Cauchy sequence, in L^p. Hence, u^ε (strongly) converges to some u in $C([0,T], L^p)$. Next, we show from (2.3.14) that $|u^\varepsilon - u|^{\frac{p}{2}}$ converges to zero in $L^2([0,T], H^1)$. The convergence in $L^2([0,T], L^2)$ is immediate. For the convergence of the gradients, we first note that (2.3.14) integrated in time together with (2.3.15) to bound the right-hand side yields

$$\iint |\nabla(u^\varepsilon - u^\eta)^{\frac{p}{2}}|^2 \le C_T \int |u_0^\varepsilon - u_0^\eta|^p. \tag{2.3.16}$$

Likewise, the bounds provided by (2.3.12) and (2.3.13) show that, up to an extraction in ε, $\nabla(u^\varepsilon)^{\frac{p}{2}} \rightharpoonup \nabla v$ weakly in $L^2([0,T], L^2)$, but since we already know that u^ε itself strongly converges, we have $v = u^{\frac{p}{2}}$ and the whole sequence (and not only the extraction) converges. We now temporarily fix ε and let η vanish in (2.3.16). By the same argument, we have

$$\nabla \big|u^\varepsilon - u^\eta\big|^{\frac{p}{2}} \rightharpoonup \nabla |u^\varepsilon - u|^{\frac{p}{2}}$$

weakly in $L^2([0,T], L^2)$. Therefore

$$\begin{aligned}\iint \Big|\nabla |u^\varepsilon - u|^{\frac{p}{2}}\Big|^2 &\le \liminf \iint \Big|\nabla |u^\varepsilon - u^\eta\Big|^{\frac{p}{2}}|^2 \\ &\le C_T \liminf \int \Big|u_0^\varepsilon - u_0^\eta\Big|^p \\ &= \int |u_0^\varepsilon - u_0|^p.\end{aligned}$$

Letting now ε vanish, we obtain the convergence of $|u^\varepsilon - u|^{\frac{p}{2}}$ to zero in $L^2([0,T], H^1)$. The bounds on u^ε and this convergence in particular show (2.3.4).

We now establish bounds that will later show (2.3.5) for the solution u we construct. As mentioned above, the case $p \ge 2$ is easy thus we focus on the case $1 < p < 2$. Indeed arguing on u^ε as we formally did on u in (2.3.6), we obtain, for all $R > 0$,

$$\iint |\nabla u^\varepsilon|^2 \mathbb{1}_{|u^\varepsilon| < R} \le CR \tag{2.3.17}$$

with a constant C independent of ε. Similarly, manipulating the difference $u^\varepsilon - u^\eta$, letting η vanish, we remark, for ε fixed,

$$\begin{aligned}\liminf_{\eta\to 0} \iint \big|\nabla(u^\varepsilon - u^\eta)\big|^2 \big|u^\varepsilon - u^\eta\big|^{p-2} &\ge \iint |\nabla(u^\varepsilon - u)|^2 |u^\varepsilon - u|^{p-2} \\ &\ge \iint |\nabla(u^\varepsilon - u)|^2 \mathbb{1}_{|u^\varepsilon - u| < R},\end{aligned}$$

since $1 < p < 2$. We thus deduce from the convergences established above that, for all $R > 0$,

$$\lim_{\varepsilon\to 0} \iint |\nabla u^\varepsilon - \nabla u|^2 \mathbb{1}_{|u^\varepsilon - u| < R} = 0. \tag{2.3.18}$$

Using (2.3.17) and (2.3.18) will allow to establish (2.3.5).

Case $p = 1$. We have explained above that we may not outright multiply the equation to obtain the suitable estimates. We have to use truncations (see S_R and T_R defined on page 72). Apart from this technicality, the argument follows the same pattern and we will obtain (2.3.5).

Step 2: Passage to the limit and existence. As ε vanishes, we now know that u^ε converges to some u in the various senses made precise in Step 1. We have to prove that u is a renormalized solution to (1). This will show the existence of a renormalized solution for an L^p initial condition. As explained above, this existence is intuitively clear by "formal" interpolation. Proving it is the purpose of Step 2.

First, we note that

$$\partial_t \Phi(u^\varepsilon) - \frac{1}{2}\Delta\Phi(u^\varepsilon) - \mathbf{b}.\nabla\Phi(u^\varepsilon) + \frac{1}{2}\Phi''(u^\varepsilon)|\nabla u^\varepsilon|^2 = 0 \tag{2.3.19}$$

for all appropriate functions Φ, since the L^∞ solution u^ε is also a renormalized solution.

Passing to the limit $\varepsilon \to 0$ in the first three terms of (2.3.19), namely $\partial_t\Phi(u^\varepsilon)$, $\Delta\Phi(u^\varepsilon)$, $\mathbf{b}.\nabla\Phi(u^\varepsilon)$ is straightforward. Step 1 gives us strong convergence of $\Phi(u^\varepsilon)$ and all derivatives of that function are similarly taken care of using the theory of distributions. The point is the convergence of the term $\Phi''(u^\varepsilon)\,|\nabla u^\varepsilon|^2$. Note that we could equally well write the latter term $\Phi''(u^\varepsilon)\,|\nabla T_R(u^\varepsilon)|^2$ since, for R sufficiently large (at Φ fixed), those two expressions agree because of the properties of Φ and T_R. We now show this term converges in L^1. Of course, the difficult case is when $p < 2$, otherwise this convergence is clear. We first show the convergence in measure (that is, in probability) of ∇u^ε to ∇u. This convergence is a consequence of the following observation:

$$\begin{aligned}
\operatorname{meas}\{x : |\nabla u^\varepsilon - \nabla u| \geq \alpha\} &= \operatorname{meas}\{x : |\nabla u^\varepsilon - \nabla u| \geq \alpha,\ |u^\varepsilon - u| < R\} \\
&\qquad + \operatorname{meas}\{x : |\nabla u^\varepsilon - \nabla u| \geq \alpha,\ |u^\varepsilon - u| \geq R\} \\
&\leq \frac{1}{\alpha^2}\int |\nabla u^\varepsilon - \nabla u|^2\,\mathbb{1}_{|u^\varepsilon - u| < R} + \operatorname{meas}\{x : |u^\varepsilon - u| \geq R\} \\
&\leq \frac{1}{\alpha^2}\int |\nabla u^\varepsilon - \nabla u|^2\,\mathbb{1}_{|u^\varepsilon - u| < R} + \frac{1}{R^p}\|u^\varepsilon - u\|_{L^p}^p
\end{aligned} \tag{2.3.20}$$

for all ε, α and R. Integrating in time, we then respectively use (2.3.18) and the convergence of u^ε to u in $C([0,T], L^p)$ to treat the two terms and show that for all α, $\operatorname{meas}\{(t,x) : |\nabla u^\varepsilon - \nabla u| \geq \alpha\}$ vanishes with ε. Then uniform integrability is proved on the sequence $|\nabla T_R(u^\varepsilon)|^2$ (we explicitly reinstate here the truncation operator, otherwise the argument is obscure!), so that we obtain that this sequence converges in L^1 to $|\nabla T_R(u)|^2$, and $T_R(u^\varepsilon)$ therefore converges to $T_R(u)$ in $L^2([0,T], H^1)$. Briefly, the uniform integrability mentioned is established as follows. We recall that the only interesting case is $p < 2$, otherwise the result is clear. Since we already know that u^ε converges to u in $C([0,T], L^p)$ and that $|u^\varepsilon - u|^{\frac{p}{2}}$ converges to zero in $L^2([0,T], H^1)$, we may use the converse of the Lebesgue Dominated Convergence Theorem to claim that, up to an extraction we omit to mention, u^ε and $\nabla|u^\varepsilon - u|^{\frac{p}{2}}$ converge almost everywhere and are bounded from above by two fixed functions in L^p and L^2, respectively. We next consider $\int |\nabla T_R(u^\varepsilon)|^2\,\mathbb{1}_{|\nabla T_R(u^\varepsilon)| \geq K}$ and wish to show it is small for K large, uniformly in ε. But we notice that $|\nabla T_R(u^\varepsilon)|^2 \leq cR^{2-p}|\nabla(|u^\varepsilon|^{\frac{p}{2}})|^2$. The latter function may be bounded from above by a fixed function in L^1, and thus the uniform integrability is clear. On the other hand, already because of the strong convergence of u^ε to u established above, we also know, and we have already mentioned it for Φ but this is equally true for Φ'', that $\Phi''(u^\varepsilon)$ converges to $\Phi''(u)$ in L^1. We now just have to remark, taking $a^\varepsilon = |\nabla u^\varepsilon|^2$ and $b^\varepsilon = \Phi''(u^\varepsilon)$, that if $a^\varepsilon \to a$ in L^1, $b^\varepsilon \to b$ in L^1, and b^ε is bounded in L^∞, then $a^\varepsilon b^\varepsilon \to ab$ in L^1. This concludes the proof that $\Phi''(u^\varepsilon)|\nabla u^\varepsilon|^2$ converges to $\Phi''(u)|\nabla u|^2$ in L^1. All the terms of the renormalized formulation of the equation thus pass to the limit.

We know that, by construction, $u \in C^0([0,T], L^p)$ and, for $p > 1$, $u^{\frac{p}{2}} \in L^2([0,T], H^1)$ and $\Phi(u) \in L^2([0,T], H^1)$. This follows from our arguments of Step 1. Likewise, (2.3.5) holds true. This concludes the proof that u is a renormalized solution to (1) in the sense of Definition 2.

Step 3: Uniqueness. Uniqueness typically proceeds by subtraction of the two equations respectively solved by the two tentative solutions u_1 and u_2 and an argument on the difference $u_1 - u_2$. The technical difficulty here is that, since we manipulate a *nonlinear* formulation (2.3.1) of equation (1), the difference $u_1 - u_2$ of two renormalized solutions is not the solution of an equation of the same type as the original equation. More precisely, choosing a smooth function Φ as in Definition 2 which will be made precise below and denoting by $w = \Phi(u_1) - \Phi(u_2)$, we obtain

$$\partial_t w - \frac{1}{2}\Delta w - \mathbf{b}.\nabla w = \frac{1}{2}\Phi''(u_2)|\nabla u_2|^2 - \frac{1}{2}\Phi''(u_1)|\nabla u_1|^2. \tag{2.3.21}$$

We now choose a function γ such that (as usual, up to an appropriate mollification) $\gamma'(w) = (w \wedge K) \vee (-K)$. We multiply (2.3.21) by $\gamma'(w)$ and integrate to obtain

$$\begin{aligned} &\frac{d}{dt}\int \gamma(w) + \frac{1}{2}\int |\nabla w|^2 \mathbb{1}_{|w|\le K} \\ &\quad \le C\int \gamma(w) + \frac{1}{2}K\int \left|\Phi''(u_2)|\nabla u_2|^2 - \Phi''(u_1)|\nabla u_1|^2\right|, \end{aligned} \tag{2.3.22}$$

where the constant C only depends on the data (that is, in the simple case we consider here, $\|[\operatorname{div}\mathbf{b}]_-\|_{L^\infty}$). We now make the function Φ specific by taking $\Phi = \Phi_R = R\Phi_1(\frac{t}{R})$, where

$$\Phi_1(t) = \begin{cases} t & \text{when } t < 1, \\ 2 & \text{when } t > 2, \end{cases}$$

with a smooth monotonic transition between $t = 1$ and $t = 2$, and suitably symmetrized so that Φ_1 is an odd function of the real line. The purpose of this construction is to have

$$\left|\Phi_R''(t) = \frac{1}{R}\Phi_1''\left(\frac{t}{R}\right)\right| \le \frac{C}{R}\mathbb{1}_{R\le t\le 2R}.$$

This property is next used to bound from above the rightmost term of (2.3.22). Recalling estimate (2.3.9) which holds for both the renormalized solutions u_1 and u_2, we readily obtain that this term vanishes as $R \to +\infty$. It is then straightforward to conclude that $w = 0$, that is, $u_1 = u_2$. This concludes the proof of uniqueness and thus the proof of Theorem 3. □

Remark 26. The nature of the arguments in the above proof clearly shows that the point is to understand existence and uniqueness for *bounded* initial conditions, as we did in Theorem 1. The case of L^p $(1 \leq p < +\infty)$ initial conditions then follows using truncation, even though some technical details can be rather subtle.

Remark 27. All the estimates we establish in the course of the proof on the regularized solution u^ε (but *not* on its gradient) carry over to the case when the initial condition u_0 is a bounded measure. This remark is in echo to Remark 20.

2.4 Miscellaneous remarks

We are interested here in considering more general assumptions on $\mathbf{b}$ than the simple assumption $\|[\operatorname{div}\mathbf{b}]_-\|_{L^\infty}$ we have considered for establishing Theorem 3. The assumption we have in mind for the generalization is of course (2.1.2), which we reproduce here for convenience:

$$\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2, \quad \mathbf{b}_1 \in \varepsilon_1 L^{d,\infty} + L^d, \ [\operatorname{div}\mathbf{b}_2]_- \in \varepsilon_2 L^{\frac{d}{2},\infty} + L^{\frac{d}{2}}.$$

For simplicity we will work in dimensions $d \geq 3$ and leave the specific case of the two-dimensional setting to the reader.

Although the manipulations performed in the course of the proof of Theorem 3 do not themselves make use of the assumption $\|[\operatorname{div}\mathbf{b}]_-\|_{L^\infty}$, they rest on properties (2.3.4)–(2.3.5) and (2.3.9) of Lemma 5 satisfied by a renormalized solution. When establishing those properties, we have used in an essential way our assumption $\|[\operatorname{div}\mathbf{b}]_-\|_{L^\infty}$. This is the case in (2.3.7)

$$-\int \mathbf{b}\,.\,\nabla u\, T_R(u) = -\int (\operatorname{div}\mathbf{b}) S_R(u),$$

where we control the right-hand side using $\|[\operatorname{div}\mathbf{b}]_-\|_{L^\infty}$, and in similar manipulations, such as (2.3.10). When assuming (2.1.2), we have to revisit such parts of the proof.

To begin with, we consider the case $\mathbf{b} = \mathbf{b}_2$ with $[\operatorname{div}\mathbf{b}_2]_- \in \varepsilon_2 L^{\frac{d}{2},\infty} + L^{\frac{d}{2}}$. In that case, the argument we have performed above readily applies for both cases $p > 1$ and $p = 1$, up to minor technical modifications.

The case when we do not have any control on $\operatorname{div}\mathbf{b}$ is more delicate.

For $p > 1$, if we assume $\mathbf{b} = \mathbf{b}_1 \in \varepsilon_1 L^{d,\infty} + L^d$, the adaptations of our argument are easy. When $p \to 1$, our arguments, together with our methodological remarks on page 39, carry over to the setting of $L \ln L$.

The special case $p = 1$ requires more work. We are unable to conclude when $\mathbf{b} = \mathbf{b}_1 \in \varepsilon_1 L^{d,\infty} + L^d$ but we can conclude under the stronger assumption $\mathbf{b} = \mathbf{b}_1(x) \in \varepsilon_1 L^{d,1}$ with $\mathbf{b}_1$ independent of time. In fact, we may, as in many places in these notes, add to this $\mathbf{b}_1$ a regular part – more specifically an L^∞ part here – and also allow for a time dependence of $\mathbf{b}_1$ if this dependency is regular – continuous here. The essential ingredient we need is the following lemma.

Lemma 6. *We assume that* $\mathbf{b} = \mathbf{b}(x)$ *is independent of time and belongs to* $L^{d,1}$. *We consider* $\partial_t u - \mathbf{b}.\nabla u - \frac{1}{2}\Delta u = 0$. *Then*

$$\|\mathbf{b}.\nabla u\|_{L^1_{t,x}} \le C\|\mathbf{b}\|_{L^{d,1}} \left\|\partial_t u - \frac{1}{2}\Delta u\right\|_{L^1_{t,x}}. \tag{2.4.1}$$

If we assume that $\mathbf{b} = \mathbf{b}_1(x) \in \varepsilon_1 L^{d,1}$ with $\mathbf{b}_1$ independent of time, and if we temporarily admit Lemma 6, then the adaptation of our arguments is easy. We observe that

$$\left\|\partial_t u - \frac{1}{2}\Delta u\right\|_{L^1_{t,x}} = \|\mathbf{b}.\nabla u\|_{L^1_{t,x}} \le C\|\mathbf{b}\|_{L^{d,1}} \left\|\partial_t u - \frac{1}{2}\Delta u\right\|_{L^1_{t,x}} \le C\varepsilon_1 \left\|\partial_t u - \frac{1}{2}\Delta u\right\|_{L^1_{t,x}},$$

successively using the equation, Lemma 6 and our assumption on $\mathbf{b}_1$. Choosing ε_1 sufficiently small (in function of the (universal) constant C from (2.4.1) in Lemma 6), we obtain $\partial_t u - \frac{1}{2}\Delta u \in L^1_{t,x}$, and readily insert this information in (2.4.1) to have $\mathbf{b}.\nabla u \in L^1_{t,x}$. The rightmost term of (2.3.7) (once integrated in time from 0 to T) is then estimated as follows:

$$\left|\int_0^T \int \mathbf{b}.\nabla S_R(u)\right| = \left|\int_0^T \int \mathbf{b}.\nabla u \,.\, T_R(u)\right| \le \|\mathbf{b}.\nabla u\|_{L^1_{t,x}} \|T_R(u)\|_{L^\infty_{t,x}} = \|\mathbf{b}.\nabla u\|_{L^1_{t,x}}.R.$$

We therefore deduce from (2.3.7) that

$$\int S_R(u)(T) + \int_0^T \int |\nabla T_R(u)|^2 \le \int S_R(u_0) + \|\mathbf{b}.\nabla u\|_{L^1_{t,x}} R = O(R)$$

and then proceed with the proof of estimate (2.3.5). A similar argument applies to (2.3.10) and the proof of Lemma 5.

We now turn to the:

Proof of Lemma 6. In order to prove (2.4.1), we observe it is sufficient to prove that the solution p to the heat equation

$$\partial_t p - \frac{1}{2}\Delta p = \delta_{x_0,t_0}$$

satisfies

$$\|\mathbf{b}.\nabla p\|_{L^1_{t,x}} \le C\|\mathbf{b}\|_{L^{d,1}},$$

and next (2.4.1) will follow by superposition of such solutions p. To this end, we first argue on the whole space and observe that

$$\|\mathbf{b}.\nabla p\|_{L^1_{t,x}} = \int |\mathbf{b}(x)| \left(\int_0^\infty |\nabla p|\, dt\right) dx \le \|\mathbf{b}\|_{L^{d,1}} \left\|\int_0^\infty |\nabla p|\, dt\right\|_{L^{d/d-1,\infty}},$$

where the rightmost norm is finite because

$$\int_0^\infty |\nabla p|\, dt \propto \int_0^\infty \frac{|x| e^{-\frac{|x|^2}{2t}}}{t^{\frac{d}{2}+1}} \propto \frac{1}{|x|^{d-1}} \int_0^\infty \frac{e^{-\frac{1}{2s}}}{s^{\frac{d}{2}+1}}\, ds.$$

Note that a similar argument was used in (2.1.38) in Section 2.1.3. When we work on a bounded domain with periodic boundary conditions, we adapt the argument as we did in the proof of Lemma 2. □

Our next lemma (and the subsequent lemmata for some related cases) shows that, in the case we consider (L^1 initial condition and $\mathbf{b}.\nabla u \in L^1_{t,x}$), more information on the solution u and its gradient ∇u is available *within* the equation. In particular, ∇u makes sense as a locally integrable function, and this needs to be put in perspective with our comments of Remark 18 on page 72.

Lemma 7. *Assume that* $\partial_t u - \frac{1}{2}\Delta u \in L^1([0,T], L^1)$ *and* $u(t=0,\cdot) = u_0 \in L^1$. *Then:*

- *u satisfies*
$$u \in L^{\alpha,\infty}([0,T], L^{\beta,1}) \tag{2.4.2}$$
for all $\alpha > 1$, $\beta \geq 1$ *such that* $\frac{2}{d}\frac{1}{\alpha} + \frac{1}{\beta} = 1$, *and in particular, for* $\alpha = \beta = \frac{d+2}{d}$, *we have*
$$u \in L^{\frac{d+2}{d},\infty}([0,T], L^{\frac{d+2}{d},1}).$$
- *∇u satisfies*
$$\nabla u \in L^{\alpha,\infty}([0,T], L^{\beta,1}) \tag{2.4.3}$$
for all $\alpha > 1$, $\beta \geq 1$ *such that* $\frac{2}{d+1}\frac{1}{\alpha} + \frac{d}{d+1}\frac{1}{\beta} = 1$, *and in particular, for* $\alpha = \beta = \frac{d+2}{d+1}$, *we have*
$$u \in L^{\frac{d+2}{d+1},\infty}([0,T], L^{\frac{d+2}{d+1},1}).$$

Proof. We begin by providing an intuition for (2.4.2). Formally, we know from the equation that $u \in L^\infty([0,T], L^1)$. In addition, and although this fact does no make sense, everything happens as if we had $u \in L^{1,\infty}([0,T], L^{\frac{d}{d-2},1})$. Assuming we indeed had that information, we would write
$$\begin{cases} \dfrac{1}{\alpha} = \dfrac{\theta}{1} + \dfrac{1-\theta}{\infty}, \\ \dfrac{1}{\beta} = \dfrac{d-2}{d}\theta + \dfrac{1-\theta}{1} \end{cases}$$
for $\theta \in (0,1)$ and the Hölder inequality would then give (2.4.2). Although the above intuitive argument is incorrect, its conclusion (2.4.2) holds true. As for the previous lemma, the actual, rigorous proof of (2.4.2) makes use of the fact that since the right-hand side of the equation is assumed to be L^1 in both space and time, proving the estimate amounts to proving it for the fundamental solution (that is, the Green function, or in that particular case, the heat kernel) p of the equation, and then using the superposition principle. Otherwise stated, once p belongs to the suitable functional space, we use that the solution u writes
$$u = p \star_{t,x} f + p \star_x u_0$$
(with the obvious notation $\star_{t,x}$ and $\star_x$ for the convolution in both the time and space variables and only the space variable, respectively) and apply the convolution estimate (2.0.12) to get the result on u itself. The heat kernel p is defined in (2.1.34) (with

the usual adaptations when we work on a bounded domain, see Remark 28). Since p is nonincreasing in $|x|$ and

$$p(t,x) \propto \frac{e^{-\frac{x^2}{2t}}}{t^{\frac{d}{2}}},$$

we have, given the definition (2.0.4) of the Lorentz spaces,

$$\|p\|_{L^{\beta,1}} \propto \int_0^\infty \sigma^{\frac{1}{\beta}-1} \frac{e^{-\frac{1}{2t}\sigma^{\frac{2}{d}}}}{t^{\frac{d}{2}}}\, d\sigma \propto t^{\frac{d}{2}(\frac{1}{\beta}-1)} \int_0^\infty s^{\frac{d}{2\beta}-1} e^{-\frac{s}{2}}\, ds.$$

We next notice that, for $\alpha > 1$, this function belongs to $L^{\alpha,\infty}$ in the time variable exactly when $\alpha\frac{d}{2}(\frac{1}{\beta}-1) = -1$, that is, $\frac{2}{d}\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

The second assertion (2.4.3) of Lemma 7 is proved similarly. We argue on the *gradient* of the heat kernel, which is not a nonincreasing function but we can formally proceed as if it were nonincreasing:

$$\|\nabla p\|_{L^{\beta,1}} \propto \int_0^\infty \sigma^{\frac{1}{\beta}-1} \frac{1}{t} \frac{e^{-\frac{1}{2t}\sigma^{\frac{2}{d}}}}{t^{\frac{d}{2}}}\, d\sigma \propto t^{\frac{d}{2}(\frac{1}{\beta}-1)-\frac{1}{2}} \int_0^\infty s^{\frac{d}{2\beta}-\frac{1}{2}} e^{-\frac{s}{2}}\, ds$$

belongs, for $\alpha > 1$, to $L^{\alpha,\infty}$ in the time variable exactly when $\frac{d}{2}(\frac{1}{\beta}-1) - \frac{1}{2} = -\frac{1}{\alpha}$, that is, $\frac{2}{d+1}\frac{1}{\alpha} + \frac{d}{d+1}\frac{1}{\beta} = 1$. □

Remark 28. The above proofs have made use of the explicit form of the heat kernel, that is, the solution to $\partial_t p - \frac{1}{2}\Delta p = \delta$ *on the whole space* $\mathbb{R}^d$. We have already seen, in the proof of Lemma 2, that the argument is easily adapted when the equation is posed on a bounded domain with periodic boundary conditions. Actually, all types of boundary conditions can be accommodated, and the results of the above lemmata hold true, even if the technical details of the proofs change. The easiest adaptation is that for a bounded domain with homogeneous Dirichlet boundary conditions because then, by monotonicity, the Green function is dominated by the heat kernel and the proof above readily applies.

Remark 29. In sharp contrast to the (easy) question of alternate boundary conditions examined in Remark 28, the question of considering another second-order operator than the Laplacian is significantly more complicated. It is well known that, when $\operatorname{div}(\sigma(x)\sigma^t(x)\nabla u)$ – with the usual conditions on $\sigma\sigma^t$ for positive definiteness – replaces Δu, the fundamental solution $G(t,x,y)$ pointwise behaves like the heat kernel $p(t, x-y)$ in the limits $|x-y| \to 0$ or $+\infty$. This is however *not* true for its gradient. The above proofs can therefore not apply immediately. Additional assumptions, such as, e.g., regularity and periodicity, allow to proceed. Or, in generality, local averages of the Green function, instead of its pointwise values, need to be considered. The *results* obtained above therefore still hold true under convenient conditions.

We now conclude this section with two lemmata that indeed show that some further integrability information, both on u and on ∇u, is implicitly contained in an estimate such as (2.3.5). The results below allow for the same type of integrability on the solution u and its gradient ∇u as above, but for a general second-order operator, *provided* we consider a solution in the renormalized sense with property (2.3.5).

Lemma 8. *In dimension $d \geq 3$, we have, for all smooth, compactly supported functions u:*

$$\|u\|_{L^{d/(d-2),\infty}} \leq C \sup_R \frac{1}{R} \int |\nabla T_R(u)|^2. \tag{2.4.4}$$

Proof. Fix $R_0 > 0$. We apply the Sobolev inequality to $T_{2R_0}(u)$:

$$\|T_{2R_0}(u)\|^2_{L^{2d/(d-2)}} \leq \|\nabla T_{2R_0}(u)\|^2_{L^2}.$$

The left-hand side is bounded from below as follows:

$$\begin{aligned}\|T_{2R_0}(u)\|^2_{L^{2d/(d-2)}} &\geq \Bigg(\int\limits_{|T_{2R_0}(u)| \geq R_0} |T_{2R_0}(u)|^{\frac{2d}{d-2}} \Bigg)^{d-\frac{2}{d}} \\ &= \Bigg(\int\limits_{|u| \geq R_0} |T_{2R_0}(u)|^{\frac{2d}{d-2}} \Bigg)^{d-\frac{2}{d}} \\ &\geq R_0^2 \operatorname{meas}\{|u| \geq R_0\}^{d-\frac{2}{d}},\end{aligned}$$

where we have used that $\{x : |T_{2R_0}(u(x))| \geq R_0\} = \{x : |u(x)| \geq R_0\}$ since $2R_0 \geq R_0$. The right-hand side is now bounded from above:

$$\|\nabla T_{2R_0}(u)\|^2_{L^2} \leq 2R_0 \sup_R \frac{1}{R} \int |\nabla T_R(u)|^2.$$

Collecting all this, we obtain

$$R_0 \operatorname{meas}\{|u| \geq R_0\}^{d-\frac{2}{d}} \leq 2 \sup_R \frac{1}{R} \int |\nabla T_R(u)|^2,$$

which, taking the supremum in R_0 of the left-hand-side and using the definition (2.0.7), yields a minoration by the $L^{\frac{d}{d-2},\infty}$ quasi-norm. □

Lemma 9. *In dimension $d \geq 3$, we have*

$$\|\nabla u\|_{L^{(d+2)/(d+1),\infty}_{t,x}} \leq C\Big(\|u\|_{L^{(d+2)/d,\infty}_{t,x}} + \sup_R \frac{1}{R} \iint |\nabla T_R(u)|^2 \Big). \tag{2.4.5}$$

Proof (sketch). For simplicity, we sketch the proof of

$$\|\nabla u\|_{L^q_{t,x}} \leq C\Big(\|u\|_{L^{(d+2)/d,\infty}_{t,x}} + \sup_R \frac{1}{R} \iint |\nabla T_R(u)|^2 \Big) \tag{2.4.6}$$

for any exponent $q < \frac{d+2}{d+1}$ instead of that of estimate (2.4.5) itself. The proof of estimate (2.4.5) is an adaptation of the proof below, manipulating measures of sets of the type $\{x : |\nabla u(x)| > s\}$ instead of integrals of $|\nabla u|^q$, in the spirit of what we have performed on page 30 to prove (2.0.15).

To establish (2.4.6), we split $\iint |\nabla u|^q \, dx \, dt$ into dyadic annular regions. For every $n \in \mathbb{N}$, we fix a constant $\lambda > 0$ (to be chosen shortly) and notice that

$$\begin{aligned}
&\iint \mathbb{1}_{2^n\lambda<|u|\leq 2^{n+1}\lambda} |\nabla u|^q \, dx \, dt \\
&\quad \leq \Big(\iint \mathbb{1}_{|u|\leq 2^{n+1}\lambda} |\nabla u|^2 \, dx \, dt \Big)^{\frac{q}{2}} (\mathrm{meas}\{|u| \geq 2^n\lambda\})^{1-\frac{q}{2}} \\
&\quad \leq C(2^n\lambda)^{\frac{q}{2}} \Big(\sup_R \frac{1}{R} \iint |\nabla T_R(u)|^2 \Big)^{\frac{q}{2}} (2^{-n}\lambda^{-1})^{(1-\frac{q}{2})\frac{d+2}{d}} \Big(\|u\|^{\frac{d+2}{d}}_{L^{(d+2)/d,\infty}_{t,x}} \Big)^{1-\frac{q}{2}} \\
&\quad = C(2^n\lambda)^{\frac{d+1}{d}(q-\frac{d+2}{d+1})} \Big(\sup_R \frac{1}{R} \iint |\nabla T_R(u)|^2 \Big)^{\frac{q}{2}} \Big(\|u\|^{\frac{d+2}{d}}_{L^{(d+2)/d,\infty}_{t,x}} \Big)^{1-\frac{q}{2}}
\end{aligned}$$

successively using the Hölder inequality and the definition of the quasinorm of the Lorentz spaces, and where the constant C does not depend on n, λ, u. Summing up these inequalities for all $n \in \mathbb{N}$, and noticing that the series in power of 2 converges since we have assumed $q - \frac{d+2}{d+1} < 0$, we obtain

$$\iint_{|u|>\lambda} |\nabla u|^q \, dx \, dt \leq C\lambda^{\frac{d+1}{d}(q-\frac{d+2}{d+1})} \Big(\sup_R \frac{1}{R} \iint |\nabla T_R(u)|^2 \Big)^{\frac{q}{2}} \Big(\|u\|^{\frac{d+2}{d}}_{L^{(d+2)/d,\infty}_{t,x}} \Big)^{1-\frac{q}{2}}.$$

We notice $\frac{d+1}{d}(q - \frac{d+2}{d+1}) = \frac{q}{2} - (1 - \frac{q}{2})\frac{d+2}{d}$, so we now choose $\lambda = \|u\|_{L^{(d+2)/d,\infty}_{t,x}}$ and obtain

$$\iint_{|u|>\|u\|_{L^{(d+2)/d,\infty}_{t,x}}} |\nabla u|^q \, dx \, dt \leq C \Big(\sup_R \frac{1}{R} \iint |\nabla T_R(u)|^2 \Big)^{\frac{q}{2}} \big(\|u\|_{L^{(d+2)/d,\infty}_{t,x}} \big)^{\frac{q}{2}}. \tag{2.4.7}$$

On the other hand, we have

$$\begin{aligned}
\iint_{|u|<\lambda} |\nabla u|^q \, dx \, dt &\leq \lambda^{\frac{q}{2}} \Big(\sup_R \frac{1}{R} \iint |\nabla T_R(u)|^2 \Big)^{\frac{q}{2}} (\mathrm{meas}\{|u| < \lambda\})^{1-\frac{q}{2}}, \\
&\leq C\lambda^{\frac{q}{2}} \Big(\sup_R \frac{1}{R} \iint |\nabla T_R(u)|^2 \Big)^{\frac{q}{2}}
\end{aligned} \tag{2.4.8}$$

successively using the Hölder inequality and the fact that we work on a bounded domain. Applying (2.4.8) to $\lambda = \|u\|_{L^{(d+2)/d,\infty}_{t,x}}$, summing up (2.4.7) and (2.4.8), putting the sum to the power $\frac{1}{q}$ and using the Young inequality gives (2.4.6). □

3 Equations in divergence form

We now proceed to cases when the coefficient matrix $\mathbf{a}$ in (1) is not constant. We devote the present chapter to the study of the equations of the general form (1) in the particular case when the second-order operator is in divergence form. Chapter 4 will address equations in non-divergence form. More precisely, and up to a slight abuse of notation with respect to (1), we consider

$$\partial_t u - b_i \partial_i u - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j u) = 0 \tag{3.0.1}$$

for a certain $d \times k$ matrix-valued function σ. We will exclusively consider the case when the initial condition u_0 we supply (3.0.1) with is L^∞. The modifications of our arguments necessary in order to address an unbounded initial condition will then directly follow from the arguments we have developed in Section 2.3. We note indeed that the two procedures, respectively renormalization and regularization, are performed sequentially, in this order. Therefore the details of the regularization we provide in the present section in order to address the presence of a *varying* coefficient matrix $\sigma\sigma^t$ can be applied to the specific setting of a renormalized equation. The complete proof is then technical and somewhat tedious, but not more difficult conceptually than the case studied here. The results in the renormalized setting are obtained by a straightforward adaptation of the results stated below.

Our study of this setting has been initiated in [65]. In Sections 3.1 and 3.2, we will essentially summarize the main results obtained there, along with an outline of the major ingredients of the proof. We next turn to various extensions of the results obtained in [65] in Section 3.5. In the latter section, we will in particular explore the link between the questions we examine and the theory of hypo-ellipticity. Note that the consideration of the probabilistic setting is postponed until Section 4.2.

Before we begin, we note that the arguments we develop throughout this section also apply mutatis mutandis to the equation

$$\partial_t u - b_i \partial_i u - \partial_i(a_{ij}\partial_j u) = 0 \tag{3.0.2}$$

for some symmetric matrix coefficient $\mathbf{a}$ that satisfies, in the sense of symmetric matrices,

$$\underline{\mu}\sigma\sigma^t \le \mathbf{a} \le \overline{\mu}\sigma\sigma^t \tag{3.0.3}$$

for some $d \times k$ matrix-valued function σ, and some positive finite constants $\underline{\mu}$ and $\overline{\mu}$. Notice also that property (3.0.3) is *equivalent* to

$$\mathbf{a} = \sigma C \sigma^t \quad \text{for } C \text{ bounded, symmetric } k \times k\text{-matrix-valued function such that } \underline{\mu} I \le C \le \overline{\mu} I. \tag{3.0.4}$$

The equivalence is easy to establish (see below). One should also note that C in (3.0.4) is only assumed bounded. We will shortly consider H^1 regular functions σ, and therefore

https://doi.org/10.1515/9783110635508-003

not all matrices of the form (3.0.4) are reducible to the form $\sigma\sigma^t$ with $\sigma \in H^1$. In this sense, the present section therefore covers the whole generality of a second-order operator $\partial_i(a_{ij}\partial_j u)$ in divergence form for *any symmetric nonnegative matrix* $\mathbf{a}$, as announced in the introduction: it suffices to choose $\sigma = \sqrt{\mathbf{a}}$. The point is, the sequel will require *regularity* on σ. In this respect, our assumption (3.0.3) (or (3.0.4)) is thus much more general than $\mathbf{a} = (\sqrt{\mathbf{a}})(\sqrt{\mathbf{a}})^t$. The specific writing used in (3.0.1) also better connects with the probability theoretic setting, where a particular format $n \times k$ of the matrix σ is fixed.

To show the above claimed equivalence between (3.0.4) and (3.0.3), we first note that (3.0.4) clearly implies (3.0.3). Conversely, we consider the following (obviously) bilinear symmetric form defined on the space $\operatorname{Im}\sigma^t$ by $(Y_1, Y_2)_{\mathbf{a}} = (\mathbf{a}X_1, X_2)$ whenever $Y_i = \sigma^t X_i$, $i = 1, 2$. This form is well-defined and continuous, because of (3.0.3). Note that $(\mathbf{a}X_0, X_0) \le \overline{\mu}(\sigma^t X_0, \sigma^t X_0)$, which proves both the well-posedness and the continuity. Consequently, this form can be represented by some nonnegative symmetric matrix C, which, up to a completion, satisfies the properties stated in (3.0.4). These properties in turn allow for our proof (performed below in the case $\mathbf{a} = \frac{1}{2}\sigma\sigma^t$) to carry over to the case (3.0.3) (or, equivalently, (3.0.4)). We refer the reader to Remark 33 for an outline of the necessary adaptations.

3.1 Possibly degenerate diffusion matrices

We recall that, for simplicity of exposition, we work throughout these notes in the simple setting of a bounded domain with periodic boundary conditions, and we generically consider time-independent coefficient $\mathbf{a}$ and $\mathbf{b}$. In [65], the case of the whole space, with appropriate growth conditions at infinity on $\mathbf{a}$ and $\mathbf{b}$, together with the possible dependency upon time, has been explicitly considered. We begin, as usual, with the natural formal a priori estimate derived from (3.0.1) upon multiplying the equation by u and integrating over the domain:

$$\frac{d}{dt}\int \frac{u^2}{2} + \int \frac{u^2}{2}\operatorname{div}\mathbf{b} + \frac{1}{2}\int |\sigma^t \nabla u|^2 = 0. \tag{3.1.1}$$

Assuming $\operatorname{div}\mathbf{b} \in L^\infty$, or even only

$$[\operatorname{div}\mathbf{b}]_- \in L^\infty, \tag{3.1.2}$$

we obtain $u \in L^\infty([0,T], L^2)$, $\sigma^t\nabla u \in L^2([0,T], L^2)$. Using the maximum principle, we also (formally) get an L^∞ bound on u. Therefore, for an initial condition $u_0 \in L^\infty$, the solution u is expected to belong to the space

$$X = \{u \in L^\infty([0,T], L^\infty) \cap C^0([0,T], L^p),\ 1 \le p < +\infty,\ \sigma^t\nabla u \in L^2([0,T], L^2)\}. \tag{3.1.3}$$

We temporarily admit that this space X has all the suitable properties for our arguments to make sense. Because the space X defined above in (3.1.3) will play an ubiquitous role in the next sections, we will study in more details some of its properties in Remark 32 right after the proof of our Proposition 1 below. For the time being, we cannot emphasize enough the key point: *even is σ is degenerate*, the gradient ∇u of the solution is expected to be controlled exactly in the directions $\sigma^t \nabla u$ needed for the proof.

We notice that an L^p estimate may also be obtained upon multiplying the equation by $\beta'(u)$ and proceeding classically, thus the solution is also expected to be continuous in time with values in L^p for all $1 \le p < +\infty$. We refer to [65] for the detail of such easy manipulations.

In view of the formal a priori estimate (3.1.1) above, it is intuitive (and indeed true) that the following result holds:

Proposition 1 (Equation in divergence form, [65, Proposition 1]). *Assume that* $\mathbf{b}$ *and* σ *are (for simplicity of exposition, we recall) time-independent and satisfy*

$$\mathbf{b} \in W^{1,1}, \quad [\operatorname{div} \mathbf{b}]_- \in L^\infty, \tag{3.1.4}$$

$$\sigma \in H^1. \tag{3.1.5}$$

Then, for all initial conditions $u_0 \in L^\infty$, (1.3.5) *has a unique solution in the space*

$$\{u \in L^\infty([0,T], L^\infty) \cap C^0([0,T], L^p),\ 1 \le p < +\infty,\ \sigma^t \nabla u \in L^2([0,T], L^2)\}.$$

Remark 30. Assumption (3.1.5) is necessary for the regularization we perform in the proof. Another reason for requiring this condition is explained in Remark 32 below.

Remark 31. The L^p estimate (2.1.3) is also true, using the divergence form and manipulations similar to (2.1.17)–(2.1.18) on the sequence of approximate solutions. Similarly, the maximum principle (2.1.4) holds.

Proof of Proposition 1 (outline). As said above, the proof of Proposition 1 is given in details in [65]. It is actually performed therein in the case of the whole space, so that all the details related to the growth at infinity of the coefficients and how they relate to the estimation of the tails of integrals are dealt with. The case of time-dependent coefficients is also discussed. For the sake of consistency, we now only reproduce the outline of the crucial regularization step.

Convoluting (1.3.5) with some regularizing kernel

$$\rho_\varepsilon = \varepsilon^{-N} \rho(\varepsilon^{-1} \cdot) \quad \text{with } \rho \in \mathcal{D}(\mathbb{R}^N),\ \rho \ge 0,\ \int \rho = 1,$$

we obtain

$$\partial_t \rho_\varepsilon \star u - \rho_\varepsilon \star (b_i \partial_i u) - \frac{1}{2} \rho_\varepsilon \star \partial_i (\sigma_{ik} \sigma_{jk} \partial_j u) = 0. \tag{3.1.6}$$

We denote by

$$[\rho_\varepsilon, c](f) = \rho_\varepsilon \star (cf) - c(\rho_\varepsilon \star f)$$

for a differential operator c, and $u_\varepsilon = \rho_\varepsilon \star u$. We readily note that

$$\rho_\varepsilon \star (b_i \partial_i u) = Q_\varepsilon + b_i \partial_i u_\varepsilon, \tag{3.1.7}$$

where we have defined

$$Q_\varepsilon = [\rho_\varepsilon, b_i\partial_i](u). \tag{3.1.8}$$

Likewise,

$$\begin{aligned}
\rho_\varepsilon \star \partial_i(\sigma_{ik}\sigma_{jk}\partial_j u) &= \partial_i(\rho_\varepsilon \star (\sigma_{ik}\sigma_{jk}\partial_j u))\\
&= \partial_i([\rho_\varepsilon, \sigma_{ik}\sigma_{jk}\partial_j](u)) + \partial_i(\sigma_{ik}\sigma_{jk}\partial_j u_\varepsilon)\\
&= \partial_i(\sigma_{ik}[\rho_\varepsilon, \sigma_{jk}\partial_j](u) + [\rho_\varepsilon, \sigma_{ik}](\sigma_{jk}\partial_j u)) + \partial_i(\sigma_{ik}\sigma_{jk}\partial_j u_\varepsilon)\\
&= \partial_i(\sigma_{ik}[\rho_\varepsilon, \sigma_{jk}\partial_j](u)) + [\rho_\varepsilon, \partial_i\sigma_{ik}](\sigma_{jk}\partial_j u)\\
&\qquad + [\rho_\varepsilon, \sigma_{ik}\partial_i](\sigma_{jk}\partial_j u) + \partial_i(\sigma_{ik}\sigma_{jk}\partial_j u_\varepsilon)\\
&= \partial_i(\sigma_{ik}R_{k,\varepsilon}) + S_\varepsilon + T_\varepsilon + \partial_i(\sigma_{ik}\sigma_{jk}\partial_j u_\varepsilon),
\end{aligned} \tag{3.1.9}$$

where we have set

$$\begin{cases}
R_{k,\varepsilon} = [\rho_\varepsilon, \sigma_{jk}\partial_j](u),\\
S_\varepsilon = [\rho_\varepsilon, \partial_i\sigma_{ik}](\sigma_{jk}\partial_j u),\\
T_\varepsilon = [\rho_\varepsilon, \sigma_{ik}\partial_i](\sigma_{jk}\partial_j u).
\end{cases} \tag{3.1.10}$$

We thus obtain from (3.1.6) equation (3.0.1) set on u_ε instead of u, but with an error term on the right-hand side:

$$\partial_t u_\varepsilon - b_i\partial_i u_\varepsilon - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j u_\varepsilon) = Q_\varepsilon + \frac{1}{2}(\partial_i(\sigma_{ik}R_{k,\varepsilon}) + S_\varepsilon + T_\varepsilon). \tag{3.1.11}$$

Our main tool is now the following commutation result (already mentioned in a simpler form in Lemma 1):

Lemma 10 (see, e.g., [41, Lemma II.1]). *Let $f \in L^r$ and $\mathbf{c} \in W^{1,\alpha}$. Set $\frac{1}{\beta} = \frac{1}{r} + \frac{1}{\alpha}$. Then, locally, as $\varepsilon \to 0$,*

$$[\rho_\varepsilon, \mathbf{c}.\nabla](f) \to 0 \quad \text{in } L^\beta \tag{3.1.12}$$

and

$$[\rho_\varepsilon, \operatorname{div}\mathbf{c}](f) \to 0 \quad \text{in } L^\beta. \tag{3.1.13}$$

If $\nabla f \in L^2$ and $\mathbf{c} \in L^\alpha$, then the same conclusion holds for $\frac{1}{\beta} = \frac{1}{2} + \frac{1}{\alpha}$.

Using the above lemma, we immediately see that, when $\mathbf{b} \in W^{1,1}$, we have the convergence to zero of the first-order error term

$$Q_\varepsilon = [\rho_\varepsilon, b_i\partial_i](u) \xrightarrow{\varepsilon\to 0} 0 \quad \text{in } L^1.$$

This convergence holds uniformly in time, because u is L^∞ and $\mathbf{b}$ is time independent. This term, already present in the classical argument for the transport equation in [41], is the standard term in (3.1.11). We now turn our attention to the, non-standard, second order term. Applying again the above lemma, we have for all k,

$$R_{k,\varepsilon} = [\rho_\varepsilon, \sigma_{jk}\partial_j](u) \xrightarrow{\varepsilon\to 0} 0 \quad \text{in } L^2 \tag{3.1.14}$$

as soon as $\sigma \in H^1$. This convergence holds uniformly in time, for $u \in L^\infty([0,T], L^\infty)$. Likewise,

$$S_\varepsilon = [\rho_\varepsilon, \partial_i\sigma_{ik}](\sigma_{jk}\partial_j u) \xrightarrow{\varepsilon\to 0} 0 \quad \text{in } L^1, \tag{3.1.15}$$

when σ is again H^1 and $\sigma^t \nabla u \in L^2$. With respect to time, the convergence (3.1.15) holds in L^2, because $\sigma^t \nabla u \in L^2([0,T], L^2)$. As for the last term, we have

$$T_\varepsilon = [\rho_\varepsilon, \sigma_{ik}\partial_i](\sigma_{jk}\partial_j u) \xrightarrow{\varepsilon\to 0} 0 \quad \text{in } L^1, \tag{3.1.16}$$

also in L^2 with respect to time. Setting

$$U_\varepsilon = Q_\varepsilon + \frac{1}{2}S_\varepsilon + \frac{1}{2}T_\varepsilon,$$

we now collect all these convergences in (3.1.11) and obtain

$$\partial_t u_\varepsilon + b_i \partial_i u_\varepsilon - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j u_\varepsilon) = U_\varepsilon + \frac{1}{2}\partial_i(\sigma_{ik}R_{k,\varepsilon}) \tag{3.1.17}$$

with

$$U_\varepsilon \xrightarrow{\varepsilon\to 0} 0 \quad \text{in } L^\infty + L^2([0,T], L^1),$$

$$R_{k,\varepsilon} \xrightarrow{\varepsilon\to 0} 0 \quad \text{in } L^\infty([0,T], L^2),$$

under the conditions $\mathbf{b} \in W^{1,1}$, $\operatorname{div}\mathbf{b} \in L^\infty$, $\sigma \in H^1$, $\sigma^t\nabla u \in L^2$. It is readily seen, multiplying (3.1.17) by u_ε and integrating, that we may now *rigorously* obtain estimate (3.1.1). This concludes our outline of the regularization step involved in the proof of Proposition 1. The remainder of the proof of Proposition 1 is classical. □

Remark 32. As announced above, we devote this remark to some comments on the functional space X defined in (3.1.3). We have considered this space X for $\sigma \in H^1$. This regularity is needed for the regularization procedure that we have just performed in the proof of Proposition 1. It is also needed, as will now be seen, to have suitable properties of the functional space at hand. Evidently, the space X involves (for each time t) functions that at least belong to the space

$$H = \{u \in L^2 : \sigma^t\nabla u \in L^2\}. \tag{3.1.18}$$

This space H is indeed, when $\sigma \in H^1$, a space of distributions, that is: smooth compactly supported functions are dense in H for the natural norm of H. In order to show this, we need to find a sequence of smooth functions approximating u for the norm of H. To this end, we first notice that it is enough to obtain this property for $u \in L^\infty \cap H$. Indeed, we know from the techniques of the previous sections that the function $T_R(u) = (u \wedge R) \vee (-R)$, bounded for all R fixed, approximates u for the norm of H, since $T_R(u) = u\mathbb{1}_{|u|<R}$ and $\nabla T_R(u) = \nabla u \mathbb{1}_{|u|<R}$. Next, for $u \in L^\infty \cap H$, we note that the regularization procedure performed in the proof of Proposition 1 has used (3.1.12) of Lemma 10, which precisely shows that the commutator $[\rho_\varepsilon, \sigma_{ik}\partial_k](u)$ vanishes in L^2 with ε upon the conditions $\sigma \in H^1$ and $u \in L^\infty$. This exactly shows that $\sigma^t\nabla u_\varepsilon$ converges to $\sigma^t\nabla u$ in L^2 (note that without the prior truncation we have performed we would only have $u \in L^2$, thus we would need the stronger assumption $\sigma \in W^{1,\infty}$ to conclude). The density claimed above follows. Note that since we know that the commutation Lemma 10 is, in some sense, sharp, the condition $\sigma \in H^1$ is thus necessary for the space H to make sense as a distribution space.

Remark 33. In the case when, following (3.0.3) or (3.0.4), $\mathbf{a} = \sigma C \sigma^t$, we modify (3.1.9) as follows:

$$\begin{aligned}\rho_\varepsilon \star \partial_i(\sigma_{ik}C_{kl}\sigma_{jl}\partial_j u) &= \partial_i(\rho_\varepsilon \star (\sigma_{ik}C_{kl}\sigma_{jl}\partial_j u)) \\ &= \partial_i([\rho_\varepsilon, \sigma_{ik}C_{kl}\sigma_{jl}\partial_j](u)) + \partial_i(\sigma_{ik}C_{kl}\sigma_{jl}\partial_j u_\varepsilon) \\ &= \partial_i(\sigma_{ik}C_{kl}[\rho_\varepsilon, \sigma_{jl}\partial_j](u) + [\rho_\varepsilon, \sigma_{ik}C_{kl}](\sigma_{jl}\partial_j u)) \\ &\qquad + \partial_i(\sigma_{ik}C_{kl}\sigma_{jl}\partial_j u_\varepsilon) \\ &= \partial_i(\sigma_{ik}C_{kl}R_{k,\varepsilon}) + \partial_i([\rho_\varepsilon, \sigma_{ik}C_{kl}](\sigma_{jl}\partial_j u)) \\ &\qquad + \partial_i(\sigma_{ik}C_{kl}\sigma_{jl}\partial_j u_\varepsilon), \end{aligned} \tag{3.1.19}$$

where we have, like in (3.1.10), denoted by $R_{k,\varepsilon} = [\rho_\varepsilon, \sigma_{jl}\partial_j](u)$. We next write

$$\begin{aligned}\partial_i([\rho_\varepsilon, \sigma_{ik}C_{kl}](\sigma_{jl}\partial_j u)) &= \partial_i([\rho_\varepsilon, \sigma_{ik}](C_{kl}\sigma_{jl}\partial_j u)) + \partial_i(\sigma_{ik}[\rho_\varepsilon, C_{kl}](\sigma_{jl}\partial_j u)) \\ &= [\rho_\varepsilon, \partial_i\sigma_{ik}](C_{kl}\sigma_{jl}\partial_j u) + [\rho_\varepsilon, \sigma_{ik}\partial_i](C_{kl}\sigma_{jl}\partial_j u) \\ &\qquad + \partial_i(\sigma_{ik}[\rho_\varepsilon, C_{kl}](\sigma_{jl}\partial_j u)) \\ &= \widetilde{S}_\varepsilon + \widetilde{T}_\varepsilon + \partial_i(\sigma_{ik}[\rho_\varepsilon, C_{kl}](\sigma_{jl}\partial_j u)), \end{aligned} \tag{3.1.20}$$

where, similarly to (3.1.10), we have denoted by

$$\widetilde{S}_\varepsilon = [\rho_\varepsilon, \partial_i\sigma_{ik}](C_{kl}\sigma_{jl}\partial_j u), \qquad \widetilde{T}_\varepsilon = [\rho_\varepsilon, \sigma_{ik}\partial_i](C_{kl}\sigma_{jl}\partial_j u).$$

Inserting (3.1.20) into (3.1.19), we obtain

$$\begin{aligned}\rho_\varepsilon \star \partial_i(\sigma_{ik}C_{kl}\sigma_{jl}\partial_j u) &= \partial_i(\rho_\varepsilon \star (\sigma_{ik}C_{kl}\sigma_{jl}\partial_j u)) \\ &= \partial_i(\sigma_{ik}C_{kl}R_{k,\varepsilon}) + \widetilde{S}_\varepsilon + \widetilde{T}_\varepsilon + \partial_i(\sigma_{ik}[\rho_\varepsilon, C_{kl}](\sigma_{jl}\partial_j u)) \\ &\qquad + \partial_i(\sigma_{ik}C_{kl}\sigma_{jl}\partial_j u_\varepsilon). \end{aligned} \tag{3.1.21}$$

We now compare (3.1.21) with (3.1.9). Since the matrix C is uniformly positive definite and bounded, both $\sigma^t\nabla u$ and $C\sigma^t\nabla u$ are L^2 and the terms $R_{k,\varepsilon}$, $\widetilde{S}_\varepsilon$, $\widetilde{T}_\varepsilon$ behave like the terms $R_{k,\varepsilon}$, S_ε, T_ε of the proof of Proposition 1. The extra term $\partial_i(\sigma_{ik}[\rho_\varepsilon, C_{kl}](\sigma_{jl}\partial_j u))$ is easily taken care of, given the boundedness of C. It is then straightforward to proceed with the proof and conclude that the result of Proposition 1 carries over to the case (3.0.3). Notice that, as stated in (3.0.4), no assumption of weak differentiability of C is needed.

An interesting extension of Proposition 1 is:

Proposition 2 (Girsanov-type transform, [65, Proposition 3]). *All the results of Proposition 1 hold true when σ satisfies* (3.1.5) *and the two conditions*

$$\mathbf{b} = \widetilde{\mathbf{b}} + \sigma\widetilde{\theta}, \quad \widetilde{\theta} \in L^2, \qquad \widetilde{\mathbf{b}} \in W^{1,1}, \tag{3.1.22}$$

$$\mathbf{b} = \beta + \sigma\Theta, \quad \Theta \in L^\infty, \ [\operatorname{div}\beta]_- \in L^\infty. \tag{3.1.23}$$

In short, condition (3.1.22) in Proposition 2 expresses that a part of the transport field $\mathbf{b}$ that lives in the image space of σ can be easily accommodated provided it is L^2. Of course, when σ is a symmetric positive definite matrix, one finds again that the theory covers the case of a transport field $\mathbf{b} \in L^2$.

Remark 34. Proposition 2 has actually been proven in our previous work [65]. However, one particular claim within our argument, and thus also our statement of the result (see [65, Proposition 3]), were erroneous: we assumed $\Theta \in L^2 + L^\infty$ while $\Theta \in L^\infty$ is the correct assumption to make. Otherwise, one cannot control the transport term in the a priori estimate. We correct our mistake in the above Proposition 2, and its proof below.

Proof of Proposition 2 (outline). The proof of Proposition 2 is a simple modification of that of Proposition 1. In the regularization step, we have the additional error term

$$\begin{aligned}\rho_\varepsilon \star \sigma_{ik}\widetilde{\theta}_k\,.\,\partial_i u - \sigma_{ik}\widetilde{\theta}_k\,.\,\partial_i u_\varepsilon &= [\rho_\varepsilon, \sigma_{ik}\widetilde{\theta}_k\,.\,\partial_i](u),\\ &= \widetilde{\theta}_k[\rho_\varepsilon, \sigma_{ik}\,.\,\partial_i](u) + [\rho_\varepsilon, \widetilde{\theta}_k](\sigma_{ik}\partial_i u),\\ &= \widetilde{\theta}\,.\,Y_\varepsilon + Z_\varepsilon. \end{aligned} \tag{3.1.24}$$

We have

$$Y_\varepsilon = [\rho_\varepsilon, \sigma^t\nabla](u) \xrightarrow{\varepsilon\to 0} 0 \quad \text{in } L^2([0,T], L^2), \tag{3.1.25}$$

because $\sigma \in L^2([0,T], H^1)$ and u is L^∞. On the other hand,

$$Z_\varepsilon = [\rho_\varepsilon, \widetilde{\theta}](\sigma^t\nabla u) \xrightarrow{\varepsilon\to 0} 0 \quad \text{in } L^1([0,T], L^1), \tag{3.1.26}$$

because $\sigma^t\nabla u \in L^2([0,T], L^2)$ and $\widetilde{\theta} \in L^2([0,T], L^2)$. These two convergences show that the error term (3.1.24) vanishes, as ε goes to zero, in $L^1([0,T], L^1)$. This allows to proceed with the regularization step as above.

As for the (first formal and then rigorous when the regularization has been performed) a priori estimate, we again only mention how to bound the additional term, using an Hölder inequality,

$$\begin{aligned}\left|\int \sigma\Theta\,.\,\nabla u u\right| &= \left|\int \sigma^t\nabla u\,.\,\Theta u\right|\\ &\le \|\Theta\|_{L^\infty}\|u\|_{L^2}\|\sigma^t\nabla u\|_{L^2}, \end{aligned} \tag{3.1.27}$$

and the proof of the following estimate, similar to the a priori estimate (3.1.1),

$$\frac{d}{dt}\int \frac{u^2}{2} + \int \frac{u^2}{2}\operatorname{div}\beta + \frac{1}{4}\int \left|\sigma^t\nabla u\right|^2 \le C\int \frac{u^2}{2}$$

for some irrelevant constant C depending upon $\|\Theta\|_{L^\infty}$, is then clear. The proof then follows the exact pattern of the proof of Proposition 1. □

Remark 35. Proposition 2 also applies to the case when $\mathbf{a} = \sigma C\sigma^t$ as in (3.0.4) without modifying assumptions (3.1.22) and (3.1.23). The above proof indeed only makes use of the fact that $\sigma^t\nabla u \in L^2$, a fact which is also true in that case, given the properties of the matrix C in (3.0.4).

Remark 36. The setting of Proposition 2 gives a convenient opportunity to return to our remarks of page 39 regarding the limit $p \to 1$ of our arguments in L^p, which gives estimations in $L \ln L$. Assume (for simplicity) that we address the case of the equation

$$\partial_t u - \partial_i(b_i u) - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j u) = 0$$

(obviously a variant of (3.0.1)) where, as in Proposition 2 and again to keep things simple, we assume $\mathbf{b} = \sigma\Theta$ with $\Theta \in L^\infty$. Then, upon formally multiplying the above equation by $\ln u$, we obtain

$$\frac{\partial}{\partial t}\int (u \ln u - u) + \int \frac{|\sigma^t \nabla u|^2}{u} = -\int \Theta\sigma^t \nabla u \leq \|\Theta\|_{L^\infty} \int \frac{|\sigma^t \nabla u|}{\sqrt{u}}\sqrt{u},$$

thus, using the Cauchy–Schwarz inequality and $\int u \leq C \int (u \ln u - u)$, the estimation

$$\int (u \ln u - u) \leq e^{Ct} \int (u_0 \ln u_0 - u_0).$$

This is, as announced, the formal extension of our classical L^p estimates. The above argument may of course be made rigorous by regularization, as usual.

In terms of probability theory, the invariance of our assumptions on $\mathbf{b}$ with respect to the addition of $\sigma\theta$ has a significance. Let us briefly recall the well known Girsanov transform. Consider $(\mathbf{X}_t, \mathbf{W}_t)$ a solution to the stochastic differential equation

$$d\mathbf{X}_t = (\mathbf{b}(\mathbf{X}_t) + \sigma\theta(\mathbf{X}_t))\, dt + \sigma(\mathbf{X}_t)\, d\mathbf{W}_t \tag{3.1.28}$$

for $\mathbf{b}$, σ regular and θ in L^2. We may define

$$\overline{\mathbf{W}_t} = \mathbf{W}_t + \int_0^t \theta(\mathbf{X}_s)\, ds, \tag{3.1.29}$$

which is a (k-dimensional) Brownian motion under the probability $\mathbb{P}^\theta$ defined by

$$\frac{d\mathbb{P}^\theta}{d\mathbb{P}} = \exp\left(-\int_0^t \theta(\mathbf{X}_s) \,.\, d\mathbf{W}_s - \frac{1}{2}\int_0^t |\theta|^2(\mathbf{X}_s)\, ds\right). \tag{3.1.30}$$

Then $(\mathbf{X}_t, \overline{\mathbf{W}_t})$ is a (weak) solution to the stochastic differential equation

$$d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)\, dt + \sigma(\mathbf{X}_t)\, d\overline{\mathbf{W}_t}. \tag{3.1.31}$$

Intuitively, this transform is an elaborate, infinite-dimensional version of the following elementary, finite-dimensional observation. Given that

$$\exp\left(\sum_{i=1}^N (\mu_i x_i - \frac{1}{2}\mu_i^2)\right) \exp\left(-\frac{1}{2}\sum_{i=1}^N x_i^2\right) = \exp\left(-\sum_{i=1}^N \frac{1}{2}(x_i - \mu_i)^2\right), \tag{3.1.32}$$

a collection of N independent Gaussian random variables with unit variance can be understood either as centered, that is, with density $\exp(-\frac{1}{2}\sum_{i=1}^N x_i^2)$, or with means μ_i, that is, with density $\exp(-\sum_{i=1}^N \frac{1}{2}(x_i - \mu_i)^2)$ provided one possibly modifies the underlying probability by the density $\exp(\sum_{i=1}^N (\mu_i x_i - \frac{1}{2}\mu_i^2))$. A Brownian motion being precisely a collection of infinitely many independent Gaussian increments, equation (3.1.30) is the infinite-dimensional analogue of the transformation (3.1.32).

The Girsanov transform therefore amounts to replacing the drift vector **b** by **b** – $\sigma\theta$ so that each given trajectory $\mathbf{X}_t$ still solves the equation (but for another Brownian motion, now implicitly depending on $\mathbf{X}_t$). The control of the law of $\mathbf{X}_t$ is not affected by this transform. Thus the invariance that the assumptions of our Proposition 2 express in terms of the theory of existence and uniqueness for the associated partial differential equation. Note, however, that our manipulations involve functions $\theta \in L^2$ that are less regular than those admissible for the classical Girsanov transform, for which $\theta \in L^\infty$. Our terminology "Girsanov transform" then only refers to an analogy.

3.2 Non-degenerate cases

We may obtain a result stronger than that of Proposition 1 when the matrix $\sigma\sigma^t$ is uniformly positive definite, i.e. there exists some constant $C_0 > 0$ such that

$$|\sigma^t(x)\zeta|^2 \geq C_0|\zeta|^2 \quad \text{for all } \zeta \in \mathbb{R}^N \text{ and for almost all } x. \tag{3.2.1}$$

The details are given in [65]. A summary of the key steps of the argument goes as follows. Inequality (3.2.1) clearly implies that

$$\int_{\mathbb{R}^N} |\sigma^t \nabla u|^2 \geq C_0 \int_{\mathbb{R}^N} |\nabla u|^2, \tag{3.2.2}$$

and thus the a priori estimate (3.1.1) implies that the functions of X defined in (3.1.3) also belong to $L^2([0,T], H^1)$. With such an H^1 integrability of the solution, it is now well known (see [64] and our arguments of the previous sections) that the regularization step can then accommodate a drift vector that is only L^2 and not necessarily $W^{1,1}$ in space. In the estimates of the regularization procedure (outlined in the previous section), we write, instead of (3.1.9),

$$\rho_\varepsilon \star \partial_i(\sigma_{ik}\sigma_{jk}\partial_j u) - \partial_i(\sigma_{ik}\sigma_{jk}\partial_j u_\varepsilon) = \partial_i([\rho_\varepsilon, \sigma_{ik}\sigma_{jk}\partial_j](u)), \tag{3.2.3}$$

where we now denote by $\partial_i R_\varepsilon$ the right-hand side. If we next assume $\sigma \in L^\infty$, we see, by application of Lemma 10, that

$$R_\varepsilon = [\rho_\varepsilon, \sigma_{ik}\sigma_{jk}\partial_j](u) \xrightarrow{\varepsilon \to 0} 0 \quad \text{in } L^2([0,T], L^2).$$

Multiplying the regularized equation by u_ε and integrating in space and time yields the following specific contribution:

$$\int_0^T \int [\rho_\varepsilon, \sigma_{ik}\sigma_{jk}\partial_j](u)\partial_i u_\varepsilon = \int_0^T \int R_\varepsilon \partial_i u_\varepsilon,$$

where the right-hand side vanishes when ε goes to zero because

$$u \in L^2([0,T], H^1) \qquad \text{and} \qquad R_\varepsilon \xrightarrow{\varepsilon \to 0} 0 \quad \text{in } L^2([0,T], L^2).$$

Consequently, the regularization may be performed, and the remainder of the argument for existence and uniqueness follows unchanged. The following proposition therefore holds.

Proposition 3 (Positive definite diffusion matrices, [65, Proposition 4]). *Assume that the matrix $\sigma\sigma^t$ is uniformly positive definite (i.e. that* (3.2.1) *holds) and that*

$$\mathbf{b} \in L^2 + W^{1,1}, \quad [\operatorname{div} \mathbf{b}]_- \in L^\infty, \quad \sigma \in L^\infty.$$

Then, for each initial condition in L^∞, (1.3.5) *has a unique solution in the space*

$$\{u \in L^\infty([0,T], L^\infty) \cap L^2([0,T], H^1) \cap C^0([0,T], L^p) : 1 \le p < +\infty\}.$$

Although we will not go in that direction, we mention that many extensions of the results we have obtained for a *constant* matrix **a** in Chapter 2 can be adapted in the case of the operator in divergence form when the matrix is positive definite. Extensions of Proposition 3, especially regarding the assumptions of type (2.1.2) on **b** that we may consider, are thus possible.

Other remarks in the same vein are contained in the next section.

3.3 Remarks

We give here various comments.

First, we note that the theory developed above for equation (3.0.1) with the second-order operator in divergence form has immediate consequences on the equation that is not in divergence form, that is, (1). It suffices to replace the transport field **b** by the field of coordinates $b_i - \partial_j a_{ij}$ and assume the suitable conditions on that field. We will return to this in more details in Section 4.1.

We also note that, like in all the results established above, we have indeed proved more than existence and uniqueness. We have proven convergence of the regularized solutions.

Like in the case of a second-order operator with constant coefficients, we may consider all the estimates that come together with existence and uniqueness. As in Theorem 1, we may prove, under the appropriate assumptions, estimates (2.1.3) and (2.1.4). For brevity, we skip the proofs of these estimates, which are essentially similar to the proofs given above in the case of second-order operators with constant coefficients. As mentioned above, the case of an initial condition $u_0 \in L^p$, $1 < p < +\infty$, instead of L^∞, may also be considered, upon some modifications of our argument, based on a renormalization procedure. In contrast, notice that in our proof of the L^1 estimate (2.1.24) established in Lemma 2, we explicitly made use of the heat kernel, and so that estimate requires a specific adaptation, which we will address in Section 3.4.

An interesting remark is that, for a generic σ enjoying no additional property, the H^1 regularity of σ is actually necessary for uniqueness to hold. To illustrate the critical role played by this regularity, we again consider the static variant of our problem, and more precisely the following simple one-dimensional, stationary problem:

$$-\frac{1}{2}\frac{d}{dx}\left(|x|^{2\beta}\frac{du}{dx}\right)+u=f \tag{3.3.1}$$

for an exponent $\beta > 0$, and, say, a smooth right-hand side f. Equation (3.3.1) is, e.g., set on $[-1,+1]$ with homogeneous Dirichlet boundary conditions at -1 and $+1$. In that case, $\sigma(x)=|x|^{\beta}$, thus $\sigma'=\beta|x|^{\beta-2}x$. Note that $\sigma\in H^1$ if and only if $\beta>\frac{1}{2}$. The natural solution space to consider is the space $Y := \{u\in L^{\infty} : |x|^{\beta}u'\in L^2,\ u(\pm 1)=0\}$. In the definition of this space Y, $|x|^{\beta}u'$ is well defined in the sense of distributions using $|x|^{\beta}u'=(|x|^{\beta}u)'-\beta|x|^{\beta-2}xu$, where $|x|^{\beta-1}$ is in some L^p since $\beta>0$.

For $\beta>\frac{1}{2}$, one may show there is a unique solution to (3.3.1) in this space. Indeed, like in Remark 32, one shows that $|x|^{\beta}(\rho_{\varepsilon}\star u)'$ converges in L^2 to $|x|^{\beta}u'$, since then $|x|^{\beta}\in H^1$. Using $\rho_{\varepsilon}\star u$ (with some surgery at ± 1 to agree with the homogeneous boundary conditions) in the formulation of (3.3.1) (for $f=0$) in the distribution sense, and next letting ε vanish, proves uniqueness.

In the limit case $\beta=\frac{1}{2}$, this regularization does not work and the space Y is somewhat peculiar. However, the coefficient $|x|^{\beta}=|x|^{\frac{1}{2}}$ is $C^{0,\frac{1}{2}}$ and one therefore knows, by arguments that are specific to the one-dimensional setting, that the stochastic flow of the stochastic differential equation associated to (3.0.1) is unique. Unfortunately, this does not say anything on the uniqueness for equation (3.3.1) itself in a distribution space of solutions.

On the other hand, for $0<\beta<\frac{1}{2}$, we now show that uniqueness cannot hold. This will prove that our assumption $\nabla\sigma\in L^2$ *cannot* be weakened into $\nabla\sigma\in L^p$ for any $p<2$. To this end, we first note that the regularity of f and the L^{∞} character of u, once inserted in (3.3.1), show that necessarily $|x|^{2\beta}u'$ is Lipschitz continuous, thus continuous at the origin. Consequently, u' behaves as $|x|^{-2\beta}$ and thus, since $\beta<\frac{1}{2}$, is integrable. We may therefore *prescribe* an arbitrary value of u on either side of the origin, and solve *separately* (3.3.1) on either side $[-1,0]$ and $[0,+1]$, with the respective Dirichlet boundary conditions $(u(-1)=0, u(0^-))$ and $(u(0^+), u(1)=0)$. On either side, we thus get a one-parameter family of solutions, indexed by $u(0^-)$ and $u(0^+)$, respectively. To form a solution on the entire interval $[-1,+1]$, the fluxes $|x|^{2\beta}u'$ need to agree at the origin. This eliminates one of the two degrees of freedom and leaves us with a one-parameter family of solutions for (3.3.1). Note indeed that the left-hand side solution and the right-hand side solution thus constructed need not agree at the origin: a non-zero jump $u(0^+)-u(0^-)$ translates in a Dirac mass δ_0 at the origin, which is not seen in the term $|x|^{2\beta}u'$ since $\beta>0$ and $|x|^{2\beta}\delta_0=0$ in the distribution sense. We thus obtain infinitely many solutions to (3.3.1). Of course, the absence of uniqueness mentioned above does not contradict the uniqueness of the variational solution to (3.3.1), that is, the minimizer of the corresponding, strictly convex energy functional $\frac{1}{4}\int_{-1}^{+1}|x|^{2\beta}|u'|^2+\frac{1}{2}\int_{-1}^{+1}u^2-\int_{-1}^{+1}fu$ on the suitable functional space. The variational problem indeed *selects* one, unique, solution among the infinitely many solutions to (3.3.1): for instance, for $f \equiv 0$, the minimizer $u \equiv 0$ of the above functional on $H_0^1([-1, 1])$ is clearly unique, while our proof of non-uniqueness performed above shows the existence of infinitely many solutions to the equation.

3.4 About the L^1 estimate

Our purpose in this section is to discuss the adaptation of the L^1 estimate contained in Lemma 2 to the present context of an equation in divergence form (3.0.1) with a varying diffusion matrix $a_{ij} = \frac{1}{2}\sigma_{ik}\sigma_{jk}$. Our comments and results follow up on those of Section 2.1.3.

General comments. We recall that the issue with the L^1 estimate is that, the function $|t|$ being not strongly convex, it is delicate to bootstrap the expected improved integrability and regularity on the solution u from the presence of the second-order term. In particular, it is thus difficult to allow for more general transport fields $\mathbf{b}$ than those already suitable in the pure transport case, which we know must be $W^{1,1}$ and have their divergence (or, more precisely the negative part of that divergence) controlled. We are going to see that, as in the case of Lemma 2 for a second-order operator with constant coefficients, it is indeed possible to extend the setting.

First of all, we mention that, for discussing the L^1 estimate in the case of a varying diffusion matrix, we have to distinguish between the case where the operator is in divergence form (which is the setting of the present Chapter 3), and the case where it is not (which will be examined in Section 4.1). We also have to distinguish, within each category, between the case where the operator is elliptic, and where it is not (and distinguish also “intermediate” cases of hypoellipticity and related settings, which will be dealt with in Section 3.5). In the present section, we will mainly consider the divergence form, first under the assumption of ellipticity and next allowing for degeneracies. We will next proceed with some remarks on the equation not in divergence form, anticipating the content of Section 4.1.

We also mention that all estimates we will discuss below are of course obvious when σ is assumed sufficiently regular. If we for instance assume integrability of the second derivatives $\partial^2_{ij}\mathbf{a}$, then “everything” which is true for the Laplacian operator is true for the operator in divergence form. However our purpose throughout these notes is to avoid such assumptions of second-order differentiability on $\mathbf{a}$. In passing, we emphasize that this purpose is bold. In the case of the linear transport equation, we *need* to put conditions on the first derivatives of the field $\mathbf{b}$, namely conditions on $\operatorname{div}\mathbf{b}$, in order to be able to control the L^1 norm of the solution, and then we get an *exponential* growth of that norm. So, intuitively, avoiding conditions on the second derivatives of $\mathbf{a}$ in the parabolic case is more demanding.

Before we begin our discussion, we mention an important issue, which is closely related to establishing an L^1 estimate on the solution u. We have already observed that the purpose of the argument is to obtain a better regularity on the solution in order to accommodate more general transport fields than in pure transport equations. This is in essence based upon proving integrability of the gradient ∇u of the solution u. More precisely in our setting, this is about understanding the integrability of $\sigma^t \nabla u$. The least we can expect to establish is $\sigma^t \nabla u \in L^1$. Put differently, the issue is not so much to obtain an L^1 estimate on u than to obtain this estimate *together with* the property $\sigma^t \nabla u \in L^1$.

A comment in line with the above discussion concerns the Girsanov-type transform we addressed in Proposition 2 above. Let us consider, for simplicity, the equation

$$\frac{\partial u}{\partial t} + \sigma\Theta \,.\, \nabla u + Au = 0, \tag{3.4.1}$$

where $\Theta \in L^\infty$ (as in (3.1.23)) and A denotes the second-order operator

$$A = -\frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j)$$

(but it could equally well be, for this specific part of our discussion, $A = -\frac{1}{2}\sigma_{ik}\sigma_{jk}\partial^2_{ij}$ or $A = -((\sigma^k)^t \nabla)^2$ in the Hörmander notation (anticipating on the next section)). Assume that, for *all* functions $\Theta \in L^\infty$, we wish to establish an L^1 estimate on the solution u to that equation (3.4.1) (or more precisely, as is always the case in these notes, for that equation with possibly a right-hand side f in the suitable functional space, here $f \in L^1_{t,x}$). This can indeed be expected to be true and is particularly relevant for the applications to the probabilistic setting we have in mind. Then, at least formally (and the argument we are about to outline can be made rigorous using adequate assumptions and regularization techniques), this is equivalent to establishing both that u solution to (3.4.1) satisfies $\sigma^t \nabla u \in L^1_{t,x}$ and that the L^1 estimate holds for any solution to (3.4.1) for $\Theta = 0$. Indeed, first suppose the latter two properties hold true. Then, using $\Theta \in L^\infty$ and $\sigma^t \nabla u \in L^1_{t,x}$, we write $\frac{\partial u}{\partial t} + Au = -\sigma\Theta \,.\, \nabla u \in L^1_{t,x}$, thus, given the L^1 estimate for the equation without drift term, we obtain that same estimate on u. Conversely, if we know that the L^1 estimate holds for (3.4.1) for all $\Theta \in L^\infty$, then we apply this to the particular function $\Theta = \frac{\sigma^t \nabla u}{|\sigma^t \nabla u|}$. This gives the L^1 estimate for the solution u to

$$\frac{\partial u}{\partial t} + |\sigma^t \,.\, \nabla u| + Au = 0. \tag{3.4.2}$$

In turn, by formal integration at least, this L^1 estimate yields $\sigma^t \nabla u \in L^1_{t,x}$.

In the sequel, with a view to investigating conditions under which we may include in our theory all functions $\Theta \in L^\infty$ (as in the Girsanov transform) at no additional cost, we therefore examine both the question of establishing an L^1 estimate on u and that of the L^1 integrability of $\sigma^t \nabla u$. We notice that, in the case of an operator in divergence form as in (3.0.1) and for $\mathbf{b} = 0$, the L^1 estimate on the solution u is formally evident, integrating the equation (remark we may always assume that the solution is nonnegative). The actual proof proceeds by regularization. The integrability of $\sigma^t \nabla u$ is intuitively less clear and is more demanding technically. We will investigate the question below. For operators that are not in divergence form, we will, except for some very specific situations (see below the case of an operator in the Hörmander form), have to transform the second-order term into divergence form, at the price of some additional first-order corrections which will be dealt with as the original transport field is. The detail of such manipulations will be made precise in Section 4.1.4, where we will specifically return to the issue of the L^1 estimate.

Notice also that, for simplicity, we will proceed assuming that the transport field **b** vanishes. Reinstating that term in our arguments is easy, precisely because the terms of dominant order in the estimates will have been dealt with. We will not make precise the assumptions needed on **b**, for the regularization step and for the estimation step, since they are immediate consequences, by duality and simple inequalities, of the estimates we establish in the absence of transport, and they very much depend on the combination of assumptions we take. Once the bounds obtained below are made clear, the rest of the argument consists in technicalities. We spare the reader those technicalities.

In view of the above discussion, we consider henceforth in this section the equation

$$\frac{\partial u}{\partial t} - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j u) = 0, \tag{3.4.3}$$

supplied with the initial condition $u_0 \in L^1$, and examine the issue of an L^1 estimate on u and on $\sigma^t \nabla u$. We know from the above that this is the key issue for our theory for what regards the L^1 estimate in the general case.

Equation in divergence form, uniformly elliptic case. We first assume that the matrix $\sigma\sigma^t$ is uniformly elliptic. Formally, and this is made rigorous using regularization, we have

$$\sup_{t\geq 0} \int u(t) = \int u_0, \tag{3.4.4}$$

that is, in particular $u \in L^\infty_t(L^1_x)$, simply integrating equation (3.4.3) and assuming without loss of generality that $u_0 \geq 0$, and thus $u(t) \geq 0$ for all times. The point is now to obtain better integrability on u and integrability on ∇u. Introducing $S_R(u)$ and $T_R(u)$ as in Section 2.3, we obtain, multiplying the equation by $S_R(u)$ and integrating both in time and space,

$$\begin{aligned}
\int T_R(u(t)) + \int_0^t \int |\nabla T_R(u)|^2 &\leq \int T_R(u_0) \\
&= \int T_R(u_0)\mathbb{1}_{|u_0|<R} + \int T_R(u_0)\mathbb{1}_{|u_0|>R} \\
&\leq \int \frac{|u_0|^2}{2}\mathbb{1}_{|u_0|<R} + \int \left(R|u_0| - \frac{R^2}{2}\right)\mathbb{1}_{|u_0|>R} \\
&= O(R)\int |u_0|.
\end{aligned}$$

It follows that

$$\int_0^\infty \int |\nabla T_R(u)|^2 \leq C_0 R, \tag{3.4.5}$$

where the constant C_0 depends on u_0. Using the Sobolev inequality, we readily infer from this bound that

$$\int_0^\infty \|T_R(u)\|^2_{L^{\frac{2d}{(d-2)}}} \leq C_0 R. \tag{3.4.6}$$

If the function $T_R(u)$ were not depending upon time, and if the bounds (3.4.4) and (3.4.5) were reading as $\int u \leq C$ and $\int |\nabla T_R(u)|^2 \leq CR$, respectively, we would immediately deduce that

$$\|u\|_{L^{\frac{d}{(d-2)},\infty}} \leq C, \tag{3.4.7}$$

using Lemma 8, and

$$\|\nabla u\|_{L^{\frac{d+2}{(d+1)},\infty}} \leq C, \tag{3.4.8}$$

using Lemma 9 and, on the right-hand side of (2.4.5), the inequality $\frac{d}{d-2} > \frac{d+2}{d}$.

For functions depending upon time, such informations of integrability on u and ∇u are more intricate to establish.

In order to get the bound on u, the simplest possible argument is based upon symmetrization. We know (see [4]) that the solution u to (3.4.3) satisfies, for all p, q,

$$\|u(t)\|_{L^{p,q}} \leq \|v(t)\|_{L^{p,q}},$$

where v solves

$$\frac{\partial v}{\partial t} - \frac{1}{2}\Delta v = 0, \tag{3.4.9}$$

and $v(0) = (u_0)^*$, the Schwarz symmetrization of u_0. Notice that, for simplicity, we have assumed that the coercivity constant of $\sigma\sigma^t$ is one. The solution to (3.4.9) may be readily expressed in terms of the heat kernel (2.1.34), and bounds on v, thus on u, follow. In particular, simply using that $\|p(t)\|_{L^\infty} = O(\frac{1}{t^{d/2}})$ we obtain

$$\|u(t)\|_{L^\infty} \leq \frac{C}{t^{\frac{d}{2}}}\|(u_0)^*\|_{L^1} = \frac{C}{t^{\frac{d}{2}}}\|u_0\|_{L^1}. \tag{3.4.10}$$

The Hölder inequality then yields from (3.4.4) and (3.4.10)

$$\|u(t)\|_{L^{\frac{d}{(d-2)}}} \leq \|u(t)\|^{\frac{2}{d}}_{L^\infty}\|u(t)\|^{\frac{(d-2)}{d}}_{L^1} \leq \frac{C}{t}. \tag{3.4.11}$$

We have therefore obtained, intuitively, $u \in$ "$L_t^{1,\infty}(L^{\frac{d}{(d-2)}})$" $\cap L_t^\infty(L^1)$.

There remains to now bound ∇u. We begin by two remarks and then proceed with outlining the essential ingredients of proof for two different bounds, namely (3.4.16) and (3.4.17) below.

A first remark is that, in the particular case when σ does not depend on time, a bound on ∇u can be readily deduced from (3.4.10). Indeed, we first obtain by the Hölder inequality that

$$\int |u(t)|^2 \leq \|u(t)\|_{L^\infty}\|u(t)\|_{L^1} \leq \frac{C}{t^{\frac{d}{2}}}.$$

Next, formally multiplying (3.4.3) by $\frac{\partial u}{\partial t}$ (but again, this can be made rigorous using regularization), we have

$$\int \left(\frac{\partial u}{\partial t}\right)^2 + \frac{1}{4}\frac{d}{dt}\int |\sigma^t \nabla u|^2 = 0$$

from where we infer that $\int |\sigma^t \nabla u|^2$ is a decreasing function of time. But then, starting from the classical a priori L^2 estimate of (3.4.3), namely

$$\frac{d}{dt}\int u^2 + \frac{1}{2}\int |\sigma^t \nabla u|^2 = 0,$$

we obtain, integrating from $\frac{t}{2}$ to t, using the decay, and the bound $\int |u(t)|^2 \leq \frac{C}{t^{d/2}}$, that

$$c\frac{t}{2}\int |\nabla u(t)|^2 \leq \frac{1}{2}\int_{\frac{t}{2}}^{t}\int |\sigma^t \nabla u|^2 \leq \int \left|u\left(\frac{t}{2}\right)\right|^2 \leq Ct^{-\frac{d}{2}},$$

thus the following bound on the gradient:

$$\int |\nabla u(t)|^2 \leq Ct^{-(1+\frac{d}{2})}.$$

A second remark is the following. If we knew that

$$\iint \frac{|\nabla u|^2}{|u|} < \infty, \tag{3.4.12}$$

then, using (3.4.4) and the Cauchy–Schwarz inequality,

$$\int |\nabla u| \leq \left(\int \frac{|\nabla u|^2}{|u|}\right)^{\frac{1}{2}}\left(\int |u|\right)^{\frac{1}{2}}, \tag{3.4.13}$$

we would obtain

$$\nabla u \in L_t^2(L_x^1). \tag{3.4.14}$$

But we know the latter assertion is not correct. Otherwise it would carry over to the case of an initial condition that is a bounded measure, thus it would in particular apply to the heat kernel (2.1.34). Now, $\nabla(\frac{1}{t^{d/2}}e^{-\frac{|x|^2}{2t}})$ cannot be $L_t^2(L_x^1)$ since its L_x^1 norm scales like $\frac{1}{\sqrt{t}}$. On the other hand, one can easily establish that the bound (3.4.12) does hold for an initial condition $u_0 \in L\ln L$. When we only assume $u_0 \in L^1$, what we can only establish, mimicking the above argument, is the following. Multiplying the equation by $u^{-\delta}$ for $\delta > 0$ arbitrarily small and integrating in space and time, we obtain, instead of (3.4.12),

$$\int_0^T\int \frac{|\nabla u|^2}{|u|^{1+\delta}} \leq \int |u(T)|^{1-\delta},$$

where the rightmost integral is bounded *if we assume the domain is bounded*. We then obtain, instead of (3.4.13),

$$\int |\nabla u| \leq \left(\int \frac{|\nabla u|^2}{|u|^{1+\delta}}\right)^{\frac{1}{2}}\left(\int |u|^{1+\delta}\right)^{\frac{1}{2}}.$$

The first factor is L_t^2. Because of (3.4.4) and (3.4.10), the second factor is in some L_t^q for all $q < q(\delta)$, with $q(\delta) \to +\infty$ as $\delta \to 0^+$ (actually, $q(\delta) = \frac{4}{\delta d}$). Choosing δ arbitrarily small, we therefore obtain, instead of (3.4.14),

$$\nabla u \in L_t^p(L_x^1) \quad \text{for all } p < 2. \tag{3.4.15}$$

We shall see below that the optimal information we can get is

$$\nabla u \in L_t^{2,\infty}(L_x^1). \tag{3.4.16}$$

In passing, we notice that, since $u \in$ "$L_t^{1,\infty}(L^{\frac{d}{(d-2)}})$" $\cap L_t^\infty(L^1)$ (a property rigorously formalized in (3.4.4) and (3.4.11)), we may also write

$$\int |\nabla u|^{\frac{d}{d-1}} \leq \left(\int \frac{|\nabla u|^2}{|u|}\right)^{\frac{d}{2(d-1)}} \left(\int |u|^{\frac{d}{d-2}}\right)^{\frac{(d-2)}{2(d-1)}},$$

where, for simplicity of exposition, we have taken $\delta = 0$. Formally, the first factor is $L_t^{2(d-1)/d}$ and the second is $L_t^{2(d-1)/(d-2)}$, and these two exponents are conjugate to one another. Thus, the bound that is "almost" true reads as

$$\text{“}\nabla u \in L_t^1(L_x^{\frac{d}{d-1}})\text{”}, \tag{3.4.17}$$

and formally agrees with what we would obtain for the heat kernel itself. Making the above formal argument rigorous simply requires to reinstate $\delta > 0$, successively use (3.4.4) and (3.4.11), and prove an explicit bound on ∇u corresponding to a formalization of the loosely stated property (3.4.17). In some sense, the estimates (3.4.16) and (3.4.17) are the two limit cases of integrability of ∇u, in time and space, respectively.

The above discussion mainly consisted of remarks and formal proofs of bounds on ∇u. We mentioned those of these proofs that could be made rigorous, and we mentioned how and when. In order to actually establish general bounds, we can proceed in at least two ways, which we briefly mention now. In some sense, the proofs omitted below are the formalizations of the general ideas developed so far in this section. The details we skip, and for which we refer the reader to the specific bibliography mentioned or to classical arguments of the literature, resemble arguments we sketched above in the proofs of Lemma 8 and Lemma 9.

We may obtain bounds on ∇u proceeding as in [70, proofs of Theorems 3.7 and 3.8, pp. 101–105]. Based on estimates (3.4.4), (3.4.10), (3.4.11) on u and estimate (3.4.5) on the truncated version of ∇u, the arguments of [70, pp. 101–105] may be adapted to obtain $L_t^p(L_x^1)$ bounds on ∇u, for all $1 \leq p < 2$, and also, with further adaptations, $L_t^p(L_x^q)$ bounds, for all $1 \leq p \leq 2$, $1 \leq q \leq \frac{d}{d-1}$, $\frac{d}{p} + \frac{2}{p} < d + 1$. We also can obtain, specifically with [70, proof of Theorem 3.7], a bound in $L_{t,x}^{5/2,\infty}$ in dimension $d = 3$.

In addition, we may also proceed using duality arguments to get further bounds. In particular, proving the $L_t^{2,\infty}(L_x^1)$ bound on ∇u announced in (3.4.16) amounts to proving that the solution v to

$$\frac{\partial v}{\partial t} - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j v) = \operatorname{div} \mathbf{f}, \tag{3.4.18}$$

starting from the initial condition $v(0) = 0$ and for a vector field $\mathbf{f}$ in $L_t^{2,1}(L_x^\infty)$ is bounded.

Indeed, combining the two equations (3.4.3) and (3.4.18), we have

$$\int_0^T \int \nabla u \,.\, \mathbf{f} = -\int_0^T \int u \,\mathrm{div}\,\mathbf{f} = -\int u(T)v(T).$$

Bounding from above the rightmost term by

$$\|u(T)\|_{L^1}\|v(T)\|_{L^\infty} \le C\|\mathbf{f}\|_{L_t^{2,1}(L_x^\infty)},$$

for all $\mathbf{f}$, will then yield the desired bound on ∇u. For this purpose, it suffices to proceed by symmetrization. The result will follow as soon as it holds for the heat equation and the right-hand side $\mathrm{div}(\frac{x}{|x|}g)$, where g is an arbitrary spherically symmetric function in $L_t^{2,1}(L_x^\infty)$. The latter property holds by a direct application of the Young inequality, given the properties of the heat kernel.

Remark 37. We conclude this discussion on the uniformly elliptic case by a remark on hypoelliptic operators, which will be addressed in Section 3.5. Although we did not work out the L^1 estimate in this setting, we anticipate that all the arguments we outlined above may be adapted to hypoelliptic operators, at the price of sometimes substantial additional technical difficulties.

Equation in divergence form in the degenerate case. Alternative particular settings. As mentioned above, obtaining per se an L^1 estimate on the solution u is not an issue when dealing with operators in divergence form. Formally at least, the L^1 estimate falls by integration of the equation. The issue is to obtain that estimate under assumptions on the transport field that are more general than those for the pure transport equation, so that more general drifts and also extra terms (coming, e.g., from a Girsanov-type transform) may be admissible. This in turn requires to establish some integrability property of ∇u, useful both for the a priori estimates and the regularization employed to make rigorous the formal arguments. Obtaining bounds on ∇u is, as we have just seen, already demanding in the uniformly elliptic case.

In the degenerate case, we cannot expect, generically, a better situation than in the pure transport case. Clearly, in the absence of any particular structure assumption on the transport field $\mathbf{b}$, possibly depending upon the structure of the diffusion matrix $\sigma\sigma^t$ itself, we have to assume it satisfies the classical assumptions ($\mathbf{b} \in W^{1,1}$, $[\mathrm{div}\,\mathbf{b}]_- \in L^\infty$). Of course assuming more on the transport field and on the diffusion matrix, that is, for instance degeneracy of the diffusion in some particular fixed directions, and corresponding assumptions on the transport, then we may slightly generalize the setting, but we will not proceed in this direction. On the other hand, when the transport field involves σ, and a prototypical case for this is the Girsanov-type transform we mentioned above, then more can be expected. Using some expected L^1 integrability of $\sigma^t\nabla u$, we could then accommodate such an extension. In any event, one has to understand the correct assumptions to set on σ, for an operator in divergence form that is possibly degenerate.

In order to figure out what minimal regularity on σ is necessary to proceed, we begin by considering the elliptic variant of our parabolic equation (3.4.3), namely

$$-\operatorname{div}(\sigma\sigma^t\nabla u) + u = f, \tag{3.4.19}$$

and investigate the regularity conditions on σ necessary to hope for $\sigma^t\nabla u \in L^1$. We insert $f = \delta_0$ in the right-hand side, thereby considering the critical situation. It suffices, in fact, to consider the one-dimensional version

$$-(\sigma^2(r)u'(r))' + u = \delta_0 \tag{3.4.20}$$

of that equation. Then, we may adapt the arguments below for spherically symmetric functions and obtain the same critical setting in a generic dimension d. We notice that $u = \delta_0$ is solution to (3.4.20) as soon as we have $\sigma^2(r)(\delta_0)'(r) = 0$, which, for the prototype function $\sigma(r) = r^\gamma$, occurs when $\gamma > \frac{1}{2}$. On the other hand, we have for that choice of function, $\sigma^t\nabla u = r^\gamma(\delta_0)'(r)$. The latter distribution is a bounded measure if and only if $\gamma \geq 1$. Therefore, for $\frac{1}{2} < \gamma < 1$, we cannot expect the solution u to (3.4.19) to satisfy $\sigma^t\nabla u \in L^1$ in general. We immediately realize that the same setting shows that the solution u to the parabolic equation (3.4.3) starting from $u_0 = \delta_0$ is $u(t) = \delta_0$ for all times t, and thus cannot be expected to always satisfy $\sigma^t\nabla u \in L^1$ either. This shows that, at least, σ must be Lipschitz continuous to expect $\sigma^t\nabla u \in L^1$, even for operators in divergence form. In the present state of our understanding, we cannot prove this property indeed holds under those assumptions, in the degenerate case.

An alternative setting we may consider, different from the divergence-form, is that of operators in the Hörmander form $A = -\sum_{k=1}^K((\sigma^k)^t\nabla)^2$. The only case we can essentially address then is the case when there is only *one* term, that is, $A = -((\sigma^1)^t\nabla)^2$ for σ^1 a vector-valued function. Our argument may be extended to the case of several components σ^k *provided* the vector fields σ^k commute with one another, that is, for all k and $k' \neq k$, for all $1 \leq i, j \leq d$, $\sigma_j^k\partial_j\sigma_i^{k'} = \sigma_j^{k'}\partial_j\sigma_i^k$. This extension, which of course includes the case of *constant* fields σ^k, is very peculiar and we therefore concentrate on the case of only one component. In the general case, the situation is unclear.

Our proof for the particular setting $A = -((\sigma^1)^t\nabla)^2$ proceeds probabilistically. For simplicity, we denote by $\sigma = \sigma^1$. We assume $\operatorname{div}\sigma \in L^\infty$ and $\sigma \in H^1$. The equation

$$\frac{\partial u}{\partial t} - \frac{1}{2}(\sigma^t\nabla)^2 u = 0 \tag{3.4.21}$$

corresponds to the stochastic differential equation, in Stratonovich form,

$$d\mathbf{X}_t = \sigma(\mathbf{X}_t) \circ d\mathbf{W}_t. \tag{3.4.22}$$

Introducing the deterministic flow $\mathbf{Z}(t,x)$ solution at time t to $\dot{\mathbf{Z}} = \sigma(\mathbf{Z})$ for the initial condition x, the solution to (3.4.22) reads as $\mathbf{X}_t = \mathbf{Z}(\mathbf{W}_t, x)$ and the solution u to (3.4.21) for the initial condition u_0 (which, without loss of generality and as usual in these notes, we assume nonnegative in the sequel, otherwise one needs to proceed in all estimations below with $|u_0|$ instead of u_0) is the expectation

$$u(t,x) = \mathbb{E}[u_0(\mathbf{Z}(\mathbf{W}_t, x))]. \tag{3.4.23}$$

Differentiating this expression with respect to x allows us to write our quantity of interest as

$$\begin{aligned}\sigma^t . \nabla u(t,x) &= \lim_{h\to 0} \mathbb{E}\Big[\frac{u_0(\mathbf{Z}(\mathbf{W}_t, \mathbf{Z}(h,x))) - u_0(\mathbf{Z}(\mathbf{W}_t,x))}{h}\Big] \\ &= \lim_{h\to 0} \mathbb{E}\Big[\frac{u_0(\mathbf{Z}(\mathbf{W}_t + h, x)) - u_0(\mathbf{Z}(\mathbf{W}_t,x))}{h}\Big],\end{aligned} \tag{3.4.24}$$

using the semi-group property of the flow $\mathbf{Z}$ in the latter equality. Next, by Gaussian integration, we know that, for any sufficiently regular function $t \to \varphi(t)$,

$$\lim_{h\to 0} \mathbb{E}\Big[\frac{\varphi(\mathbf{W}_t + h) - \varphi(\mathbf{W}_t)}{h}\Big] = \mathbb{E}[\varphi'(\mathbf{W}_t)] = \mathbb{E}\Big[\varphi(\mathbf{W}_t)\frac{\mathbf{W}_t}{t}\Big].$$

Applying this to $\varphi = u_0(\mathbf{Z}(\,\cdot\,,x))$, we deduce from (3.4.24) that

$$\sigma^t . \nabla u(t,x) = \mathbb{E}\Big[u_0(\mathbf{Z}(\mathbf{W}_t,x))\frac{\mathbf{W}_t}{t}\Big]. \tag{3.4.25}$$

At this stage, we notice that, ω by ω,

$$\int u_0(\mathbf{Z}(\mathbf{W}_t,x))\,dx \le e^{C|\mathbf{W}_t|}\int u_0(x)\,dx, \tag{3.4.26}$$

since the Jacobian J of the deterministic flow of $\dot{\mathbf{Z}} = \sigma(\mathbf{Z})$ is bounded because of our assumption $\operatorname{div}\sigma \in L^\infty$. Taking the expectation yields

$$\int u(t,x)\,dx = \mathbb{E}\int u_0(\mathbf{Z}(\mathbf{W}_t,x))\,dx \le \mathbb{E}(e^{C|\mathbf{W}_t|})\int u_0\,dx.$$

We notice at this point that $\mathbb{E}(e^{C|\mathbf{W}_t|})$ is much larger than an exponential e^{Kt} for small times t, which is consistent with our earlier remarks on page 40 regarding the fact that for an exponential-in-time L^1 bound to hold, the divergence of the drift has to be bounded. This would imply here a control of the second-order derivatives of σ, which we do not assume.

On the other hand, taking the absolute value and integrating (3.4.25) in x, next inserting for "each" ω, estimation (3.4.26) into the right-hand side, we obtain

$$\begin{aligned}\int |\sigma^t . \nabla u|\,dx &\le \int \mathbb{E}_x\Big[u_0(\mathbf{Z}(\mathbf{W}_t,x))\frac{|\mathbf{W}_t|}{t}\Big]\,dx \\ &\le \mathbb{E}\Big[e^{C|\mathbf{W}_t|}\frac{|\mathbf{W}_t|}{t}\Big]\int u_0(x)\,dx \\ &= O\Big(\frac{1}{\sqrt{t}}\Big)\end{aligned} \tag{3.4.27}$$

for finite times. This shows the desired L^1_x bound on $\sigma^t . \nabla u$. This more precisely shows that $\sigma^t . \nabla u \in L^{2,\infty}_t(L^1_x)$ as could be intuitively be expected from (3.4.16).

We notice, to conclude this section, that in dimension $d = 1$, the above proof shows that, for σ Lipschitz continuous, the estimate on $\sigma^t \nabla u$ in L^1 holds true along with the L^1 estimate on u.

3.5 "Intermediate" cases and, in particular, their relation to hypoellipticity

As noticed in Section 1.3.3, and this is of course not unexpected, the fact that, in the parabolic setting, we are able to prove well-posedness for transport fields more general than for pure transport equations is related to the gain in regularity we observe on the solution u. In particular, if we focus on the assumptions needed for the a priori estimate to hold and if we leave aside the assumptions, such as (2.1.1), specifically related to the regularization procedure, assumptions (2.1.2) are more general than the classical assumption $[\operatorname{div} \mathbf{b}]_- \in L^\infty$ because in the estimate of the term

$$\int \mathbf{b}\,.\,\nabla u\,.\,u = -\int \operatorname{div} \mathbf{b}\,.\,\frac{u^2}{2},$$

we may, on one side of the above equality or the other, use the Sobolev inequality (1.3.15)

$$\|u\|_{L^{2d/(d-2)}} \leq C\|\nabla u\|_{L^2},$$

which states an L^p integrability of the solution u for $p = \frac{2d}{d-2} > 2$, thereby allowing for $\mathbf{b}$ such that $\operatorname{div} \mathbf{b} \in L^{\frac{d}{2}}$, or satisfying the variant of this integrability condition mentioned in (2.1.2).

We would like to investigate in this section some situations for which, even though the second-order operator is not elliptic (and consequently does not allow for (1.3.15) to hold), this operator still provides some additional integrability of the solution u, in L^p for some $p > 2$, in the sense that (somewhat vaguely stated; see precise statements below)

$$\|u\|_{L^p} \leq C\|\sigma^t \nabla u\|_{L^2}. \tag{3.5.1}$$

This in turn allows for establishing an a priori estimate, only assuming $[\operatorname{div} \mathbf{b}]_- \in L^q$ for $q = \frac{p}{p-2}$ and not necessarily L^∞.

Likewise, in the vein of Proposition 2, we may extend the result to fields $\mathbf{b} = \beta + \sigma\Theta$ with $\Theta \in L^\infty$ but $[\operatorname{div} \beta]_-$ not necessarily in L^∞ and Θ not necessarily in L^∞ either.

As will be clear below, we will not attempt to write a general theory to answer such questions. We will proceed considering *some examples*, and then will show, in Section 3.5.3, how these examples are indeed *prototypical* and achieve some *genericness* for the problem considered.

Remark 38. Inequalities of type (3.5.1) are actually related to the property of ultracontractivity of the semi-group associated to the operator.

Before we proceed, we would like to make two comments on the assumptions we need.

Firstly, we emphasize that we assume, throughout this Section 3.5, that the field $\mathbf{b}$ has the regularity prescribed in (3.1.22), that is,

$$\mathbf{b} = \widetilde{\mathbf{b}} + \sigma\widetilde{\theta}, \quad \widetilde{\theta} \in L^2, \; \widetilde{\mathbf{b}} \in W^{1,1}.$$

The regularization procedure can then be applied to the equation in the standard way repeatedly described above. On the other hand, we concentrate ourselves on the derivation of the a priori estimate.

Secondly, we will assume throughout this Section 3.5 that

$$\sigma \in L^\infty. \tag{3.5.2}$$

The reason is the following. We will first establish inequalities of the type (3.5.1) for smooth, compactly supported functions defined on the whole space $\mathbb{R}^d$. This is the part of our argument we specifically outline in the remainder of this section. Once such an inequality is established, it is immediate to infer the specific inequality useful for our purpose. For brevity, we will actually skip that second step. Indeed, admit for the time being that we have proven that

$$\|v\|_{L^p} \le C\|\sigma^t \nabla v\|_{L^2}$$

for $v \in \mathcal{D}(\mathbb{R}^d)$. We may next extend the inequality by density (recall our Remark 32 on page 91) to all functions such that both sides are finite and next apply it to $v = u\varphi$, with u our tentative solution and φ a smooth compactly supported function that has value 1 on the domain $[0,1]^d$. The triangle inequality applied to $\nabla(u\varphi) = u\nabla\varphi + \varphi\nabla u$ then readily yields

$$\|u\|_{L^p([0,1]^d)} \le \|u\varphi\|_{L^p} \le C\|\sigma^t \nabla u\|_{L^2} + C\|u\|_{L^2}, \tag{3.5.3}$$

where the constant C only depends on universal constants, on the C^1 norm of the cut-off function φ, and on $\|\sigma\|_{L^\infty}$. This is where we make use of our additional assumption (3.5.2). This assumption is in particular satisfied by all the prototypical examples we will exhibit. The latter inequality (3.5.3) is the one we will use for estimating the term $-\int \operatorname{div} \mathbf{b} \,.\, \frac{u^2}{2}$ in the derivation of our a priori estimate. Note indeed that, for this latter purpose, we proceed with a Gronwall-type argument which is not perturbed by the addition of the rightmost term $C\|u\|_{L^2}$ in (3.5.3).

Before we concentrate, throughout this section, on which estimates of type (3.5.1) hold true in which setting, and how to establish those estimates, let us give slightly more details on how we use them.

The a priori estimate at the basis of our argument is (3.1.1), that is,

$$\frac{d}{dt}\int \frac{u^2}{2} + \int \frac{u^2}{2} \operatorname{div} \mathbf{b} + \frac{1}{2}\int |\sigma^t \nabla u|^2 = 0.$$

One way to proceed is as follows:

$$\frac{d}{dt}\int \frac{u^2}{2} + \frac{1}{2}\int |\sigma^t \nabla u|^2 = -\int \frac{u^2}{2} \operatorname{div} \mathbf{b} \le \frac{1}{2}\|\operatorname{div} \mathbf{b}\|_{L^{p/(p-2)}} \|u\|_{L^p}^2.$$

If we temporarily admit (3.5.3), then we obtain

$$\frac{d}{dt}\int \frac{u^2}{2} + \frac{1}{2}\int |\sigma^t \nabla u|^2 \le C\|\operatorname{div} \mathbf{b}\|_{L^{p/(p-2)}} \|\sigma^t \nabla u\|_{L^2}^2 + C\|\operatorname{div} \mathbf{b}\|_{L^{p/(p-2)}} \|u\|_{L^2}^2, \tag{3.5.4}$$

which, for div $\mathbf{b}$ small in $L^{\frac{p}{p-2}}$ gives us the result. But that assumption may readily be extended into div $\mathbf{b}$ arbitrarily large in $L^{\frac{p}{p-2}}$ since, by a standard argument we already employed in the course of the proof of Theorem 1, we may always write $\operatorname{div}\mathbf{b} = f_1 + f_2$ with f_1 small in $L^{\frac{p}{p-2}}$ and $f_2 \in L^\infty$, the latter term being easily treated in the a priori estimate. The assumption

$$\operatorname{div}\mathbf{b} \in L^{\frac{p}{p-2}} \tag{3.5.5}$$

thus allows for the results to hold. Next, we also notice that, by the same argument, we may also accommodate in the equation a term $\sigma\Theta\,.\,\nabla u$, since it amounts to adding $\int \sigma\Theta\,.\,\nabla u\, u$ to the right-hand side of (3.5.4), a term which is in turn bounded from above as follows:

$$\begin{aligned}\left|\int \sigma\Theta\,.\,\nabla u\, u\right| &= \left|\int \Theta\,.\,\sigma^t\nabla u\, u\right| \\ &\le \|\Theta\|_{L^{\frac{2p}{p-2}}}\|\sigma^t\nabla u\|_{L^2}\|u\|_{L^p} \\ &\le C\|\Theta\|_{L^{\frac{2p}{p-2}}}\|\sigma^t\nabla u\|_{L^2}^2 + C\|\Theta\|_{L^{\frac{2p}{p-2}}}\|u\|_{L^2}^2.\end{aligned}$$

We observe that, again first for Θ small in $L^{\frac{2p}{p-2}}$ and then for any $\Theta \in L^{\frac{2p}{p-2}}$ using the same argument as above, we may conclude. The assumption

$$\mathbf{b} = \Theta\sigma \quad \text{with } \Theta \in L^{\frac{2p}{p-2}} \tag{3.5.6}$$

is therefore also convenient. Further extensions of the above conditions (3.5.5) or (3.5.6), using Lorentz spaces, may also be obtained.

Remark 39. In the hypoelliptic case examined in Section 3.5.1, the proof of estimates, such as (3.5.1), made precise in the statement of Lemma 11 below, rests on an intermediate step that proves some additional (fractional) differentiability of u. In the simplest case, this is (3.5.18). When the first-order term is involved, this differentiability is expressed in (3.5.24). It might be possible, but, in the present state of our understanding, it is unclear to us how this should be done, to use this extra differentiability in order to *also* weaken the assumptions needed for the regularization step (in the same spirit as we are able to perform the regularization for more general fields $\mathbf{b}$ when the second-order operator is positive definite).

3.5.1 Hypoellipticity in the non-regular setting

Hypoellipticity with the second-order term. The first case we are going to consider here is the case when the tensor-valued function σ defining our second-order operator in divergence form $-\frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j u)$ satisfies the following (pointwise) inequality

$$|\sigma^t\nabla u|^2 \ge \nu\big(|\nabla_x u|^2 + |x|^{2\beta}|\nabla_y u|^2\big) \quad \text{for all } (x,y) \in \mathbb{R}^k \times \mathbb{R}^m = \mathbb{R}^d \tag{3.5.7}$$

for some constant $\nu > 0$, some (not necessarily integer) exponent $\beta > 0$, and for all smooth compactly supported functions u defined on $\mathbb{R}^d$. Evidently, since the function u

is arbitrary, this property is an intrinsic property of the function σ and may also be written

$$|\sigma^t(x,y)Z|^2 \geq \nu(|X|^2 + |x|^{2\beta}|Y|^2), \quad Z = (X,Y),\ z = (x,y) \in \mathbb{R}^k \times \mathbb{R}^m = \mathbb{R}^d, \quad (3.5.8)$$

for some constant $\nu > 0$ and some (not necessarily integer) exponent $\beta > 0$, both independent of Z. The typical operator satisfying this condition is of course

$$-\frac{1}{2}\partial^2_{xx}u - \frac{1}{2}\partial_y(|x|^{2\beta}\partial_y u). \quad (3.5.9)$$

This particular situation is actually related to the theory of hypoellipticity. We briefly recall now some of the basic ingredients of this classical theory which we will need for the sake of comparison with the arguments we shall develop in our non-regular setting.

The notion of *hypoellipticity* has been introduced by L. Schwartz. It is said that a differential operator $A = \sum_{|\alpha|\leq m} a_\alpha(x)\partial_x^\alpha$ (where we obviously denote by α a generic multi-index of differentiation, a_α the corresponding coefficient, and m the order of the operator) is hypoelliptic (at the point x_0) if the following property holds:

$$(Au \text{ smooth around } x_0) \implies (u \text{ smooth around } x_0). \quad (3.5.10)$$

Of course, an elliptic operator is hypoelliptic. An elliptic operator of order m satisfies (we argue locally) the *elliptic estimate*

$$\|u\|_{H^{s+m}} \leq C(\|Au\|_{H^s} + \|u\|_{H^s}).$$

An hypoelliptic operator, on the other hand, may satisfy the *subelliptic estimate* with a loss of $m - m'$ derivatives:

$$\|u\|_{H^{s+m'}} \leq C(\|Au\|_{H^s} + \|u\|_{H^s}) \quad (3.5.11)$$

for some $0 < m' < m$. If (3.5.11) holds for all $s \geq 0$ with m' independent of s, then the operator is actually hypoelliptic.

A sufficient condition for hypoellipticity has been established in the works of L. Hörmander. Assume that the operator A is a sum $A = \sum_{j=1}^d X_j^2$ of vector fields X_j. That way of decomposing the operator is usually called the Hörmander notation. Consider the Lie algebra generated by all the Lie brackets

$$[X_{j_0}, [X_{j_1}, \ldots [X_{j_{m-1}}, X_{j_m}] \ldots]$$

for $\deg(j_0, \ldots, j_m) \leq r$, where $\deg(j_0, \ldots, j_m)$ denotes the sum of the degrees of each of the vector fields X_j (that is, usually, one). Assume that this Lie algebra is full for some r, that is, is equal to the ambient space $\mathbb{R}^d$. Then A is hypoelliptic (at the point x_0) and the subelliptic estimate (3.5.11) holds with $m' = \frac{2}{1+r}$. A variant of that condition exists for operators of the form $A = X_0 + \sum_{j=1}^d X_j^2$ (in which case the degree of X_0 is usually two, while that of all the other X_j, $j \geq 1$, is one). Note that, in both cases, iterating (3.5.11) gives the smoothness required in the definition (3.5.10).

A simple example of an hypoelliptic operator that is not elliptic is

$$Au = -\partial^2_{xx}u - \partial_y(|x|^{2n}\partial_y u)$$

for some integer $n \geq 1$, which we can consider in dimension $d = 2$ for simplicity. Obviously, that operator is not elliptic since $|x|^{2n}$ vanishes at the origin. We however observe that $[\partial_x, |x|^n\partial_y] = n|x|^{n-2}x\partial_y$, thus repeating this bracket n times gives the field ∂_y, when n is even (when n is odd, consider $|x|^{n-1}x\partial_y$ instead of $|x|^n\partial_y$). The algebra generated at order n is the whole ambient space $\mathbb{R}^2$. The operator is thus hypoelliptic, and the corresponding subelliptic estimate holds. For more information on the classical theory of hypoellipticity, we refer, in the case of constant-coefficients operators, to the landmark work [55] (in particular Chapter XI), and, for the general case, to [56] (in particular Chapter XXII), and to the many references therein. We also mention the classical article [85]. We notice that, although very attractive theoretically, the definition (3.5.10) remains very difficult to manipulate per se, without further explicit conditions such as the conditions on Lie brackets recalled above.

Our assumption (3.5.7) can therefore be seen as the natural extension to non-integer exponents β of the classical hypo-ellipticity theory for the operator

$$-\frac{1}{2}\partial^2_{xx}u - \frac{1}{2}\partial_y(|x|^{2n}\partial_y u). \tag{3.5.12}$$

We notice that one could consider performing a regularization of $|x|^{2\beta}$ in (3.5.9) (or, likewise, of any of our coefficients σ of the present notes), so as to apply the classical theory to the regularized operator and pass to the limit. This course of action is actually unclear, in particular because of the technicalities expected in the entangled treatment of the Lie brackets and the regularized coefficients of the operator. We prefer to follow another pathway.

In order to proceed under assumption (3.5.7), all we need is the following result.

Lemma 11. *For all smooth, compactly supported functions u defined on all points $z = (x, y) \in \mathbb{R}^k \times \mathbb{R}^m = \mathbb{R}^d$, we have*

$$\|u\|_{L^s} \leq c\big(\|\nabla_x u\|_{L^2} + \||x|^\beta \nabla_y u\|_{L^2}\big), \tag{3.5.13}$$

where $s = \frac{2D}{D-2}$ and $D = k + m(\beta + 1)$. Note that D plays the role of an “effective” dimension, and that when $\beta = 0$, we have $D = k + m = d$ and we thus recover the classical Sobolev inequality (1.3.15).

Using Lemma 11, it is immediate, using our comments at the beginning of this section and arguments similar to (1.3.14) and (3.1.27), to establish the *a priori estimate* under assumptions (3.5.7) and

$$\mathbf{b} = \beta + \sigma\Theta, \quad [\operatorname{div}\beta]_- \in L^{\frac{D}{2}}, \ \Theta \in L^\infty, \tag{3.5.14}$$

where D is defined in the statement of Lemma 11. Next, it is also straightforward to obtain well-posedness when we additionally assume (3.1.22) and (3.5.2). Note also that, arguing as we did for (2.1.12)–(2.1.13), we may further generalize assumption (3.5.14) into

$$\mathbf{b} = \beta + \sigma\Theta, \quad [\operatorname{div}\beta]_- \in \varepsilon L^{\frac{D}{2},\infty} + L^{\frac{D}{2}}, \ \Theta \in L^\infty. \tag{3.5.15}$$

This requires to show that inequality (3.5.13) also holds when using the norm $L^{s,2}$, instead of L^s, on the left-hand side, an extension which is clear on our outlined proof of page 30 since one may replace the first-order gradient ∇ by whatever first derivative operator for which a classical Sobolev-type inequality holds and follow step by step the same argument.

Proof of Lemma 11 (outline). Using a partial Fourier transform in the variable $y \in \mathbb{R}^m$, we know, by Parseval's identity, that the square of the right-hand side of (3.5.13) is equivalent to

$$\iint (|\nabla_x v|^2 + |x|^{2\beta}|\eta|^2|v|^2)\,dx\,d\eta, \tag{3.5.16}$$

where $v(x,\eta)$ denotes the partial Fourier transform of $u(x,y)$ in y. We next introduce, for all $\eta \in \mathbb{R}^m$ fixed, the minimization problem

$$\lambda_1(\eta) = \inf\left\{E(w) := \int (|\nabla_x w|^2 + |x|^{2\beta}|\eta|^2|w|^2)\,dx,\ \int_{\mathbb{R}^k} w^2 = 1,\ w \in H^1(\mathbb{R}^k)\right\}. \tag{3.5.17}$$

Since obviously, $\lambda_1(\eta)$ only depends on the norm $|\eta|$ and not on η itself, we denote by $\lambda_1(1)$ the value of $\lambda_1(\eta)$ for whichever η with norm one. We notice that $\lambda_1(1) > 0$. Indeed, a minimizing sequence w_n of (3.5.17) is clearly bounded in $H^1(\mathbb{R}^k)$, thus, extracting a subsequence if necessary, it weakly converges in that space, to some $w \in H^1(\mathbb{R}^k)$. The functional E being weakly lower semi continuous in $H^1(\mathbb{R}^k)$, proving $\lambda_1(1) > 0$ amounts to proving that $\int w^2 = 1$. To this end, we observe that

$$\int_{|x|\geq R} w_n^2\,dx \leq R^{-2\beta} \int_{|x|\geq R} |x|^{2\beta} w_n^2\,dx \leq R^{-2\beta}(\lambda_1(1) + 1)$$

for n sufficiently large. Applying the Rellich Theorem, which shows the strong convergence in $L^2(B_R)$, this proves $\int_{\mathbb{R}^k} w^2 = 1$. Since the substitution $w \to w_L(\cdot) = L^{-\frac{k}{2}} w(\frac{\cdot}{L})$ does not change the L^2 norm of w and since, by an easy change of variable $x \to \frac{x}{L}$ in the integrand,

$$E(w_L) = L^{-2} \int (|\nabla_x w|^2 + |x|^{2\beta}|L^{1+\beta}\eta|^2|w|^2)\,dx,$$

we necessarily have $\lambda_1(\eta) = L^{-2}\lambda_1(L^{1+\beta}\eta)$, thus

$$\lambda_1(\eta) = |\eta|^{\frac{2}{1+\beta}}\lambda_1(1).$$

We next use this information to bound (3.5.16) from below:

$$\iint (|\nabla_x v|^2 + |x|^{2\beta}|\eta|^2|v|^2)\,dx\,d\eta \geq c \int |\eta|^{\frac{2}{1+\beta}} \int_{\mathbb{R}^k} v^2\,dx\,d\eta$$

for some $c > 0$, which, using inverse Fourier transform, shows

$$\|\nabla_x u\|_{L^2} + \big\||x|^\beta \nabla_y u\big\|_{L^2} \geq c\|\nabla_y^\theta u\|_{L^2}^2 \tag{3.5.18}$$

for $\theta = \frac{1}{1+\beta}$. Since the same quantity trivially bounds $c\|\nabla_x u\|_{L^2}^2$ from above, we obtain a minoration by the (squared) norm in the space

$$L^2(\mathbb{R}^k, H^\theta(\mathbb{R}^m)) \cap L^2(\mathbb{R}^m, H^1(\mathbb{R}^k))$$

again for $\theta = \frac{1}{1+\beta}$. Using the Cauchy–Schwarz inequality, we have

$$L^2(\mathbb{R}^m, H^1(\mathbb{R}^k)) \subset H^1(\mathbb{R}^k, L^2(\mathbb{R}^m)),$$

thus we obtain a minoration in

$$L^2(\mathbb{R}^k, H^\theta(\mathbb{R}^m)) \cap H^1(\mathbb{R}^k, L^2(\mathbb{R}^m)).$$

Using next the Sobolev embedding $H^1(\mathbb{R}^k) \subset L^a(\mathbb{R}^k)$ with $a = \frac{2k}{k-2}$ and $H^\theta(\mathbb{R}^m) \subset L^b(\mathbb{R}^m)$ for $b = \frac{2m}{m-2\theta}$, we obtain a minoration in the norm of the space

$$L^2(\mathbb{R}^k, L^b(\mathbb{R}^m)) \cap L^a(\mathbb{R}^k, L^2(\mathbb{R}^m))$$

for the specific values of a and b mentioned above. We now interpolate

$$\begin{cases} \dfrac{1}{s} = \dfrac{\alpha}{2} + \dfrac{1-\alpha}{a}, \\ \dfrac{1}{s} = \dfrac{\alpha}{b} + \dfrac{1-\alpha}{2} \end{cases}$$

for some $\alpha \in (0,1)$ and obtain the minoration in $L^s(\mathbb{R}^{k+m})$ for the precise value of s defined in the statement of Lemma 11. Note that, by homogeneity, we may independently calculate s as the only possible exponent such that changing $u(x,y)$ into $u(\lambda^{-1}x, \mu^{-1}y)$ does not affect the inequality

$$\|u\|_{L^2(\mathbb{R}^k, L^b(\mathbb{R}^m)) \cap L^a(\mathbb{R}^k, L^2(\mathbb{R}^m))} \geq c\|u\|_{L^s(\mathbb{R}^{k+m})}.$$

Indeed, equating the scaling of the two sides, we necessarily have

$$\lambda^{\frac{k}{2}} \mu^{\frac{m-2\theta}{2}} + \lambda^{\frac{k-2}{2}} \mu^{\frac{m}{2}} \gtrsim \lambda^{\frac{k}{s}} \mu^{\frac{m}{s}},$$

thus

$$s = \frac{2(m+k\theta)}{(m+k\theta) - 2\theta} = \frac{2(k+m(1+\beta))}{(k+m(1+\beta)-2)},$$

which is the value mentioned in the statement of the lemma. This concludes the proof of this lemma. □

Before we proceed to our second example, we would like to make remarks in two additional directions related to hypoellipticity.

Firstly, we notice that, in Lemma 11, we have only established a better integrability of the solution u. In the classical hypoellipticity theory, this integrability (see (3.5.11)) is only the first step of the argument. Bootstrapping next this information in the equation, and *differentiating* the equation at every order, it is then proved that the solution is indeed smooth. In our non-regular setting, we are not able to perform this sequence of arguments.

Secondly, we have only considered the second-order operator in our assumption (3.5.7), and in Lemma 11. We have not paid any attention to the first-order operator present in the equation, namely $\mathbf{b}\,.\,\nabla u$. Again, in the classical hypoellipticity theory, this first-order term may be considered, and in some situations, critically contributes to the regularity of the solution. A famous example of the type of phenomenon observed is given by the *Vlasov Fokker–Planck equation*

$$\partial_t h + \mathbf{v}\,.\,\nabla_x h - \Delta_{\mathbf{v}} h + F(\mathbf{v})\,.\,\nabla_{\mathbf{v}} h = 0 \tag{3.5.19}$$

set in the variables time t, space x and velocity $\mathbf{v}$, for a certain force field F. That equation is not elliptic since only a Laplacian in the velocity variables is present. However, it is hypoelliptic because the algebra generated by the Lie brackets is full: $\nabla_{\mathbf{v}}$ is present in the second-order term and $\nabla_x = [\nabla_{\mathbf{v}}, \mathbf{v}\nabla_x]$ is obtained with the help of the first-order term. A simpler, academic model example for that situation is the equation

$$-(\partial_x + \partial^2_{yy})u = f. \tag{3.5.20}$$

For that particular operator, a direct a priori estimate yields

$$\|\partial_y u\|_{L^2} \le \|f\|_{L^2}$$

since the first-order term does not contribute to it. However, the algebra generated by the Lie brackets is full, precisely owing to the presence of the operator ∂_x. A better regularity than that suggested by the a priori estimate may be obtained. We will be more precise below. It turns out that in our theoretical developments, and at least in the setting of our *elliptic* variant, we may also take benefit of the presence of some first-order operator. We now briefly expose some arguments in this direction.

Hypoellipticity using the first-order term. Consider the elliptic variant (1.3.6) of our parabolic problem, to which we add a sufficiently large zeroth-order term $+\alpha u$ for reasons that have been made clear on page 23:

$$\alpha u - \mathbf{b}\,.\,\nabla u - \frac{1}{2}\,\mathrm{div}(\sigma\sigma^t\nabla u) = f \tag{3.5.21}$$

for some right-hand side $f \in L^2$. We then notice, using the definition (3.1.18) of the functional space H, that

$$\begin{aligned}
\|\mathbf{b}\,.\,\nabla u\|_{H'} &= \left\|\alpha u - \frac{1}{2}\,\mathrm{div}(\sigma\sigma^t\nabla u) - f\right\|_{H'} \\
&= \sup_{v\in H} \frac{\langle \alpha u - \frac{1}{2}\,\mathrm{div}(\sigma\sigma^t\nabla u) - f, v\rangle_{H',H}}{\|v\|_H} \\
&\le \sup_{v\in H} \frac{\alpha\|u\|_{L^2}\|v\|_{L^2} + \frac{1}{2}\|\sigma^t\nabla u\|_{L^2}\|\sigma^t\nabla v\|_{L^2} + \|f\|_{L^2}\|v\|_{L^2}}{\|v\|_H},
\end{aligned}$$

whence

$$\|\mathbf{b}\,.\,\nabla u\|_{H'} \le \alpha\|u\|_{L^2} + \frac{1}{2}\|\sigma^t\nabla u\|_{L^2} + \|f\|_{L^2} \le C\|u\|_H + \|f\|_{L^2}. \tag{3.5.22}$$

Using this estimate (3.5.22), we may bootstrap better integrability of the solution u, which may allow to consider more general transport fields $\mathbf{b}$. The assumption we will consider on the couple $(\mathbf{b}, \sigma)$ in this nonregular setting will therefore write

$$c\|u\|_{L^p} \le \|\mathbf{b}\,.\,\nabla u\|_{H'} + \|u\|_H \tag{3.5.23}$$

for some $p > 2$. This assumption plays the role of the hypoellipticity assumption of the classical setting (guaranteed by the completeness of the Lie algebra generated by the brackets). Assume that $u \in H$ is a solution to (3.5.21) for $f \in L^2$. We deduce from (3.5.22) and (3.5.23) that $u \in L^p$, whence, by a now usual argument, that $\operatorname{div}\mathbf{b}$ needs not be in L^∞ but only in $L^{\frac{2p}{p-2}}$, along with the possible usual variants. Note that, short of a smallness or a structure assumption on $\mathbf{b}$, we cannot prove *ex nihilo* that $u \in H$; we have to *assume* that fact.

The academic example (3.5.20) above gives the prototypical argument and provides a guideline for the general case. In that case, $\mathbf{b}\,.\,\nabla u = \partial_x u$, $\sigma^t \nabla u = \partial_y u$,

$$H = \{u \in L^2 : \partial_y u \in L^2\} = L^2_x(W^{1,2}_y),$$

and we have just obtained with (3.5.22) that $\partial_x u \in L^2_x(W^{-1,2}_y)$. This fact is, of course, evident, directly from equation (3.5.21) itself and the associated a priori estimate

$$\alpha\|u\|_{L^2} + \|\partial_y u\|_{L^2} \le C\|f\|_{L^2}$$

in that simple case. From $\partial_x u \in L^2_x(W^{-1,2}_y)$ together with $u \in L^2_x(W^{1,2}_y)$, we deduce by interpolation that $\partial_x^{\frac{1}{2}} u \in L^2_{x,y}$, and thus

$$u \in W^{\frac{1}{2},2}_{x,y}. \tag{3.5.24}$$

The proof is immediate in the language of Fourier transforms since $|\xi_y|^{-1}|\xi_x|\hat{u} \in L^2$ and $|\xi_y|\hat{u} \in L^2$ together imply, by the Cauchy–Schwarz inequality, that $\sqrt{|\xi_x|}\hat{u} \in L^2$. Subsequently, it follows from $u \in W^{\frac{1}{2},2}_{x,y}$ that $u \in L^{\frac{2d}{d-1}}$, a fact that can be used to weaken the assumptions on $\mathbf{b}$ for our well-posedness theory. For instance, using arguments repeatedly mentioned above, we may take $\operatorname{div}\mathbf{b} \in L^d$, instead of L^∞, and keep all the results.

We conclude this paragraph mentioning that the phenomena we have just discussed are delicate and sensitive to the equation. Whereas the situation in the elliptic case and the parabolic case are often comparable (in line with our comments of page 23), there is some subtle difference here. In contrast to the elementary elliptic example (3.5.20), where an extra regularization is obtained, that regularization is not obtained on the parabolic variant of that equation, namely

$$\frac{\partial u}{\partial t} - (\partial_x + \partial^2_{yy})u = 0, \tag{3.5.25}$$

since, in that case, the solution corresponding to the specific initial condition $u_0(x, y) = \varphi(x)$ reads as $u(t, x) = \varphi(t + x)$. No regularization effect can be expected. On the other hand, when it comes to the Vlasov equation (3.5.19), say with $F \equiv 0$,

$$\partial_t h + \mathbf{v}\,.\,\nabla_x h - \Delta_{\mathbf{v}} h = 0,$$

an equation not so much different from (3.5.25), then we have hypoellipticity.

3.5.2 A subelliptic situation

We now examine an example of a situation different from the "hypoelliptic" situation (3.5.7). More precisely, we assume, for $0 < \beta < 1$,

$$|\sigma^t \nabla u|^2 \geq \nu |x|^{2\beta} |\nabla u|^2 \quad \text{for all } x \in \mathbb{R}^d, \tag{3.5.26}$$

or, equivalently,

$$|\sigma^t(x) X|^2 \geq \nu |x|^{2\beta} |X|^2 \quad \text{for all } (X, x) \in \mathbb{R}^d \times \mathbb{R}^d. \tag{3.5.27}$$

For $\beta \neq 0$ integer, this setting is not hypoelliptic, since we cannot expect any additional regularity, given that the coefficient vanishes at least quadratically at the origin. But when β is not an integer, there is room for some extra regularity. It is actually a case of *subellipticity*.

We now claim that the following result holds, in the vein of Lemma 11.

Lemma 12. *For all smooth, compactly supported functions u defined on all $x \in \mathbb{R}^d$, we have, for $s = \frac{2d}{d+2\beta-2}$,*

$$\|u\|_{L^s} \leq c \big\| |x|^\beta \nabla_x u \big\|_{L^2}. \tag{3.5.28}$$

Temporarily admitting that this lemma holds, we see that it allows for establishing our usual a priori estimate, under assumption (3.5.26) together with

$$\mathbf{b} = \beta + \sigma\Theta, \quad [\operatorname{div}\beta]_- \in L^{\frac{d}{2(1-\beta)}}, \; \Theta \in L^\infty, \tag{3.5.29}$$

or, using the usual extension and the *proof* of Lemma 12 below (see (3.5.32)), the analogous assumption with $[\operatorname{div}\beta]_- \in \varepsilon L^{\frac{d}{2(1-\beta)},\infty} + L^{\frac{d}{2(1-\beta)}}$.

Proof of Lemma 12 (outline). Inequality (3.5.28) is a particular case of the Caffarelli–Kohn–Nirenberg inequalities [28, 31] and we prove it here for consistency. We first notice that, as above, a dimensionality argument gives the necessary value of the exponent s. Indeed, changing u into $u(\lambda^{-1}x)$ should not modify the inequality (3.5.28) and this can only hold when $\frac{d}{s} = \frac{1}{2}(d + 2\beta - 2)$ which yields the value of s mentioned in the Lemma.

We assume in our proof that the ambient dimension d is larger than or equal to 2. Therefore, because $0 < \beta < 1$, we have $2 < s < +\infty$. The case $d = 1$ can be addressed separately.

One possible argument for now establishing (3.5.28) is to first use the generalized Hölder inequality (2.0.11) as follows:

$$\|\nabla u\|_{L^{p,2}} = \left\| \frac{1}{|x|^\beta} \cdot |x|^\beta \nabla u \right\|_{L^{p,2}} \leq C \left\| \frac{1}{|x|^\beta} \right\|_{L^{\frac{d}{\beta},\infty}} \big\| |x|^\beta \nabla u \big\|_{L^{2,2}} \tag{3.5.30}$$

for $\frac{1}{p} = \frac{\beta}{d} + \frac{1}{2}$. Next, we use the following inequality:

$$\|u\|_{L^{q,2}} \leq C \|\nabla u\|_{L^{p,2}} \tag{3.5.31}$$

for $\frac{1}{q} = \frac{1}{p} - \frac{1}{d}$. The latter inequality (3.5.31) is obtained by an argument similar to that for establishing (2.0.13). It amounts to observing that $u = -\Delta u \star G = -\partial_i u \star \partial_i G$

with G the Green function of the Laplacian, whence $\partial_i G \in L^{\frac{d}{d-1},\infty}$, and next to apply the generalized Young inequality (2.0.12) with $1 + \frac{1}{q} = \frac{1}{p} + \frac{(d-1)}{d}$. We have already indicated how to rigorously formalize this schematic argument. Combining (3.5.30) and (3.5.31), we obtain

$$\|u\|_{L^{q,2}} \leq C\big\||x|^{\beta}\nabla u\big\|_{L^2} \tag{3.5.32}$$

for $\frac{1}{q} = \frac{1}{p} - \frac{1}{d} = \frac{\beta-1}{d} + \frac{1}{2} = \frac{1}{s}$, which implies (3.5.28) since $s > 2$ thus $L^{s,2} \subset L^{s,s} = L^s$. This concludes the proof of Lemma 12. □

3.5.3 Toward genericness

We would like to conclude this section by commenting upon the degree of genericness that the above two examples have. In some sense, these two examples are representative of all the possible interesting situations, at least in the two-dimensional space $\mathbb{R}^2$. In higher dimensions, geometry definitely plays a role and other phenomena than those discussed below may appear. To some extent, however, they are only technicalities for the issues we are focusing on.

Let us be given a second-order operator in the form

$$\sum_{i,j=1}^{2} a_{ij}\partial^2_{ij} \tag{3.5.33}$$

for a nonnegative symmetric matrix valued coefficient $\mathbf{a}$, as is the case throughout these notes. We focus on the situation near the origin $(x_1, x_2) = (0, 0)$, since we know, by arguments we have made precise above, that we may always *localize* our considerations, using a suitable cut-off function, a density argument, and the degree of freedom provided by the Gronwall-type argument we typically use for establishing our a priori estimate. We immediately diagonalize this operator at the origin, and put it under the form $a_1(x_1, x_2)\partial_{11} + a_2(x_1, x_2)\partial_{22}$, with $a_1 \geq 0$, $a_2 \geq 0$.

Clearly, if neither a_1 nor a_2 vanishes at the origin, standard elliptic theory applies and we may proceed as if our operator were the Laplacian. On the other hand, if both coefficients a_1 and a_2 vanish at zero, they vanish, at least, at quadratic order there (because of the nonnegativeness) and so the operator is typically dominated by $|x|^2\Delta$. There can be no regularization provided by such an operator and again, the situation is uninteresting. The only hope, however, is that a first-order term compensates for this cancellation of second-order at the origin. This is exactly what happens when considering

$$|x|^2\Delta u + 2x\,.\,\nabla u = \partial_1(|x|^2\partial_1) + \partial_2(|x|^2\partial_2),$$

which is precisely the subelliptic situation we have considered in (3.5.26). When we restrict our attention to the second-order operator of the form (3.5.33), the only interesting case is thus the case when exactly one of the coefficients a_1 or a_2 vanishes at the origin. We are thus left with the typical behavior ($a_1 \equiv 1$, $a_2 = a^2(x_1, x_2)$), that is, the operator $\partial_{11} + a^2(x_1, x_2)\partial_{22}$.

Hypoellipticity is concerned with the case $\partial_1 a(0,0) \neq 0$. Otherwise,

$$[\partial_1, a(x_1,x_2)\partial_2] = (\partial_1 a)\partial_2$$

vanishes at the origin and there is no hope to have a full Lie algebra. Put differently, and since only the first two derivatives at the origin matter, the generic case is thus of the form $\partial_{11} + (x_1 + \lambda x_2)^2 \partial_{22}$. Up to an affine change of variable, performed *locally* at the origin, this is also $\partial_{11} + |x_1|^2 \partial_{22}$, thus exactly the case we have considered in (3.5.12), for non-integer exponents, namely $\partial_{11} + |x_1|^{2\beta} \partial_{22}$, and which we call our "hypoelliptic" situation in the non-regular setting.

In a similar spirit, we may notice that in all our arguments of this section and the preceding one, the prototypical factor $|x|^\beta$ in terms such as $|x|^\beta \nabla_1 u$ or $|x|^\beta \nabla_2 u$ does not play any intrinsic role. Although we have not worked out all the details, a term of the form $a(x_1,x_2)\nabla u$ is likely to be equally well accommodated, provided this term is such that $\frac{1}{a}$ enjoys the suitable integrability properties.

4 Extensions

4.1 Non-divergence form equations

We devote this section to the study of equation (1) when it is *not* in the conservative form (3.0.2) (or equivalently (3.0.1)). We first mention immediate corollaries of the results obtained above for the conservative form, and then proceed to more specific arguments that show further results.

4.1.1 Immediate extensions of the results previously obtained

We remark that the equation

$$\partial_t u - b_i \partial_i u - a_{ij} \partial^2_{ij} u = 0$$

also reads

$$\partial_t u - b_i^{\mathrm{mod}} \partial_i u - \partial_i (a_{ij} \partial_j u) = 0, \tag{4.1.1}$$

where we have denoted by

$$\mathbf{b}^{\mathrm{mod}} = \mathbf{b} - \operatorname{div} \mathbf{a}, \tag{4.1.2}$$

i.e. in coordinates $b_i^{\mathrm{mod}} = b_i - \partial_j a_{ij}$ (recall that $\mathbf{a}$ is symmetric). The divergence of this modified drift reads

$$\operatorname{div} \mathbf{b}^{\mathrm{mod}} = \partial_i b_i - \partial^2_{ij} a_{ij}.$$

Remark 40. When $\mathbf{a} = \frac{1}{2} \sigma \sigma^t$, note that this modified drift vector $\mathbf{b}^{\mathrm{mod}}$ is (in general) different from the drift vector that is obtained when considering the Stratonovich form of the stochastic differential equation (1.2.1), namely

$$d\mathbf{X}_t = \mathbf{b}^{\mathrm{Strat}}(\mathbf{X}_t)\, dt + \sigma(\mathbf{X}_t) \circ d\mathbf{W}_t, \tag{4.1.3}$$

where

$$(\mathbf{b}^{\mathrm{Strat}})_i = b_i - \frac{1}{2} \sigma_{jk} \partial_j \sigma_{ik}. \tag{4.1.4}$$

The difference indeed writes $(\mathbf{b}^{\mathrm{mod}} - \mathbf{b}^{\mathrm{Strat}})_i = -\frac{1}{2} \sigma_{ik} \partial_j \sigma_{jk}$. As we already have mentioned in [65, Remark 18], it is unclear to us how to interpret the role of the modified drift vector $\mathbf{b}^{\mathrm{mod}}$ in terms of the theory of stochastic differential equations, and why this drift, instead of the Stratonovich drift vector $\mathbf{b}^{\mathrm{Strat}}$, naturally appears. A longer discussion may be found in [65].

It follows from the above observations that the results of the previous sections readily apply to (1) up to the modification of $\mathbf{b}$ into $\mathbf{b}^{\mathrm{mod}}$.

Of course, when $\mathbf{a} = \frac{1}{2} \sigma \sigma^t$ (or is like in (3.0.4)), all the extensions of the results are still valid *modulo* the transformation of the drift vector using a Girsanov-type transform as in Proposition 2.

https://doi.org/10.1515/9783110635508-004

Remark 41. Similar considerations hold on the Fokker–Planck equation

$$\partial_t u + \partial_i(u b_i) - \frac{1}{2}\partial_{ij}^2(\sigma_{ik}\sigma_{jk}u) = 0,$$

which may be written

$$\partial_t u + \partial_i(u b_i^{\mathrm{mod}}) - \frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j u) = 0, \tag{4.1.5}$$

and analogous results follow.

4.1.2 The non-divergence case per se

The immediate extension we have just obtained at the price of modifying the drift is interesting, but definitely not optimal. It is in particular concerning that conditions on *second-order derivatives* of $\mathbf{a}$, or σ, appear. Note indeed that our assumptions of the previous section, namely (3.1.4), involve the first derivatives of the drift and now apply to $\mathbf{b}^{\mathrm{mod}}$ defined by (4.1.2), a quantity where $\operatorname{div}\mathbf{a}$ appears.

We would like to now turn to an approach that, in the particular case $\mathbf{a} = \frac{1}{2}\sigma\sigma^t$ (or its extension mentioned above), allows to establish existence and uniqueness with assumptions on *only* the first derivatives. Although the equation we consider in this section is *not* necessarily in divergence form, we show it is associated, when $\mathbf{a} = \frac{1}{2}\sigma\sigma^t$, to a natural energy estimate (see (4.1.11) below), involving only these first derivatives. We likewise show that the regularization may be performed when assuming only some integrability of some first derivatives.

As a matter of fact, the key ingredient for the energy estimate we are about to establish has already been introduced in our earlier work [65, Section 8.1]. We made use there of a specific manipulation of the second-order term. It was in particular employed in [65] to establish existence and uniqueness for an H^1 initial condition and $\mathbf{b}$ and σ both Lipschitz continuous. For simplicity of exposition, we take $\mathbf{b} = 0$ and will argue on the L^2 a priori estimate, but our discussion below is entirely general. We will state our main result, Proposition 4 below, reinstating the transport field $\mathbf{b}$. Let us at once emphasize that our main new assumption, which will allow for our proof to hold, reads (see (4.1.14) below)

$$\operatorname{div}\sigma \in L^\infty, \quad \operatorname{Tr}[(D\sigma)^2] \le C,$$

where the divergence and trace operators are defined in (4.1.12)–(4.1.13) below.

A priori estimate. The L^2 formal a priori estimate associated to (1) (with, we recall, we assume $\mathbf{b} = 0$ for simplicity of exposition) writes

$$\frac{d}{dt}\int \frac{u^2}{2} - \int a_{ij}\partial_{ij}u u = 0,$$

that is,

$$\frac{d}{dt}\int u^2 - \int \sigma_{ik}\sigma_{jk}\partial_{ij}u u = 0,$$

where we have, for each triple (i, j, k),

$$\int \sigma_{ik}\sigma_{jk}\partial_{ij}uu = -\int \partial_j(\sigma_{ik}\sigma_{jk})\partial_i\frac{u^2}{2} - \int \sigma_{ik}\sigma_{jk}\partial_i u\partial_j u. \tag{4.1.6}$$

We then observe that

$$\int \partial_j(\sigma_{ik}\sigma_{jk})\partial_i\frac{u^2}{2} = 2\int(\partial_i\sigma_{ik})(\sigma_{jk}\partial_j u)u + \frac{1}{2}\int\left((\partial_i\sigma_{ik})^2 - (\partial_j\sigma_{ik})(\partial_i\sigma_{jk})\right)u^2, \tag{4.1.7}$$

where we now use our usual convention of summation over repeated indices. Note that in (4.1.7), $(\partial_i\sigma_{ik})^2$ therefore denotes $\sum_k(\sum_i \partial_i\sigma_{ik})^2$. Indeed, for each triple (i, j, k), we have, integrating by parts,

$$\int(\partial_j\sigma_{ik})\sigma_{jk}\partial_i\frac{u^2}{2} = -\int(\partial_i\sigma_{jk})(\partial_j\sigma_{ik})\frac{u^2}{2} - \int \sigma_{jk}\partial^2_{ij}\sigma_{ik}\frac{u^2}{2}, \tag{4.1.8}$$

and likewise

$$\int \sigma_{ik}(\partial_j\sigma_{jk})\partial_i\frac{u^2}{2} = -\int(\partial_i\sigma_{ik})(\partial_j\sigma_{jk})\frac{u^2}{2} - \int \sigma_{ik}\partial^2_{ij}\sigma_{jk}\frac{u^2}{2}. \tag{4.1.9}$$

Subtracting (4.1.9) from (4.1.8) and summing over i, j and k, yields

$$\int(\partial_j\sigma_{ik})\sigma_{jk}\partial_i\frac{u^2}{2} = \int \sigma_{ik}(\partial_j\sigma_{jk})\partial_i\frac{u^2}{2} - \int\left((\partial_i\sigma_{jk})(\partial_j\sigma_{ik}) - (\partial_i\sigma_{ik})^2\right)\frac{u^2}{2}, \tag{4.1.10}$$

with summation over repeated indices. Equation (4.1.10) is equivalent to (4.1.7). Inserting now (4.1.7) into (4.1.6) and the a priori estimate, we obtain

$$\frac{d}{dt}\int u^2 + \int|\sigma^t\nabla u|^2 = -2\int \sigma^t\nabla u\,.\,\mathrm{div}\,\sigma u - \frac{1}{2}\int(|\mathrm{div}\,\sigma|^2 - \mathrm{Tr}[(D\sigma)^2])u^2,$$

which may be written under the form

$$\frac{d}{dt}\int u^2 + \int|\mathrm{div}(\sigma u)|^2 = \frac{1}{2}\int(|\mathrm{div}\,\sigma|^2 + \mathrm{Tr}[(D\sigma)^2])u^2. \tag{4.1.11}$$

In (4.1.11), we recall that $\mathrm{div}\,\sigma$ is the vector field of components $\partial_i\sigma_{ik}$, that

$$|\mathrm{div}\,\sigma|^2 = (\partial_i\sigma_{ik})^2 = \sum_{k=1}^{d}\left|\sum_{i=1}^{d}\partial_i\sigma_{ik}\right|^2 \tag{4.1.12}$$

denotes its squared Euclidean norm, that

$$\mathrm{div}(\sigma u) = \sigma^t\,.\,\nabla u + (\mathrm{div}\,\sigma)u,$$

and that we *denote by*

$$\mathrm{Tr}[(D\sigma)^2] = (\partial_i\sigma_{jk})(\partial_j\sigma_{ik}) = \sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{k=1}^{d}(\partial_i\sigma_{jk})(\partial_j\sigma_{ik}). \tag{4.1.13}$$

Notice that, despite what our notation suggests, (4.1.13) is not necessarily nonnegative. The right-hand side of (4.1.11) is readily estimated using the Hölder inequality:

$$\int(|\mathrm{div}\,\sigma|^2 + \mathrm{Tr}[(D\sigma)^2])u^2 \le \|[|\mathrm{div}\,\sigma|^2 + \mathrm{Tr}[(D\sigma)^2]]_+\|_{L^\infty}\|u\|^2_{L^2}.$$

Therefore, assuming

$$\operatorname{div}\sigma \in L^\infty, \quad \operatorname{Tr}[(D\sigma)^2] \le C, \tag{4.1.14}$$

we obtain in (4.1.11) an estimate of the solution u in terms of $\|u\|^2_{L^\infty(L^2)} + \|\sigma^t\nabla u\|^2_{L^2(L^2)}$, of course assuming that the initial condition is L^2. Adding to this regularity the bound provided by the maximum principle when starting from a bounded initial condition, we (formally) obtain, for such a bounded initial condition, a solution in the space X defined in (3.1.3), that is,

$$X = \{u \in L^\infty([0,T], L^\infty) \cap C^0([0,T], L^p),\ 1 \le p < +\infty,\ \sigma^t\nabla u \in L^2([0,T], L^2)\}.$$

Regularization. We now turn to the regularization step. As above, we assume for simplicity that $\mathbf{b} = 0$ since we know well how to regularize the term $\mathbf{b}\nabla u$ using the standard assumptions on $\mathbf{b}$. The point is to regularize the additional term $\mathbf{b}^{\mathrm{mod}} - \mathbf{b}$, that is, the term

$$\partial_j(\sigma_{ik}\sigma_{jk})\partial_i u.$$

We observe that this term writes

$$\begin{aligned}
\partial_j(\sigma_{ik}\sigma_{jk})\partial_i u &= \partial_j\sigma_{ik}\,.\,\sigma_{jk}\,.\,\partial_i u + \sigma_{ik}\,.\,\partial_j\sigma_{jk}\,.\,\partial_i u \\
&= \partial_i(\sigma_{jk}\,.\,\partial_j\sigma_{ik}\,.\,u) - \partial_i(\sigma_{ik}\,.\,\partial_j\sigma_{jk}\,.\,u) \\
&\qquad - \big(\partial_j\sigma_{ik}\,.\,\partial_j\sigma_{ik} - (\partial_j\sigma_{jk})^2\big)u + 2\sigma_{ik}\,.\,\partial_j\sigma_{jk}\,.\,\partial_i u \\
&= \partial_i(\sigma_{jk}\,.\,\partial_j\sigma_{ik}\,.\,u) - \partial_i(\sigma_{ik}\,.\,(\operatorname{div}\sigma)_k\,.\,u) \\
&\qquad - \big(\operatorname{Tr}[(D\sigma)^2] - |\operatorname{div}\sigma|^2\big)u + 2(\operatorname{div}\sigma)_k\,.\,(\sigma^t\nabla u)_k.
\end{aligned}$$

The last three terms are easy to handle in the regularization procedure, because $\sigma \in H^1$, $\operatorname{div}\sigma \in L^\infty$, $u \in L^\infty$ and $\sigma^t\nabla u \in L^2$. For the second term, one only has to proceed with an integration by part, as we did for the term R_ε in the proof of Proposition 1 (and many times elsewhere already). Performing the regularization therefore amounts to regularizing the term

$$\partial_i(\sigma_{jk}\,.\,\partial_j\sigma_{ik}\,.\,u).$$

The difficulty is that this term is of the form $\operatorname{div}(\mathbf{c}u)$ for a vector field $\mathbf{c}$ with components $[\mathbf{c}]_i = \sigma_{jk}\,.\,\partial_j\sigma_{ik}$ on which we would like to avoid assuming weak differentiability. And we will indeed be able to do so because of the specific structure of that term and the fact that $\sigma^t\nabla u \in L^2$. In the classical setting, say when $\mathbf{c} \in W^{1,1}$ to fix the ideas, the procedure to regularize such a term is (see the proof of the commutation Lemma II.1 in [41]) to observe that

$$\begin{aligned}
\rho_\varepsilon \star \operatorname{div}(\mathbf{c}u) - \operatorname{div}(\mathbf{c}\rho_\varepsilon \star u) = &-\int (\mathbf{c}(x) - \mathbf{c}(y))u(y)\,.\,\nabla\rho_\varepsilon(x-y)\,dy \\
&- (\operatorname{div}\mathbf{c})(\rho_\varepsilon \star u).
\end{aligned} \tag{4.1.15}$$

One proves that

(a) as ε vanishes, the integral converges in L^1 to $(\operatorname{div}\mathbf{c})u$ when $\mathbf{c}$ and u are both smooth,
(b) this integral can be estimated in L^1 using $\|\mathbf{c}\|_{W^{1,1}}\|u\|_{L^\infty}$.

The result follows by density. We are going to follow the pattern of the classical proof here, except that we will essentially manipulate a weak form of the above argument (meaning, we integrate against test functions $\varphi(x)$) and that we will use the specific duality of the functional spaces that correspond to our setting. More precisely, we manipulate here solutions u that are in the space (3.1.3) and we will use test functions φ in the same space. We immediately remark that the rightmost term of (4.1.15) converges, as ε vanishes, to $(\operatorname{div}\mathbf{c})u$, which is then to be understood in the following sense. When u, σ and φ are smooth, we have

$$\begin{aligned}
\int \varphi\,.\,(\operatorname{div}\mathbf{c})\,.\,u &= \int \varphi\,.\,\partial_i(\sigma_{jk}\,.\,\partial_j\sigma_{ik})\,.\,u \\
&= \int \varphi\,.\,\partial_i\sigma_{jk}\,.\,\partial_j\sigma_{ik}\,.\,u + \int \varphi\,.\,\sigma_{jk}\,.\,\partial_{ij}\sigma_{ik}\,.\,u \\
&= \int \varphi\,.\,\partial_i\sigma_{jk}\,.\,\partial_j\sigma_{ik}\,.\,u - \int \varphi\,.\,\partial_j\sigma_{jk}\,.\,\partial_i\sigma_{ik}\,.\,u - \int \partial_i\sigma_{ik}\,.\,\sigma_{jk}\,.\,\partial_j(\varphi\,.\,u) \\
&= \int \varphi\,.\,(\operatorname{Tr}[(D\sigma)^2] - |\operatorname{div}\sigma|^2)u - \int (\operatorname{div}\sigma)_k\,.\,(u\,.\,(\sigma^t\nabla\varphi)_k + \varphi\,.\,(\sigma^t\nabla u)_k).
\end{aligned}$$

We notice that when both u and φ belong to the space (3.1.3) all terms still make sense, given our assumptions on σ. Consequently, performing the regularization exactly amounts to proving that the integral

$$\begin{aligned}
I_\varepsilon &:= \int \varphi(x)(\mathbf{c}(x) - \mathbf{c}(y))u(y)\,.\,\nabla\rho_\varepsilon(x-y)\,dx\,dy \\
&= \int \varphi(x)(\sigma_{jk}(x)\,.\,\partial_j\sigma_{ik}(x) - \sigma_{jk}(y)\,.\,\partial_j\sigma_{ik}(y))u(y)\,.\,\partial_i\rho_\varepsilon(x-y)\,dx\,dy \qquad (4.1.16)
\end{aligned}$$

can be estimated in terms of the norms of u and φ in the space defined by (3.1.3). The scheme described above of the proof of the standard setting, as detailed in [41], then proceeds easily. To start with, we split I_ε as follows:

$$\begin{aligned}
I_\varepsilon &= \int \varphi(x)\sigma_{jk}(x)\,.\,(\partial_j\sigma_{ik}(x) - \partial_j\sigma_{ik}(y))u(y)\,.\,\partial_i\rho_\varepsilon(x-y)\,dx\,dy \\
&\quad + \int \varphi(x)(\sigma_{jk}(x) - \sigma_{jk}(y))\,.\,\partial_j\sigma_{ik}(y)\,.\,u(y)\,.\,\partial_i\rho_\varepsilon(x-y)\,dx\,dy \\
&=: J_\varepsilon + K_\varepsilon. \qquad (4.1.17)
\end{aligned}$$

We temporarily leave K_ε aside (this term will actually cancel out with another term below) and focus on the term J_ε. We integrate by parts (notice that the integration needs to be performed cautiously, separately in the two variables x and y)

$$\begin{aligned}
J_\varepsilon &= \int \varphi(x)\sigma_{jk}(x)\,.\,(\partial_j\sigma_{ik}(x) - \partial_j\sigma_{ik}(y))u(y)\,.\,\partial_i\rho_\varepsilon(x-y)\,dx\,dy \\
&= -\int \partial_j(\varphi(x)\sigma_{jk}(x))(\sigma_{ik}(x) - \sigma_{ik}(y))u(y)\,.\,\partial_i\rho_\varepsilon(x-y)\,dx\,dy \\
&\quad - \int \varphi(x)\sigma_{jk}(x)\,.\,(\sigma_{ik}(x) - \sigma_{ik}(y))\partial_j u(y)\,.\,\partial_i\rho_\varepsilon(x-y)\,dx\,dy \\
&=: -L_\varepsilon - M_\varepsilon. \qquad (4.1.18)
\end{aligned}$$

We notice that

$$L_\varepsilon = \int \big((\sigma^t \nabla\varphi)_k(x) + (\operatorname{div}\sigma)_k \varphi(x)\big)\big(\sigma_{ik}(x) - \sigma_{ik}(y)\big)u(y)\,.\,\partial_i\rho_\varepsilon(x-y)\,dx\,dy$$

would appear when treating the commutator $\sigma_{jk}\partial_j([\rho_\varepsilon, \sigma_{ik}\partial_i](u))$ and can be estimated as

$$|L_\varepsilon| \le C\big(\|\sigma^t\nabla\varphi\|_{L^2} + \|\operatorname{div}\sigma\|_{L^\infty}\|\varphi\|_{L^2}\big)\|\sigma\|_{H^1}\|u\|_{L^\infty}. \tag{4.1.19}$$

We now estimate M_ε. To this end, we write M_ε as follows:

$$\begin{aligned} M_\varepsilon = &\int \varphi(x)\,.\,(\sigma_{ik}(x) - \sigma_{ik}(y))\,.\,(\sigma_{jk}(x) - \sigma_{jk}(y))\,.\,\partial_j u(y)\,.\,\partial_i\rho_\varepsilon(x-y)\,dx\,dy \\ &+ \int \varphi(x)\,.\,(\sigma_{ik}(x) - \sigma_{ik}(y))\,.\,\sigma_{jk}(y)\partial_j u(y)\,.\,\partial_i\rho_\varepsilon(x-y)\,dx\,dy \\ =:&\ N_\varepsilon + O_\varepsilon, \end{aligned} \tag{4.1.20}$$

where the second term O_ε, which would appear when developing the commutator $[\rho_\varepsilon, \sigma_{ik}\partial_i](\sigma_{jk}\partial_j u)$, can be estimated using

$$|O_\varepsilon| \le \|\varphi\|_{L^\infty}\|\sigma^t\nabla u\|_{L^2}\|\sigma\|_{H^1}. \tag{4.1.21}$$

To treat the term N_ε, we integrate by parts $\partial_i u(y)$ and find

$$\begin{aligned} N_\varepsilon = &\int \varphi(x)\,.\,(\sigma_{ik}(x) - \sigma_{ik}(y))\,.\,(\sigma_{jk}(x) - \sigma_{jk}(y))\,.\,u(y)\,.\,\partial_{ij}\rho_\varepsilon(x-y)\,dx\,dy \\ &+ \int \varphi(x)\,.\,(\sigma_{ik}(x) - \sigma_{ik}(y))\,.\,\partial_j\sigma_{jk}(y)\,.\,u(y)\,.\,\partial_i\rho_\varepsilon(x-y)\,dx\,dy \\ &+ \int \varphi(x)\,.\,(\sigma_{jk}(x) - \sigma_{jk}(y))\,.\,\partial_j\sigma_{ik}(y)\,.\,u(y)\,.\,\partial_i\rho_\varepsilon(x-y)\,dx\,dy \\ =:&\ P_\varepsilon + Q_\varepsilon + K_\varepsilon, \end{aligned} \tag{4.1.22}$$

where we have readily noticed that the last term is indeed K_ε defined in (4.1.17), and will therefore cancel out when collecting all integrals to form I_ε. The second term, namely Q_ε, involves the commutator $[\rho_\varepsilon, \sigma_{ik}\partial_i]((\operatorname{div}\sigma)_k u)$ and is easy to estimate

$$|Q_\varepsilon| \le \|\varphi\|_{L^\infty}\|\operatorname{div}\sigma\|_{L^\infty}\,\|u\|_{L^2}\|\sigma\|_{H^1}. \tag{4.1.23}$$

Our focus is now the one remaining term

$$P_\varepsilon = \int \varphi(x)\,.\,(\sigma_{ik}(x) - \sigma_{ik}(y))\,.\,(\sigma_{jk}(x) - \sigma_{jk}(y))\,.\,u(y)\,.\,\partial_{ij}\rho_\varepsilon(x-y)\,dx\,dy.$$

To show that this term is bounded, we observe that it contains a "double" cancellation, namely $(\sigma_{ik}(x) - \sigma_{ik}(y))\,.\,(\sigma_{jk}(x) - \sigma_{jk}(y))$ when x approaches y. We proceed similarly to the proof of [41, Lemma II.1]. We have

$$\begin{aligned} &\left|\int \varphi(x)\,.\,(\sigma_{ik}(x) - \sigma_{ik}(y))\,.\,(\sigma_{jk}(x) - \sigma_{jk}(y))\,.\,u(y)\,.\,\partial_{ij}\rho_\varepsilon(x-y)\,dx\,dy\right| \\ &\le \|\nabla^2\rho\|_{L^\infty}\|\varphi\|_{L^\infty}\|u\|_{L^\infty}\int dx\,\varepsilon^{-d}\int\limits_{|x-y|\le c\varepsilon} \frac{|\sigma_{ik}(x) - \sigma_{ik}(y)|}{\varepsilon}\,\frac{|\sigma_{jk}(x) - \sigma_{jk}(y)|}{\varepsilon}\,dy, \end{aligned} \tag{4.1.24}$$

using the Hölder inequality, and using that $\rho_\varepsilon(\frac{z}{\varepsilon}) = \varepsilon^{-d}\rho(\frac{z}{\varepsilon})$, where the smooth function ρ is fixed and has compact support (say in the ball of radius c). Next,

$$\begin{aligned}
&\int dx\, \varepsilon^{-d} \int\limits_{|x-y|\leq c\varepsilon} \frac{|\sigma_{ik}(x)-\sigma_{ik}(y)|}{\varepsilon}\frac{|\sigma_{jk}(x)-\sigma_{jk}(y)|}{\varepsilon}\,dy \\
&\leq \int dx \int\limits_{|z|\leq c}\Big(\int_0^1 |\nabla\sigma_{ik}(x+t\varepsilon z)|\,dt\Big)\Big(\int_0^1 |\nabla\sigma_{jk}(x+t\varepsilon z)|\,dt\Big)\,dz \\
&\leq \Big(\iint\limits_{|z|\leq c}\Big(\int_0^1 |\nabla\sigma_{ik}(x+t\varepsilon z)|\,dt\Big)^2 d\,x\,dz\Big)^{\frac12}\Big(\iint\limits_{|z|\leq c}\Big(\int_0^1 |\nabla\sigma_{jk}(x+t\varepsilon z)|\,dt\Big)^2 dx\,dz\Big)^{\frac12} \\
&\leq \Big(\iint\limits_{|z|\leq c}\int_0^1 |\nabla\sigma_{ik}(x+t\varepsilon z)|^2\,dt\,dx\,dz\Big)^{\frac12}\Big(\iint\limits_{|z|\leq c}\int_0^1 |\nabla\sigma_{jk}(x+t\varepsilon z)|^2\,dt\,dx\,dz\Big)^{\frac12}, \quad (4.1.25)
\end{aligned}$$

successively using the weak differentiability of σ and the Cauchy–Schwarz inequality. Collecting (4.1.24) and (4.1.25), we deduce that

$$|P_\varepsilon| \leq C\|\varphi\|_{L^\infty}\|u\|_{L^\infty}\|\sigma\|_{H^1}^2. \qquad (4.1.26)$$

Collecting (4.1.19), (4.1.21), (4.1.23) and (4.1.26), we obtain our estimate of I_ε:

$$\begin{aligned}
|I_\varepsilon| \leq{}& (\|\sigma^t\nabla\varphi\|_{L^2} + \|\operatorname{div}\sigma\|_{L^\infty}\|\varphi\|_{L^2})\|u\|_{L^\infty}\|\sigma\|_{H^1} + \|\varphi\|_{L^\infty}\,\|\sigma^t\nabla u\|_{L^2}\|\sigma\|_{H^1} \\
&+ \|\varphi\|_{L^\infty}\,\|\operatorname{div}\sigma\|_{L^\infty}\,\|u\|_{L^2}\|\sigma\|_{H^1} + \|\varphi\|_{L^\infty}\|u\|_{L^\infty}\|\sigma\|_{H^1}^2, \qquad (4.1.27)
\end{aligned}$$

up to irrelevant, universal multiplicative constants. We readily note that since all our arguments to establish (4.1.27) are based on Hölder-type inequalities, it is a straightforward modifications of the above manipulations to establish an estimate analogous to (4.1.27) using $\|\varphi\|_{L^p}$ instead of $\|\varphi\|_{L^\infty}$, $\|u\|_{L^p}$ instead of $\|u\|_{L^\infty}$ and $\|\sigma\|_{W^{1,r}}$ instead of $\|\sigma\|_{H^1}$, provided $\frac{1}{p}+\frac{1}{r}=\frac{1}{2}$. As announced above, it suffices then to argue by density (notice that we have previously established the density of smooth functions in X defined by (3.1.18)): for smooth functions, the commutator (4.1.15) vanishes with ε, and the bounds we have allows to approximate u, φ and σ in the appropriate functional spaces. In summary, we have therefore established, up to technical details we have omitted for simplicity, that

$$\Big|\int [\rho_\varepsilon, \sigma_{ik}\sigma_{jk}\partial^2_{ij}](u)\,\varphi\Big| \leq C\|u\|_{L^\infty\cap H}\,\|\varphi\|_{L^\infty\cap H}$$

for some constant C, which actually typically depends on $1 + \|\operatorname{div}\sigma\|_{L^\infty}\|\sigma\|_{H^1} + \|\sigma\|_{H^1}^2$ and where the space H is defined in (3.1.18). This also writes, in a more abstract form

$$\|[\rho_\varepsilon, \sigma_{ik}\sigma_{jk}\partial^2_{ij}](u)\|_{L^1+H'} \leq C\|u\|_{L^\infty\cap H}.$$

It is now easy, keeping track of the time variable everywhere in the above arguments and in particular in estimate (4.1.27), to reinstate the time variable when the functions u and φ depend on time (and possibly also $\mathbf{b}$ and σ, then assuming that $\mathbf{b} \in L^1([0,T], W^{1,1})$, $\sigma \in L^2([0,T], H^1)$, $\operatorname{div}\sigma \in L^1([0,T], L^\infty)$). It follows that we have proven that, if $u \in X$ (X being defined in (3.1.3)) solves

$$\partial_t u - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial^2_{ij}u = 0,$$

then $u_\varepsilon = \rho_\varepsilon \star u$ solves

$$\partial_t u_\varepsilon - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial^2_{ij}u_\varepsilon = R_\varepsilon, \tag{4.1.28}$$

where the remainder R_ε satisfies

$$R_\varepsilon \xrightarrow{\varepsilon\to 0} 0 \quad \text{in } L^1([0,T], L^1) + L^2([0,T], H). \tag{4.1.29}$$

The above arguments (together with a now classical treatment of the term $\mathbf{b}\,.\,\nabla u$) outline the proof of the following Proposition 4, which summarizes the results of this section.

Proposition 4. *Assume* (3.1.4), (3.1.5), (4.1.14)*, that is,*

$$\begin{cases} \mathbf{b} \in W^{1,1}, & [\operatorname{div}\mathbf{b}]_- \in L^\infty, \\ \sigma \in H^1, & \\ \operatorname{div}\sigma \in L^\infty, & [\operatorname{Tr}[(D\sigma)^2]]_+ \in L^\infty, \end{cases} \tag{4.1.30}$$

where we recall the notation (4.1.12)–(4.1.13)*. Consider then equation* (1) *with* $\mathbf{a} = \frac{1}{2}\sigma\sigma^t$*. Then, for all initial conditions* $u_0 \in L^\infty$*,* (1) *has a unique solution in the space* X*, that is,*

$$\{u \in L^\infty([0,T], L^\infty) \cap C^0([0,T], L^p),\ 1 \le p < +\infty,\ \sigma^t\nabla u \in L^2([0,T], L^2)\}.$$

Remark 42. The case of an initial condition $u_0 \in L^p$, for $1 < p < +\infty$, instead of L^∞ can be treated using renormalization. The proof follows the pattern of that in Section 2.3. We do not proceed in this direction.

4.1.3 Probabilistic interpretation of our assumptions

We now comment upon our couple of assumptions ($\operatorname{div}\sigma \in L^\infty$, $[\operatorname{Tr}[(D\sigma)^2]]_+ \in L^\infty$) in (4.1.30) from the perspective of probability theory. We again argue for simplicity of exposition in the case when $\mathbf{b} \equiv 0$, but our argument is general. We consider

$$d\mathbf{X}_t = \sigma(\mathbf{X}_t)\,.\,d\mathbf{W}_t. \tag{4.1.31}$$

We would like to understand what it means to control the Jacobian $J_t = |\det \frac{\partial \mathbf{X}_t(x)}{\partial x}|$ (and actually its inverse) because we expect that, as in the deterministic setting, this control is related to the preservation of norms, thus the issue of well-posedness, in the associated partial differential equation. To this end, we assume all the regularity necessary for our manipulations to make sense and now calculate the evolution of J_t.

We could use Itô differential calculus to deduce dJ_t from $d\mathbf{X}_t$ in (4.1.31), but it is actually *technically* simpler to first write (4.1.31) in the Stratonovich form

$$d\mathbf{X}_t = \mathbf{c}(\mathbf{X}_t)\,dt + \sigma(\mathbf{X}_t) \circ d\mathbf{W}_t, \tag{4.1.32}$$

where we denote by $[\mathbf{c}]_i = -\frac{1}{2}\sigma_{jk}\partial_j\sigma_{ik}$, and use Stratonovich differential calculus, which, as is well known, proceeds exactly alike classical deterministic differential calculus. We deduce from (4.1.32) that

$$dJ_t = \operatorname{div}\mathbf{c}(\mathbf{X}_t)J_t\,dt + (\operatorname{div}\sigma(\mathbf{X}_t))J_t \circ d\mathbf{W}_t, \tag{4.1.33}$$

where

$$\operatorname{div}\mathbf{c} = -\frac{1}{2}\partial_i(\sigma_{jk}\partial_j\sigma_{ik}) = -\frac{1}{2}(\operatorname{Tr}[(D\sigma)^2] + \sigma_{jk}\partial_{ij}\sigma_{ik}). \tag{4.1.34}$$

From (4.1.32)–(4.1.33), that is,

$$d\begin{pmatrix}\mathbf{X}_t\\ J_t\end{pmatrix} = \begin{pmatrix}\mathbf{c}(\mathbf{X}_t)\\ \operatorname{div}\mathbf{c}(\mathbf{X}_t)J_t\end{pmatrix}dt + \begin{pmatrix}\sigma(\mathbf{X}_t)\\ (\operatorname{div}\sigma(\mathbf{X}_t))J_t\end{pmatrix} \circ d\mathbf{W}_t,$$

we now infer to the Itô form of the equation on J_t:

$$\begin{aligned} dJ_t = \Big(&\operatorname{div}\mathbf{c}(\mathbf{X}_t)J_t + \frac{1}{2}(\operatorname{div}\sigma(\mathbf{X}_t))J_t \,.\, \partial_J((\operatorname{div}\sigma(\mathbf{X}_t))J_t) \\ &+ \frac{1}{2}\operatorname{Tr}[\sigma^t(\mathbf{X}_t)\,.\,\partial_X((\operatorname{div}\sigma(\mathbf{X}_t))J_t)]\Big)\,dt + (\operatorname{div}\sigma(\mathbf{X}_t))J_t\,.\,d\mathbf{W}_t, \end{aligned} \tag{4.1.35}$$

where

$$\frac{1}{2}(\operatorname{div}\sigma(\mathbf{X}_t))J_t\,.\,\partial_J((\operatorname{div}\sigma(\mathbf{X}_t))J_t) = \frac{1}{2}|\operatorname{div}\sigma|^2 J_t \tag{4.1.36}$$

and

$$\begin{aligned} \frac{1}{2}\operatorname{Tr}[\sigma^t(\mathbf{X}_t)\,.\,\partial_X((\operatorname{div}\sigma(\mathbf{X}_t))J_t)] &= \frac{1}{2}\sigma_{jk}\partial_{X_j}((\partial_i\sigma_{ik})J_t) \\ &= \frac{1}{2}\sigma_{jk}\partial_{ij}\sigma_{ik}J_t. \end{aligned} \tag{4.1.37}$$

Collecting (4.1.34) through (4.1.37), we obtain

$$dJ_t = \frac{1}{2}(|\operatorname{div}\sigma|^2 - \operatorname{Tr}[(D\sigma)^2])J_t\,dt + (\operatorname{div}\sigma)J_t\,.\,d\mathbf{W}_t.$$

We next obtain the evolution of J_t^{-1} by Itô calculus:

$$d(J_t^{-1}) = \frac{1}{2}(|\operatorname{div}\sigma|^2 + \operatorname{Tr}[(D\sigma)^2])J_t^{-1}\,dt - (\operatorname{div}\sigma)J_t^{-1}\,.\,d\mathbf{W}_t. \tag{4.1.38}$$

We recognize in this equation the term $|\operatorname{div}\sigma|^2 + \operatorname{Tr}[(D\sigma)^2]$ which also arises in our theory for the associated parabolic equation, on the right-hand side of (4.1.11). We notice that, indeed, no second derivative of σ appears in this equation. Taking the expectation of (4.1.38), we have

$$d\mathbb{E}_x(J_t^{-1}) = \mathbb{E}_x\Big(\frac{1}{2}(|\operatorname{div}\sigma|^2 + \operatorname{Tr}[(D\sigma)^2])J_t^{-1}\Big)\,dt \tag{4.1.39}$$

and our bounds therefore allow to estimate the evolution of J_t^{-1} on average. As in our calculations (1.1.5) in the deterministic setting for *positive* times, J_t^{-1} indeed somehow controls the preservation of norms. Actually, that claim is not *exactly* true. And the discussion in the present paragraph must therefore be understood as at most an analogy, which confirms the instrumental role played by the quantities $|\operatorname{div}\sigma|^2$ and $\operatorname{Tr}[(D\sigma)^2]$. Indeed, we certainly have, using (1.2.3),

$$\int |u(t,x)|\,dx = \int \mathbb{E}_x(|u_0(\mathbf{X}_t(x,\omega))|\,dx = \iint |u_0(\mathbf{X}_t(x,\omega))|\,dx\,d\mathbb{P}\omega,$$

Formally using the Fubini Theorem, and next changing x into $y = \mathbf{X}_t(x,\omega)$ for each ω fixed, this yields

$$\int |u(t,x)|\,dx = \int |u_0(y)|\mathbb{E}\left(\left|\det \frac{\partial \mathbf{X}_t(x,\omega)}{\partial x}\bigg|_{x=\mathbf{X}_t^{-1}(y,\omega)}\right|^{-1}\right) dy. \tag{4.1.40}$$

The rightmost quantity does however not involve $\mathbb{E}\,(J_t^{-1})$ (in the notation we have used throughout this paragraph), since the Jacobian is *evaluated* at x such that $\mathbf{X}_t(x,\omega) = y$. The control we have in (4.1.38) and (4.1.39) does not provide any information on the integrand in (4.1.40). We thus speak of an *analogy* and not of a mathematical *argument*.

4.1.4 Back to the L^1 estimate

As announced in Section 3.4, we devote this short section to the L^1 estimate on the solution, and the corresponding L^1 estimate on $\sigma^t\nabla u$ as explained there, in the case of an operator *not* in divergence form.

Very much in line with the manipulations we just made in Section 4.1.2 above, we write

$$-\frac{1}{2}\sigma_{ik}\partial_i(\sigma_{jk}\partial_j u) = -\frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j u) + \frac{1}{2}(\partial_i\sigma_{ik})\sigma_{jk}\partial_j u.$$

Denoting by $\Theta = \frac{1}{2}\operatorname{div}\sigma$, we observe that the rightmost term of the above expression reads as $\sigma\Theta\,.\,\nabla u$ and therefore that, assuming

$$\operatorname{div}\sigma \in L^\infty \tag{4.1.41}$$

allows to include this term in our discussion of Section 3.4 regarding the Girsanov-type transform. Alternate assumptions on this term $\operatorname{div}\sigma$ can possibly be imposed, depending on the specific context.

Likewise, we may write

$$\begin{aligned}-\frac{1}{2}\sigma_{ik}\sigma_{jk}\partial_{ij}^2 u = &-\frac{1}{2}\partial_i(\sigma_{ik}\sigma_{jk}\partial_j u) + (\partial_i\sigma_{ik})\sigma_{jk}\partial_j u\\ &+\frac{1}{2}(\sigma_{ik}\partial_i\sigma_{jk} - \sigma_{jk}\partial_i\sigma_{ik})\partial_j u.\end{aligned} \tag{4.1.42}$$

The second term is the term in $\operatorname{div}\sigma$ we already saw above. We notice that the new rightmost term is of the form $\widetilde{\mathbf{b}}\,.\,\nabla u$ with

$$\operatorname{div}\widetilde{\mathbf{b}} = -\frac{1}{2}|\operatorname{div}\sigma|^2 + \frac{1}{2}\operatorname{Tr}[(D\sigma)^2].$$

The assumption

$$[\mathrm{Tr}[(D\sigma)^2]]_+ \in L^\infty, \tag{4.1.43}$$

in addition to (4.1.41), therefore allows to also address this term as a “regular” transport term that has controlled divergence.

We note that our assumptions (4.1.41)–(4.1.43) in this section are of course reminiscent of (4.1.30). They allow to apply to operators that are not in divergence form the considerations of Section 3.4. And, as emphasized in Section 3.4, they do not involve second derivatives of σ.

4.2 Application to the theory of stochastic differential equations

We have already seen in Section 2.2 the consequences on the theory of stochastic differential equations of our results obtained in the partial differential equation setting when the matrix σ is constant. We therefore only discuss here the necessary adaptations when σ varies, and satisfies the assumption of our Proposition 1, namely $\sigma \in H^1$. Of course, we take $\mathbf{b}$ satisfying (3.1.4), but this is not our focus here. In its principle, the construction of solutions to the stochastic differential equation

$$d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)\,dt + \sigma(\mathbf{X}_t)\,d\mathbf{W}_t, \tag{4.2.1}$$

under the above assumptions on $(\mathbf{b}, \sigma)$ follows the same pattern as that performed in Section 2.2 for the solutions to (2.2.1) in the case of a constant σ. The properties of uniqueness, both in law and pathwise, follow similarly. Some additional arguments are however necessary for our mathematical proofs, and we sketch those arguments in this section. Using the result of the partial differential equation setting (here Proposition 4 instead of Theorem 1), we are going to show that, up to the elimination of a set of initial conditions of zero Lebesgue measure, equation (4.2.1) is well-posed. The proofs show that the limits are independent of the specific regularization procedure employed and we will obtain a characterization of the solution constructed, via the Feynman–Kac formula.

This is precisely the point we now focus on. In Section 2.2, assumption (2.2.3), namely

$$\int \mathbb{E}_x(\mathbb{1}_A(\mathbf{X}_t))\,dx \le c|A|$$

for all Borel sets A, is precisely useful to eliminate remainders in the approximation of the Itô formula. Now that σ is not constant, we need to complement Definition 1, and we do this in the following manner. We assume that

$$\left\| \mathbb{E}_x \int_0^t \varphi(t-s, \mathbf{X}_s)\,ds \right\|_{L^1_{t,x}} \le c\|\varphi\|_{L^1([0,T],L^1)+L^2([0,T],H)} \tag{4.2.2}$$

for all $\varphi \in L^1([0,T], L^1) + L^2([0,T], H)$, where we recall that H is defined in (3.1.18). Note the latter space has dual space $L^\infty([0,T], L^\infty) \cap L^2([0,T], H)$. This additional assumption (4.2.2) complements (2.2.3). It is satisfied by regular solutions and, therefore, will be satisfied by the solution we construct below by regularization.

Definition 3. We say that $\mathbf{X}_t(x)$ (abbreviated as $\mathbf{X}_t$ when there is no ambiguity) is a (family of) solution(s) to (2.2.1), parameterized by the initial condition x, or a solution flow to (4.2.1), that is,

$$d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)\,dt + \sigma(\mathbf{X}_t)\,d\mathbf{W}_t,$$

when

- $\mathbf{X}_t$ satisfies (2.2.3), that is, there exists a constant c such that
$$\int \mathbb{E}_x(\mathbb{1}_A(\mathbf{X}_t))\,dx \le c|A|$$
for all Borel sets A,
- $\mathbf{X}_t$ satisfies
$$\mathbf{X}_t = x + \int_0^t \mathbf{b}(\mathbf{X}_s)\,ds + \int_0^t \sigma(\mathbf{X}_s)\,d\mathbf{W}_s \tag{4.2.3}$$
for all times $t \ge 0$, almost surely and for almost all initial condition x,
- $\mathbf{X}_t$ satisfies (4.2.2) that is, there exists a constant c such that
$$\left\| \mathbb{E}_x \int_0^t \varphi(t-s, \mathbf{X}_s)\,ds \right\|_{L^1_{t,x}} \le c\|\varphi\|_{L^1([0,T],L^1)+L^2([0,T],H)}$$
for all $\varphi \in L^1([0,T], L^1) + L^2([0,T], H)$,
- $\mathbf{X}_t$ satisfies the semi-group property
$$\mathbf{X}_{t+s} = \mathbf{X}_t(\mathbf{X}_s)$$
for all times $s, t \ge 0$ (where our notation assumes that the Brownian taken for $\mathbf{X}_t$ is that taken for $\mathbf{X}_{t+s}$ from time s).

Remark 43. As in Remark 5, we notice that the first two properties in Definition 3 imply the continuity of trajectories stated in (2.2.14). Our argument on page 54 anticipated the present case where σ varies.

Remark 44. In the particular case when σ is constant, condition (4.2.2) does not exactly agree with assumption (2.2.3) of Definition 1. Condition (4.2.2) is "weaker" in the sense that it is integrated in time. This feature was already sufficient for our argument of uniqueness of the case σ constant, as we observed in Remark 12, but since in that case we were indeed able to prove the *existence* of a solution satisfying the stronger condition (2.2.3), our assumption there is consistent and therefore the right one. Condition (4.2.2) is also "weaker" than (2.2.3) because, on bounded domains, $L^2([0,T], H)$ is a subspace of $L^1_{t,x}$ and thus (4.2.2) is then implied by (2.2.3). In full generality, the conditions are slightly different and not really comparable to one another. We focus here on the general case of a varying σ that in addition needs not be positive definite. Our arguments are thus intrinsically different from those of the case σ constant (and positive).

We now get to the study of equation (4.2.1). Our argument of Section 2.2 successively dealt with uniqueness in law, existence in law, pathwise uniqueness, and strong existence, based on similar considerations in nature. We know from our study there of the case of a constant σ that the key step is to be able to regularize the equation, perform Itô calculus and establish a Feynman–Kac formula for our solution. For brevity, we therefore only mention here the modifications of our previous arguments of Section 2.2 that are necessary to proceed with this key step. Likewise, we do not make precise the main results, which are the necessary adaptations of those stated in Corollaries 2, 3, 5, 6 to the present setting. They of course are based on the results obtained on the partial differential equation in Chapter 4 instead of Theorem 1, and on Definition 3 of solution to the stochastic differential equation, instead of Definition 1.

We start with *uniqueness in law*. We observe that, in the solution process we consider

$$X_t - x = \int_0^t \mathbf{b}(\mathbf{X}_s)\,ds + \int_0^t \sigma(\mathbf{X}_s)\,d\mathbf{W}_s,$$

the rightmost integral is a stochastic integral. Indeed, because of the L^2 integrability of σ and property (2.2.3) of $\mathbf{X}_t$, we have

$$\int\int_0^T \mathbb{E}_x|\sigma(\mathbf{X}_t)|^2\,dt\,dx \le CT\|\sigma\|_{L^2}^2 \tag{4.2.4}$$

thus, for almost all x, $\sigma(\mathbf{X}_t) \in L^2(\Omega\times[0,T])$. It follows that the proof of the classical Itô inequality therefore applies and we have, for ε fixed, and a regularization u_ε of u, unique solution we have constructed for the parabolic equation with initial condition $u_0 \in L^\infty$:

$$\begin{aligned}
-u_\varepsilon(t,x) + u_\varepsilon(0,\mathbf{X}_t) &= \int_0^t \frac{d}{ds}\,u_\varepsilon(t-s,\mathbf{X}_s)\,ds \\
&= -\int_0^t \frac{\partial u_\varepsilon}{\partial t}(t-s,\mathbf{X}_s)\,ds + \int_0^t \mathbf{b}(\mathbf{X}_s)\,.\,\nabla u_\varepsilon(t-s,\mathbf{X}_s)\,ds \\
&\qquad + \frac{1}{2}\int_0^t \sigma\sigma^t(\mathbf{X}_s)D^2u_\varepsilon(t-s,\mathbf{X}_s)\,ds \\
&\qquad + \int_0^t \nabla u_\varepsilon(t-s,\mathbf{X}_s)\,.\,\sigma(\mathbf{X}_s)\,d\mathbf{W}_s, \\
&= -\int_0^t R_\varepsilon(t-s,\mathbf{X}_s)\,ds + \int_0^t \sigma^t(\mathbf{X}_s)\nabla u_\varepsilon(t-s,\mathbf{X}_s)\,.\,d\mathbf{W}_s,
\end{aligned} \tag{4.2.5}$$

where we notice that, because $\sigma^t\nabla u \in L^2$ and for the same reason as in (4.2.4), the rightmost integral is a stochastic integral, and where we have denoted by R_ε the remainder term for the regularization procedure. The regularization of the parabolic equation we have performed in the previous section (see equation (4.1.28)) indeed shows that

$$\frac{\partial u_\varepsilon}{\partial t} - \mathbf{b}\,.\,\nabla u_\varepsilon - \frac{1}{2}\sigma\sigma^t D^2 u_\varepsilon = R_\varepsilon$$

for a remainder term that satisfies (see equation (4.1.29))

$$R_\varepsilon \xrightarrow{\varepsilon\to 0} 0 \quad \text{in } (L^\infty([0,T],L^\infty) \cap L^2([0,T],H))'.$$

We next take the expectation at fixed x of both sides of (4.2.5) and obtain

$$u_\varepsilon(t,x) - \mathbb{E}_x(u_\varepsilon(0,\mathbf{X}_t)) = \mathbb{E}_x \int_0^t R_\varepsilon(t-s,\mathbf{X}_s)\,ds.$$

Using (4.2.2) and letting ε vanish, we deduce that

$$u(t,x) = \mathbb{E}_x(u_0(\mathbf{X}_t)),$$

which yields uniqueness, and, in fact again, the characterization of the solution process. Consequently, the analogous result to that stated in Corollary 2 holds true in the present setting.

For *existence in law*, our argument of Section 2.2 actually immediately applies to the present case of a varying coefficient σ, because we have already collected there all the material pertaining to that case. The key ingredients, in turn used for the tightness of the sequence of probabilities solutions to the regularized problem, were estimations (2.2.6) and (2.2.12). The latter was precisely established for the present context of a varying σ. Both estimates are based upon (2.2.3). The fact that this estimation is *uniform in the regularization parameter* is itself a consequence of the L^1 estimate on the parabolic equation. So *we assume* that $\mathbf{b}$ and σ are *such that* the L^1 estimate (2.1.24) holds true for the solution to the parabolic equation. So is then estimate (2.2.3) and, likewise, estimate (4.2.2), for the solutions we construct and consider. Then it is easy to proceed with proving existence of a solution in law, using the formalism we introduced and the ingredients we made clear in Section 2.2. Note that in the absence of an L^1 estimate of the type (2.1.24), one could imagine to proceed with "only" the L^p estimate (2.1.3) which holds true more generally (see the related Remarks 31 and 8, which also apply to Definition 3).

To proceed with *pathwise uniqueness* (and *strong existence* falls using the same techniques), we introduce as in Sections 1.2 and 2.2 the parabolic equation in the space of doubled dimension

$$\begin{aligned}\frac{\partial u}{\partial t} - b_i(x)\frac{\partial u}{\partial x_i} - b_i(y)\frac{\partial u}{\partial y_i} - \frac{1}{2}\sigma_{ik}(x)\sigma_{jk}(x)\frac{\partial^2 u}{\partial x_i \partial x_j}& \\ - \sigma_{ik}(x)\sigma_{jk}(y)\frac{\partial^2 u}{\partial x_i \partial y_j} - \frac{1}{2}\sigma_{ik}(y)\sigma_{jk}(y)\frac{\partial^2 u}{\partial y_i \partial y_j} &= 0. \qquad (4.2.6)\end{aligned}$$

It may be written in the more compact form

$$\partial_t u - \mathbf{B}(x,y)\,.\,\nabla_{x,y} u - \frac{1}{2}\mathbf{\Sigma}\mathbf{\Sigma}^t(x,y) D^2_{x,y} u = 0, \tag{4.2.7}$$

where $\mathbf{B}$ and $\mathbf{\Sigma}$ have been defined in (1.2.10). In order to apply our existence and uniqueness result of Proposition 4 to equation (4.2.7), we need to check that assumptions (4.1.30) hold true.

The first three assumptions, namely $\mathbf{B} \in W^{1,1}_{x,y}$, $[\operatorname{div}_{x,y} \mathbf{B}]_- \in L^\infty_{x,y}$, $\mathbf{\Sigma} \in H^1_{x,y}$, are evidently equivalent, given the specific form of $\mathbf{B}$ and $\mathbf{\Sigma}$, to the standard assumptions $\mathbf{b} \in W^{1,1}_x$, $[\operatorname{div}_x \mathbf{b}]_- \in L^\infty_x$, $\sigma \in H^1_x$. Similarly, we have

$$\operatorname{div}_{x,y} \mathbf{\Sigma} = \operatorname{div}_x \sigma(x) + \operatorname{div}_y \sigma(y),$$

and thus $\operatorname{div}_{x,y} \mathbf{\Sigma} \in L^\infty_{x,y}$ amounts to $\operatorname{div} \sigma \in L^\infty_x$. The argument is again similar for the last assumption of (4.1.30), namely $\operatorname{Tr}_{x,y}[(D_{x,y}\mathbf{\Sigma})^2] \le C$ which is equivalent to $\operatorname{Tr}[(D\sigma)^2] \le C$. Under this set of assumptions, that is, $\mathbf{b} \in W^{1,1}_x$, $[\operatorname{div}_x \mathbf{b}]_- \in L^\infty_x$, $\sigma \in H^1_x$, $\operatorname{div} \sigma \in L^\infty_x$, $\operatorname{Tr}[(D\sigma)^2] \le C$, pathwise uniqueness follows, thereby extending the result of Corollary 5 of Section 2.2.3 to the present setting of a varying coefficient σ. Finally, under the same set of assumptions, the existence of a strong solution is proved as in Section 2.2.3, and the analogue to Corollary 6 holds true.

We conclude the section by emphasizing that, clearly, there is room for improvement in our understanding of this case when σ is not constant, notably for what regards the most general assumptions to make.

4.3 Further extensions

We mention in this final section two possible research directions that follow up on the results of these notes. For both directions, we only briefly outline the issues and suggest possible arguments to solve those issues.

More regular initial conditions. For simplicity of exposition, we consider the case $\mathbf{b} = 0$. It is straightforward, if necessary, to reinstate $\mathbf{b}$ and treat the corresponding term in the following argument (using ingredients developed in [64] and which we briefly recall below).

A natural question, in particular in echo to the classical (Hölder) theory of parabolic equations for which such a result holds true, is to ask whether a better regularity of the initial condition allows to prove a better regularity of the solution. A prototypical question is, does $u_0 \in W^{1,p}$ imply $u \in C([0,T], W^{1,p})$? We now give some indications to establish that this is indeed the case. First, at least formally, the question can be stated and studied within the probability framework. Consider the couple $(\mathbf{X}_t, \frac{\partial \mathbf{X}_t}{\partial x})$ solution to the following system of stochastic differential equations:

$$\begin{cases} d\mathbf{X}_t = \sigma(\mathbf{X}_t)\, d\mathbf{W}_t, \\ d\Big(\dfrac{\partial \mathbf{X}_t}{\partial x}\Big) = \nabla\sigma(\mathbf{X}_t)\dfrac{\partial \mathbf{X}_t}{\partial x}\, d\mathbf{W}_t. \end{cases} \tag{4.3.1}$$

The solution to the parabolic equation, and its derivative in space, can be expressed in terms of this couple as follows:

$$\begin{cases} u(t,x) = \mathbb{E}\left(u_0(\mathbf{X}_t)\right), \\ \nabla u(t,x) = \mathbb{E}\Big(\nabla u_0(\mathbf{X}_t)\dfrac{\partial \mathbf{X}_t}{\partial x}\Big). \end{cases} \tag{4.3.2}$$

This is of course true in the regular setting at least but the arguments performed throughout these notes indicate that this will also hold for more general settings. It follows from (4.3.2) that the properties of ∇u can be directly inferred from the understanding of system (4.3.1).

On the other hand, a direct approach on the parabolic equation is also possible. It is an observation by Krylov that, differentiating in space the equation

$$\partial_t u - a_{ij}\partial^2_{ij} u = 0, \tag{4.3.3}$$

one obtains

$$\partial_t u_{,k} - a_{ij}\partial^2_{ij} u_{,k} - \partial_k(a_{ij})\partial^2_{ij} u = 0, \tag{4.3.4}$$

where $u_{,k}$ evidently denotes $\frac{\partial u}{\partial x_k}$. Since equation (4.3.4) is not closed in $u_{,k}$ (it indeed involves u itself), the idea is to introduce the function

$$w(t,x,\xi) = \frac{\partial u(t,x)}{\partial x_k}\,.\,\xi_k,$$

with, as usual, the convention of summation over repeated indices, and write equation (4.3.4) in the form

$$\frac{\partial w}{\partial t} - a_{ij}\frac{\partial^2 w}{\partial x_j \partial x_j} - \frac{\partial a_{ij}(x)}{\partial x_k}\xi_k \frac{\partial^2 w}{\partial x_i \partial \xi_j} = 0. \tag{4.3.5}$$

That equation is now a parabolic-type equation in w, in a higher-dimensional space, and its study will provide information on the differentiability of the solution u to (4.3.3).

Equation (4.3.5), along with actually its probability theoretic companion equation (4.3.1), are both amenable to the techniques developed in these notes. One useful guideline for adapting our proofs to that setting is to consider our previous work [64] that establishes similar additional regularity of solutions when initial conditions are more regular. We have considered in that work both the system

$$\begin{cases} \dot{Y}(t) = \mathbf{b}(Y(t)), \\ \dot{R}(t) = \nabla_y \mathbf{b}(Y(t))R(t), \\ Y(0) = y, \quad R(0) = r, \end{cases} \tag{4.3.6}$$

and the specific transport equation

$$\frac{\partial u}{\partial t} + \mathbf{b}_1(x_1)\,.\,\nabla_{x_1} u + \mathbf{b}_2(x_1,x_2)\,.\,\nabla_{x_2} u = 0. \tag{4.3.7}$$

System (4.3.6), where (Y, R) stand for $(X(t,x), \frac{\partial X(t,x)}{\partial x})$, is clearly the deterministic analogous system to system (4.3.1). It is of course obtained by differentiating the equation in the first line with respect to the initial condition x. On the other hand, the structure of the transport equation (4.3.7) (actually, of course, the transport equation associated to the differential system (4.3.6)) is similar in nature to that of (4.3.5), up to the formal transformation of first-order terms into second-order terms. The point in (4.3.7) is that the coefficient field $b_1(x_1)$ along ∇_1 *only* depends on the variable x_1. Somewhat similarly in (4.3.5), the term $\frac{\partial a_{ij}(x)}{\partial x_k}\xi_k \frac{\partial^2 w}{\partial x_i \partial \xi_j}$, although involving a first derivative of a_{ij} and thus expectedly difficult to address, is a smooth function of ξ, since it is linear in that variable. Our technique of [64], consisting in regularizing *separately* in the two variables, and which allows to prove well-posedness of (4.3.7) (and, subsequently of (4.3.6)) under the assumptions

$$\begin{cases} b_1 = b_1(x_1) \in W^{1,1}_{x_1}, & \operatorname{div}_{x_1} b_1 = 0, \\ b_2 = b_2(x_1, x_2) \in L^1_{x_1}(W^{1,1}_{x_2}), & \operatorname{div}_{x_2} b_2 = 0 \end{cases}$$

is likely to apply to the case of (4.3.5) and (4.3.1).

Nonlinear equations. We mention in this final paragraph a possible pathway for extending the results of these notes to the case of nonlinear scalar conservation laws

$$\partial_t u + \operatorname{div}_x(\mathbf{F}(x, u(x))) = 0, \tag{4.3.8}$$

where $\mathbf{F}(x,u)$ denotes a nonlinear flux. To relate to the contents of these notes, think of, say, $\mathbf{F}(x,u) = f(u)\,\mathbf{b}(x)$, where f is a nonlinear function, or the even simpler case $\mathbf{F}(x,u) = f(u)\,\mathbf{b}$, for a fixed vector $\mathbf{b}$, which allows us to focus on the difficulty related to the nonlinearity in the solution u. In that case, we briefly recall that, in the classical setting, the typical proof of uniqueness of entropic solutions of equation (4.3.8) follows the method introduced by Kruzkov. That method can be formally summarized as follows. We consider solutions that additionally satisfy the entropy condition: for all $k \in \mathbb{R}$ and all smooth, compactly supported, nonnegative functions φ,

$$\iint \left(|u(t,x) - k|\partial_t \varphi + \operatorname{sgn}(u(t,x) - k)(f(u(t,x)) - f(k))\mathbf{b}\,.\nabla\varphi\right) dx\,dt \geq 0. \tag{4.3.9}$$

Condition (4.3.9) is actually equivalent to the alternate usual formulation of the entropy condition, namely for all couples (η, ϕ) of (entropy, flux) (meaning that η is convex and $\partial\phi(u) = \eta'(u)\partial(\mathbf{b}f(u))$),

$$\partial_t \eta(u) + \operatorname{div}_x \phi(u) \leq 0$$

in the sense of distributions. Suppose now we have two candidates u and v solutions, both entropic in the sense of (4.3.9), and associated to the same initial condition. We take $k = v(\tau, y)$ in condition (4.3.9) stated for u, and, symmetrically, $k = u(t,x)$ in the condition stated for v. We obtain

$$\begin{aligned} \iiiint & \big(|u(t,x) - v(\tau,y)|(\partial_t\varphi + \partial_\tau\varphi) \\ & + \operatorname{sgn}(u(t,x) - v(\tau,y))(f(u(t,x)) - f(v(\tau,y)))(\mathbf{b}\,.\nabla_x\varphi + \mathbf{b}\,.\nabla_y\varphi)\big)\,dx\,dt\,dy\,d\tau \geq 0. \end{aligned}$$

A suitable choice of test function φ then allows to show that

$$\iint \big(|u(t,x) - v(t,x)|\partial_t \psi \\ + \operatorname{sgn}(u(t,x) - v(t,x))(f(u(t,x)) - f(v(t,x)))\mathbf{b}\,.\,\nabla\psi\big)\,dx\,dt \geq 0 \tag{4.3.10}$$

for all smooth, compactly supported, nonnegative functions ψ. It is then deduced that, in some formal notation (but the statement is indeed established and made precise upon using a sequence of test functions ψ in (4.3.10)), for $t_1 \leq t_2$,

$$\iint |u(t,x) - v(t,x)|(\delta(t - t_1) - \delta(t - t_2))\,dx\,dt \geq 0.$$

This immediately writes

$$\int |u(t_1,x) - v(t_1,x)|\,dx - \int |u(t_2,x) - v(t_2,x)|\,dx \geq 0$$

thereby showing the decay in time of the norm $\|u - v\|_{L^1}$, and thus, given that u and v agree at initial time, uniqueness.

The above outline shows that the classical argument for uniqueness of the entropic solution is based upon manipulations of integrals using smooth test functions, duplication of the variables and related techniques. Such arguments are of a similar type, and therefore compatible with, the techniques of proof developed in these notes, the latter being essentially based upon a regularization by convolution using a kernel $\rho_\varepsilon(\cdot - x)$. This suggests that, when one reinstates the dependency upon x through fields that are possibly not regular, and starting from initial conditions likewise, one may thus envision a possible proof of uniqueness consisting first in establishing the estimates of the classical argument on a regularized version of the equation, and letting next the regularization parameter vanish to obtain uniqueness in the generalized case, in the spirit of what has been completed both for the transport and the parabolic equations discussed above. For instance, in the case $\mathbf{F}(x,u) = f(u)\mathbf{b}(x)$, if we assume (at least) an initial condition u_0 bounded, a continuous flux f and a $W^{1,1}$ transport field $\mathbf{b}$, uniqueness should hold and can be expected. For such questions, along with related questions in the general nonlinear case, we refer to the lectures [73].

Bibliography

[1] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, 2nd ed., Pure Appl. Math. 140, Elsevier, Amsterdam, 2003.
[2] S. Agmon, A. Douglis and L. Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. I, *Comm. Pure Appl. Math.* **12** (1959), 623–727.
[3] S. Agmon, A. Douglis and L. Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. II, *Comm. Pure Appl. Math.* **17** (1964), 35–92.
[4] A. Alvino, G. Trombetti and P.-L. Lions, Comparison results for elliptic and parabolic equations via Schwarz symmetrization, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **7** (1990), no. 2, 37–65.
[5] L. Ambrosio, Transport equation and Cauchy problem for *BV* vector fields, *Invent. Math.* **158** (2004), no. 2, 227–260.
[6] L. Ambrosio, The flow associated to weakly differentiable vector fields: Recent results and open problems, in: *Nonlinear Conservation Laws and Applications*, IMA Vol. Math. Appl. 153, Springer, New York (2011), 181–193.
[7] L. Ambrosio and G. 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*, Lect. Notes Unione Mat. Ital. 5, Springer, Berlin (2008), 3–57.
[8] L. Ambrosio and A. Figalli, On flows associated to Sobolev vector fields in Wiener spaces: An approach à la DiPerna–Lions, *J. Funct. Anal.* **256** (2009), no. 1, 179–214.
[9] L. Ambrosio, M. Lecumberry and S. Maniglia, Lipschitz regularity and approximate differentiability of the DiPerna–Lions flow, *Rend. Semin. Fis. Mat. Padova* **114** (2005), 29–50.
[10] D. G. Aronson, Bounds for the fundamental solution of a parabolic equation, *Bull. Amer. Math. Soc.* **73** (1967), 890–896.
[11] S. Attanasio and F. Flandoli, Zero-noise solutions of linear transport equations without uniqueness: An example, *C. R. Math. Acad. Sci. Paris* **347** (2009), no. 13–14, 753–756.
[12] S. Attanasio and F. Flandoli, Renormalized solutions for stochastic transport equations and the regularization by bilinear multiplication noise, *Comm. Partial Differential Equations* **36** (2011), no. 8, 1455–1474.
[13] G. Barles and P.-L. Lions, Remarques sur les problèmes de réflexion oblique, *C. R. Acad. Sci. Paris Sér. I Math.* **320** (1995), no. 1, 69–74.
[14] J. Bergh and J. Löfström, *Interpolation Spaces. An Introduction*, Grundlehren Math. Wiss. 223, Springer, Berlin, 1976.
[15] D. Blanchard and F. Murat, Renormalised solutions of nonlinear parabolic problems with L^1 data: Existence and uniqueness, *Proc. Roy. Soc. Edinburgh Sect. A* **127** (1997), no. 6, 1137–1152.
[16] D. Blanchard, F. Murat and H. Redwane, Existence and uniqueness of a renormalized solution for a fairly general class of nonlinear parabolic problems, *J. Differential Equations* **177** (2001), no. 2, 331–374.
[17] V. I. Bogachev, G. Da Prato and M. Röckner, On parabolic equations for measures, *Comm. Partial Differential Equations* **33** (2008), no. 1–3, 397–418.
[18] V. I. Bogachev, G. D. Prato and M. Röckner, Uniqueness for solutions of Fokker–Planck equations on infinite dimensional spaces, *Comm. Partial Differential Equations* **36** (2011), no. 6, 925–939.

https://doi.org/10.1515/9783110635508-005

[19] V. I. Bogachev, M. Röckner and S. V. Shaposhnikov, On uniqueness problems related to the Fokker–Planck–Kolmogorov equation for measures, *J. Math. Sci. (N.Y.)* **179** (2011), no. 1, 7–47.

[20] V. I. Bogachev, M. Röckner and S. V. Shaposhnikov, On uniqueness of solutions to the Cauchy problem for degenerate Fokker–Planck–Kolmogorov equations, *J. Evol. Equ.* **13** (2013), no. 3, 577–593.

[21] A. Bohun, F. Bouchut and G. Crippa, Lagrangian flows for vector fields with anisotropic regularity, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **33** (2016), no. 6, 1409–1429.

[22] A. Bohun, F. Bouchut and G. Crippa, Lagrangian solutions to the 2D Euler system with L^1 vorticity and infinite energy, *Nonlinear Anal.* **132** (2016), 160–172.

[23] A. Bohun, F. Bouchut and G. Crippa, Lagrangian solutions to the Vlasov–Poisson system with L^1 density, *J. Differential Equations* **260** (2016), no. 4, 3576–3597.

[24] F. Bouchut and G. Crippa, Uniqueness, renormalization, and smooth approximations for linear transport equations, *SIAM J. Math. Anal.* **38** (2006), no. 4, 1316–1328.

[25] F. Bouchut and G. Crippa, Lagrangian flows for vector fields with gradient given by a singular integral, *J. Hyperbolic Differ. Equ.* **10** (2013), no. 2, 235–282.

[26] F. Boyer, Trace theorems and spatial continuity properties for the solutions of the transport equation, *Differential Integral Equations* **18** (2005), no. 8, 891–934.

[27] F. Boyer and P. Fabrie, *Mathematical Tools for the Study of the Incompressible Navier–Stokes Equations and Related Models*, Appl. Math. Sci. 183, Springer, New York, 2013.

[28] L. Caffarelli, R. Kohn and L. Nirenberg, First order interpolation inequalities with weights, *Compos. Math.* **53** (1984), no. 3, 259–275.

[29] 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** (1996), no. 2, 689–703.

[30] E. A. Carlen and M. Loss, Optimal smoothing and decay estimates for viscously damped conservation laws, with applications to the 2-D Navier–Stokes equation. A celebration of John F. Nash, Jr., *Duke Math. J.* **81** (1995), no. 1, 135–157.

[31] F. Catrina and Z.-Q. Wang, On the Caffarelli–Kohn–Nirenberg inequalities: Sharp constants, existence (and nonexistence), and symmetry of extremal functions, *Comm. Pure Appl. Math.* **54** (2001), no. 2, 229–258.

[32] N. Champagnat and P.-E. Jabin, Well posedness in any dimension for Hamiltonian flows with non BV force terms, *Comm. Partial Differential Equations* **35** (2010), no. 5, 786–816.

[33] N. Champagnat and P.-E. Jabin, Strong solutions to stochastic differential equations with rough coefficients, *Ann. Probab.* **46** (2018), no. 3, 1498–1541.

[34] A. S. Cherny and H.-J. Engelbert, *Singular Stochastic Differential Equations*, Lecture Notes in Math. 1858, Springer, Berlin, 2005.

[35] M. G. Crandall, H. Ishii and P.-L. Lions, User's guide to viscosity solutions of second-order partial differential equations, *Bull. Amer. Math. Soc. New Ser.* **27** (1992), no. 1, 1–67.

[36] G. Crippa, *The flow associated to weakly differentiable vector fields*, Tesi Sc. Norm. Super. Pisa (N. S.) 12, Edizioni della Normale, Pisa, 2009.

[37] G. Crippa and C. De Lellis, Estimates and regularity results for the DiPerna–Lions flow, *J. Reine Angew. Math.* **616** (2008), 15–46.

[38] G. Crippa, C. Donadello and L. V. Spinolo, Initial-boundary value problems for continuity equations with *BV* coefficients, *J. Math. Pures Appl. (9)* **102** (2014), no. 1, 79–98.

[39] D. Cruz-Uribe and A. Fiorenza, $L \log L$ results for the maximal operator in variable L^p spaces, *Trans. Amer. Math. Soc.* **361** (2009), no. 5, 2631–2647.

[40] R. J. DiPerna and P.-L. Lions, On the Fokker–Planck–Boltzmann equation, *Comm. Math. Phys.* **120** (1988), no. 1, 1–23.

[41] R. J. DiPerna and P.-L. Lions, Ordinary differential equations, transport theory and Sobolev spaces, *Invent. Math.* **98** (1989), no. 3, 511–547.
[42] H. Doss, Liens entre équations différentielles stochastiques et ordinaires, *C. R. Acad. Sci. Paris Sér. A-B* **283** (1976), no. 13, A939–A942.
[43] H. Doss, Liens entre équations différentielles stochastiques et ordinaires, *Ann. Inst. H. Poincaré Sect. B (N. S.)* **13** (1977), no. 2, 99–125.
[44] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*, Stud. Adv. Math., CRC Press, Boca Raton, 1992.
[45] E. Fedrizzi and F. Flandoli, Noise prevents singularities in linear transport equations, *J. Funct. Anal.* **264** (2013), no. 6, 1329–1354.
[46] A. Figalli, Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients, *J. Funct. Anal.* **254** (2008), no. 1, 109–153.
[47] F. Flandoli, The interaction between noise and transport mechanisms in PDEs, *Milan J. Math.* **79** (2011), no. 2, 543–560.
[48] F. Flandoli, M. Gubinelli and E. Priola, Well-posedness of the transport equation by stochastic perturbation, *Invent. Math.* **180** (2010), no. 1, 1–53.
[49] E. Fournié, J.-M. Lasry, J. Lebuchoux and P.-L. Lions, Applications of Malliavin calculus to Monte-Carlo methods in finance. II, *Finance Stoch.* **5** (2001), no. 2, 201–236.
[50] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions and N. Touzi, Applications of Malliavin calculus to Monte Carlo methods in finance, *Finance Stoch.* **3** (1999), no. 4, 391–412.
[51] A. Friedman, *Stochastic Differential Equations and Applications. Vol. 2*, Probab. Math. Statist. 28, Academic Press, New York, 1976,
[52] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 2nd ed., Grundlehren Math. Wiss. 224, Springer, Berlin, 1983.
[53] M. Hauray and C. Le Bris, A new proof of the uniqueness of the flow for ordinary differential equations with *BV* vector fields, *Ann. Mat. Pura Appl. (4)* **190** (2011), no. 1, 91–103.
[54] M. Hauray, C. Le Bris and P.-L. Lions, Deux remarques sur les flots généralisés d'équations différentielles ordinaires, *C. R. Math. Acad. Sci. Paris* **344** (2007), no. 12, 759–764.
[55] L. Hörmander, *The Analysis of Linear Partial Differential Operators. II*, Classics Math., Springer, Berlin, 2005.
[56] L. Hörmander, *The Analysis of Linear Partial Differential Operators. III*, Classics Math., Springer, Berlin, 2007.
[57] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, 2nd ed., North-Holland Math. Libr. 24, North-Holland, Amsterdam, 1989.
[58] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Grad. Texts in Math. 113, Springer, New York, 1988.
[59] N. V. Krylov, *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*, Grad. Stud. Math. 12, American Mathematical Society, Providence, 1996.
[60] N. V. Krylov, *Lectures on Elliptic and Parabolic Equations in Sobolev Spaces*, Grad. Stud. Math. 96, American Mathematical Society, Providence, 2008.
[61] H. Kunita, Stochastic differential equations and stochastic flows of diffeomorphisms, in: *École d'été de probabilités de Saint-Flour, XII—1982*, Lecture Notes in Math. 1097, Springer, Berlin (1984), 143–303.
[62] H. Kunita, *Stochastic Flows and Stochastic Differential Equations*, Cambridge Stud. Adv. Math. 24, Cambridge University Press, Cambridge, 1997.
[63] O. A. Ladyženskaja, V. A. Solonnikov and N. N. Ural'ceva, *Linear and Quasilinear Equations of Parabolic Type*, Transl. Math. Monogr. 23, American Mathematical Society, Providence, 1968.
[64] C. Le Bris and P.-L. Lions, Renormalized solutions of some transport equations with partially $W^{1,1}$ velocities and applications, *Ann. Mat. Pura Appl. (4)* **183** (2004), no. 1, 97–130.

[65] 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** (2008), no. 7–9, 1272–1317.
[66] N. Lerner, Transport equations with partially *BV* velocities, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **3** (2004), no. 4, 681–703.
[67] P.-L. Lions, A remark on some elliptic second-order problems, *Boll. Un. Mat. Ital. A (5)* **17** (1980), no. 2, 267–270.
[68] P.-L. Lions, Remarques sur les équations linéaires elliptiques du second ordre sous forme divergence dans les domaines non bornés, *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)* **78** (1985), no. 5, 205–212.
[69] P.-L. Lions, Remarques sur les équations linéaires elliptiques du second ordre sous forme divergence dans les domaines non bornés. II, *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)* **79** (1985), no. 6, 178–183.
[70] P.-L. Lions, *Mathematical Topics in Fluid Mechanics. Vol. 1*, Oxford Lecture Ser. Math. Appl. 3, The Clarendon Press, New York, 1996.
[71] P.-L. Lions, Sur les équations différentielles ordinaires et les équations de transport, *C. R. Acad. Sci. Paris Sér. I Math.* **326** (1998), no. 7, 833–838.
[72] P.-L. Lions, Equations et systèmes paraboliques: quelques questions nouvelles, Lectures at Collège de France 2012/13, www.college-de-france.fr/site/pierre-louis-lions.
[73] P.-L. Lions, Sur les lois de conservation scalaires, Lectures at Collège de France 2017/18, www.college-de-france.fr/site/pierre-louis-lions.
[74] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, Mod. Birkhäuser Class., Birkhäuser/Springer, Basel, 1995.
[75] A. Lunardi, *Interpolation Theory*, 2nd ed., Appunti. Sc. Norm. Super. Pisa (N. S.), Edizioni della Normale, Pisa, 2009.
[76] D. Luo, Well-posedness of Fokker–Planck type equations on the Wiener space, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **13** (2010), no. 2, 273–304.
[77] D. J. Luo, Fokker–Planck type equations with Sobolev diffusion coefficients and BV drift coefficients, *Acta Math. Sin. (Engl. Ser.)* **29** (2013), no. 2, 303–314.
[78] J. Malý and L. Pick, An elementary proof of sharp Sobolev embeddings, *Proc. Amer. Math. Soc.* **130** (2002), no. 2, 555–563.
[79] P. B. A. Mucha, Transport equation: Extension of classical results for div $b \in \mathrm{BMO}$, *J. Differential Equations* **249** (2010), no. 8, 1871–1883.
[80] A. Novotny and I. Straškraba, *Introduction to the Mathematical Theory of Compressible Flow*, Oxford Lecture Ser. Math. Appl. 27, Oxford University Press, Oxford, 2004.
[81] R. O'Neil, Convolution operators and $L(p, q)$ spaces, *Duke Math. J.* **30** (1963), 129–142.
[82] H. Osada, Diffusion processes with generators of generalized divergence form, *J. Math. Kyoto Univ.* **27** (1987), no. 4, 597–619.
[83] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3rd ed., Grundlehren Math. Wiss. 293, Springer, Berlin, 1999.
[84] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes, and Martingales. Vol. 2*, Cambridge Math. Lib., Cambridge University, Cambridge, 2000.
[85] L. P. Rothschild and E. M. Stein, Hypoelliptic differential operators and nilpotent groups, *Acta Math.* **137** (1976), no. 3–4, 247–320.
[86] E. M. Stein, *Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton Math. Ser. 43, Princeton University, Princeton, 1993.
[87] D. W. Stroock and S. R. S. Varadhan, *Multidimensional Diffusion Processes*, Grundlehren Math. Wiss. 233, Springer, Berlin, 1979.

[88] H. J. Sussmann, An interpretation of stochastic differential equations as ordinary differential equations which depend on the sample point, *Bull. Amer. Math. Soc.* **83** (1977), no. 2, 296–298.
[89] H. J. Sussmann, On the gap between deterministic and stochastic ordinary differential equations, *Ann. Probability* **6** (1978), no. 1, 19–41.
[90] L. Tartar, *An Introduction to Sobolev Spaces and Interpolation Spaces*, Lect. Notes Unione Mat. Ital. 3, Springer, Berlin, 2007.
[91] W. P. Ziemer, *Weakly Differentiable Functions*, Grad. Texts in Math. 120, Springer, New York, 1989.

Index

https://doi.org/10.1515/9783110635508-006

De Gruyter Series in Applied and Numerical Mathematics

Volume 3
Dominic Breit, Eduard Feireisl, Martina Hofmanová
Stochastically Forced Compressible Fluid Flows
ISBN 978-3-11-049050-3, e-ISBN (PDF) 978-3-11-049255-2,
e-ISBN (EPUB) 978-3-11-049076-3

Volume 2
Zahari Zlatev, Ivan Dimov, István Faragó, Ágnes Havasi
Richardson Extrapolation. Practical Aspects and Applications
ISBN 978-3-11-051649-4, e-ISBN (PDF) 978-3-11-053300-2,
e-ISBN (EPUB) 978-3-11-053198-5

Volume 1
Anvarbek Meirmanov, Oleg V. Galtsev, Reshat N. Zimin
Free Boundaries in Rock Mechanics
ISBN 978-3-11-054490-9, e-ISBN (PDF) 978-3-11-054616-3,
e-ISBN (EPUB) 978-3-11-054504-3

www.degruyter.com